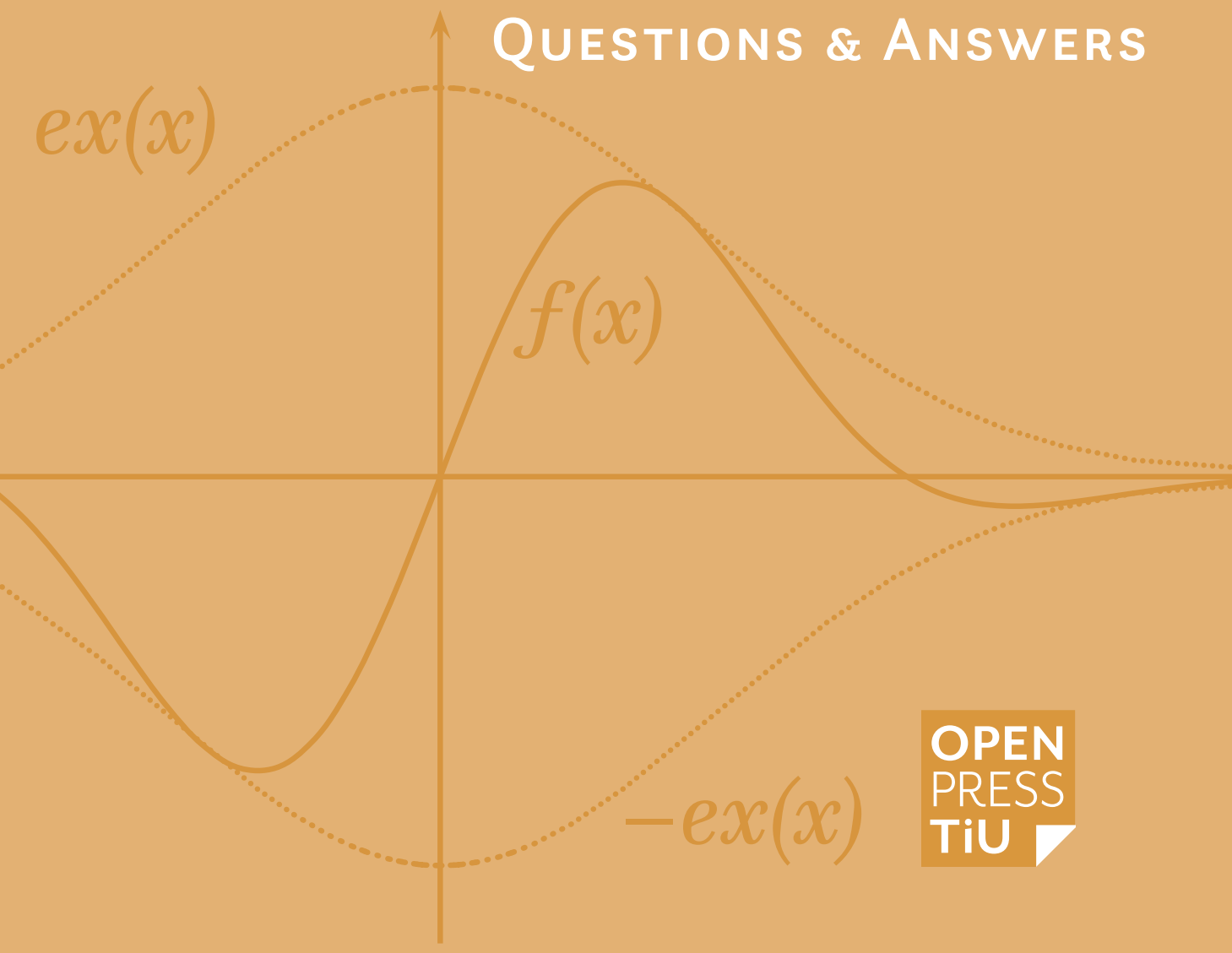
Jacob Engwerda

Introduction to Optimization for Economists

Questions & Answers



OPEN PRESS TiU

QUESTIONS AND ANSWERS INTRODUCTION TO OPTIMIZATION FOR ECONOMISTS

Questions and Answers Introduction to Optimization for Economists

Jacob Engwerda

1st edition

2026

OPEN PRESS TiU

Open Press Tilburg University
Tilburg, The Netherlands

© 2026 Jacob Engwerda

Design and typesetting: LINE UP boek en media bv
Cover design: Bas Ekkers

ISBN: 9789403866727
DOI: https://doi.org/10.56675/9789403866727



This is an Open Access book published under the terms of the Creative Commons 4.0 International license (CC BY NC ND 4.0). The full licence terms are available at creativecommons.org/licenses/by-nc-nd/4.0/legalcode.

Contents

Prologue VII

Preface IX

Notation and Symbols XI

1 Static Optimization: The Scalar Case 1
- 1.1 Functions of one variable 1
- 1.2 Compact Sets 2
- 1.3 Continuous Functions 2
- 1.4 Differentiable Functions 3
- 1.5 First-Order Criterium for an Optimum of a Scalar Function 4
- 1.6 Monotonicity of Functions 4
- 1.7 Convexity and Second-Order Criteria for an Optimum of a Scalar Function 7
- 1.8 Sexi Functions 8
- Mixed Exercises 9
- Advanced, Deepening and Broadening Exercises 17
- Problems with an Economic Content 36

2 Static Optimization: The Multi-Variable Case 43
- 2.1 Functions of more than one variable 43
- 2.2 Distance, Bounded Sets and Compact Sets 44
- 2.3 Continuous Functions 45
- 2.4 Differentiable Functions 46
- 2.5 First-Order Condition for an Optimum of a Multi-variable Function 48
- 2.6 Convexity and Second-Order Condition for an Optimum of Multi-Variable Functions 48
- Mixed Exercises 55
- Advanced, Deepening and Broadening Exercises 65
- Problems with an Economic Content 89

3 Constrained Optimization: Equality Constraints 101
- 3.1 The Substitution Method 101
- 3.2 Two variables and one constraint: a Geometric Approach 104
- 3.3 The Envelope Theorem and Lagrange Multiplier 110
- 3.4 A Second-Order Criterion for a Constrained Optimum 112
- 3.5 The Method of Lagrange: the general case 116

Mixed Exercises 119
Advanced, Deepening and Broadening Exercises 132
Problems with an Economic Content 143

4 Constrained Optimization: Inequality Constraints 157
4.1 The Directional Derivative 157
4.2 Two variables and one inequality constraint 158
4.3 The general case 161
4.4 Some pathological examples 168
4.5 Mixed constraints: a shortcut 173
4.6 Solving the problem without explicit constraint qualifications 176
4.7 A geometric approach to solve the problem 180
Mixed Exercises 181
Advanced, Deepening and Broadening Exercises 200
Problems with an Economic Content 222

5 Constrained Optimization: The Convex and Linear Case 231
5.1 Convex sets and their relationship with convex functions 231
5.2 Convex Programming and Saddle-Point Equilibria 232
5.3 Linear Programming and its Dual Problem 234
Mixed Exercises 238
Advanced, Deepening and Broadening Exercises 247
Problems with an economic content 266

6 Optimization problems with more than one decisionmaker 277
6.1 Full cooperation: the Pareto frontier 277
6.1.1 Sufficient Conditions for Pareto Efficiency 277
6.1.2 Necessary Conditions for Pareto Efficiency 281
6.1.3 An Algorithm to find Pareto Solutions 281
6.2 Noncooperative Solution 285
6.3 Hierarchical Solution 286
Mixed Exercises 295
Advanced, Deepening and Broadening Exercises 312
Problems with an Economic Content 325

References 342

Index 345

Prologue

The book *An Introduction to Optimization for Economists* originates from a crashcourse *refresher in mathematics* that I taught for many years at Tilburg University to research master and starting PhD students economics, prior to my retirement in 2019.
After my retirement, the lecture notes from this course were used as the foundation for writing this book. Additional theory was inorporated, and to make the material accessible at different levels, numerous exercises were added.

In this accompanying volume to the book, answers to all exercises are provided. A disclaimer applies here. Many exercises have not been checked yet by students. Though I did my best in elaborating them accurately, probably still typo's or other inaccuracies will be there. Readers are encouraged to report them on the site mentioned below, so that other users may benefit from this later on.

My personal hope is that students around the world will enhance their understanding through the mathematical approach to optimization presented in this book. In particular, I believe the book can serve as a useful stepping stone towards the study of more specialized areas in optimization.

I would like to thank Tilburg University, and especially Tilburg University Press, for making it possible to publish this book as open access. I am particularly grateful to Beatriz Lourenço Barrocas Neves Ferreira for her support throughout the publishing process.
Last but not least, I would like to thank my wife, Carine, for giving me the space and freedom to devote so much of my time writing this book and conducting research.

If you have any comments on the book, you are most welcome to post and share them on the DOI-site: https://doi.org/10.56675/9789403860336.

Writing this book has been a rewarding way to spend part of my retirement. If you feel inclined to show your appreciation, you may consider a small donation to the bank account below. I will use this money then to make a hiking tour with Carine. You may use next Wise bank account for that purpose: Jacob Christiaan Engwerda, IBAN: BE21967500746803, Belgium, Swift/BIC: TRWIBEB1XXX.

Hoping you may benefit from the book, Jacob Engwerda.

Preface

Economics is a social science that studies the production, distribution, and consumption of goods and services. A central concern of the discipline is how economic agents behave and interact, and how economies function as a whole. In line with this focus, a fundamental distinction in economics is made between microeconomics and macroeconomics.

Microeconomics analyzes the behavior of individual components of the economy, such as households and firms, buyers and sellers, as well as markets and their interactions. Macroeconomics, in contrast, examines the economy as a whole and addresses issues such as unemployment, inflation, economic growth, and monetary and fiscal policy.

To support such analyses, economists frequently rely on mathematical tools to model specific economic situations. By using mathematical abstractions of real-world phenomena, one aims to gain deeper insight into the underlying mechanisms at work. Ultimately, this approach is intended to improve the allocation of scarce resources.

In the book *An Introduction to Optimization for Economists*, we review a number of standard mathematical tools that are used to determine optima of functions. Accordingly, the focus lies on the second aspect of the analysis described above: given a specific mathematical model of an economic problem, how can one derive conclusions that—provided the model captures the essential features of the problem—lead to a better understanding of agent behavior and a more efficient allocation of resources.

Beginning with the concept of a function, the main text presents conditions under which optima can be derived. Both single-variable and multivariable functions are considered. Since resources are typically scarce, we then study conditions for finding optima when the domain of a function is restricted, leading to so-called constrained optimization problems. A special and important subclass of these problems consists of convex optimization problems. Owing to their favorable numerical properties—particularly in the case of linear programming—they are widely used in applications and are therefore treated in a separate section. Finally, we consider optimization problems in which the outcome is influenced by the actions of more than one agent.

This book is intended for students who wish to learn mathematical techniques that they can later apply to their own economic problems. The main text is accessible to students in economics as well as those pursuing more technical fields of study. Mathematical technicalities are deliberately kept to a minimum. The goal is to provide students with a basic understanding of the underlying ideas and an intuitive sense of the conditions under which specific mathematical tools can be applied and how they are used.

Each chapter concludes with a set of exercises, which are divided into three categories. The first category consists of exercises that allow students to test their understanding of the presented material. The second category encourages deeper engagement with mathematical rigor, including proofs of

theorems introduced in the main text and related topics. The third category aims to illustrate how the theoretical tools can be applied in economics to gain better insight into real-world phenomena.

In this accompanying volume, we provide solutions to all exercises in the book.

Notation and Symbols

$f(x) = O(x)$ as $x \to 0$	there exist $c,\ \epsilon$ such that $\|f(x)\| \leq c\|x\|$ when $0 < \|x\| \leq \epsilon$.
$\mathbb{R}$	Real numbers
$\mathbb{R}_+$	strictly positive real numbers
$\mathbb{R}^n$	set of vectors with n real entries
$\mathbb{R}^{n \times m}$	set of $n \times m$ matrices with real entries
$f^{-1}(x)$	inverse function of $f(x)$
$f^{(n)}(x)$	n^{th} order derivative of $f(x)$
x^T	transpose of the vector x
$\|x\|$	length $\sqrt{x^T x}$ of vector x, or, distance vector x to the origin
$\frac{\partial f}{\partial x_i}$, or, $D_i f(x)$, or $f'_{x_i}(x)$	partial derivative of f w.r.t. x_i
$D_u f(x)$	directional derivative $f(x)$ in the direction u
$Df(x) = [D_1 f(x) \cdots D_n(x)]$	row matrix of partial derivatives (or, Jacobian of $f(x)$)
$f'(x)$	the derivative of $f(x)$ (which equals $Df(x)$ if it exists)
C^1	set of differentiable functions which partial derivatives are continuous
C^k	set of functions which all k^{th}-order partial derivatives are continuous
$f"(x)$, or, Hessian $H(x)$	second order derivative of $f(x)$, i.e. $[(D'_1 f(x))^T \cdots (D'_n f(x))^T]^T$
A^T	transpose of matrix A
$A > 0$	positive definite matrix A
$A \geq 0$	positive semi-definite matrix A
$A < 0$	negative definite matrix A
$A \leq 0$	negative semi-definite matrix A
det A	determinant of matrix A
A^{-1}	inverse of matrix A
$I_{n \times n}$	identity matrix in $\mathbb{R}^{n \times n}$
$T_S(a)$	tangent cone to S at a
$TL_S(a)$	set of linearized feasible directions to S at a

$\subset$	subset
$\cup$	union
$\cap$	intersection
$\forall$	for all
$\exists$	there exists
$\mathbb{R}^n/\{0\}$	all $x \in \mathbb{R}^n$ except vector 0
$x \geq y$	componentwise inequality between vectors x and y
$f_D(\lambda)$	function associated with dual problem
$\mathcal{A}$	unit simplex
$S \backslash T$	all elements from S not belonging to T
$\bar{x}_{-i}$	$\bar{x}$, except entry i which is deleted
$\bar{x}_{\backslash i}$	$\bar{x}$, except entry i which is arbitrary
$\mathcal{R}(x_L)$	best response set follower against x_L
Λ	nonnegative orthant

CHAPTER 1

Static Optimization: The Scalar Case

1.1 Functions of one variable

EXERCISE 1.1 Determine for the rules f, defined below, a set D as large as possible such that f defines a function from $D \to \mathbb{R}$. Calculate the range of the corresponding function.

a. $f(x) = x + 1$. b. $f(x) = \frac{1}{x}$. c. $f(x) = \frac{1}{(x-1)^2}$. d. $f(x) = \{x \mid x^2 = y, \text{ where } y \in [1, 4]\}$.

ANSWER 1.1

a. $D = \mathbb{R}$; $f(D) = \mathbb{R}$.

b. $D = \mathbb{R} \setminus \{0\}$; $f(D) = D$.

c. $D = \mathbb{R} \setminus \{1\}$; $f(D) = \{y \mid y > 0\}$.

d. $D = [1, 2]$; $f(D) = [1, 4]$.

EXERCISE 1.2 Determine which of the functions f, defined below, is invertible. For all functions that are invertible, determine the inverse function f^{-1}.

a. $f(x) = x + 1$. b. $f(x) = x^2,\ x \in [-1, 1]$. c. $f(x) = -x^2 + 1,\ x \in [0, 1]$.

ANSWER 1.2

a. invertible; $y = x + 1 \to x = y - 1 \to f^{-1}(x) = x - 1$.

b. not invertible.

c. invertible: $y = -x^2 + 1 \to x^2 = 1 - y \to x = \sqrt{1 - y} \to f^{-1}(x) = \sqrt{1 - x}$.

EXERCISE 1.3

a. Simplify $\frac{^a\log(x)}{^a\log(y) + ^a\log(z)}$.

b. Simplify $e^{2\ln(x)} - \ln(e^{2x})$.

c. Solve $2^{x^2} = 4^{-3x-4}$.

d. Solve $\left(\frac{1}{4}\right)^{x^2} = 64\left(2^{-2x}\right)^4$.

e. Solve $2\ln(x - \frac{1}{2}e) - \ln(x - \frac{3}{4}e) = 1$.

ANSWER 1.3

a. $\frac{^a\log(x)}{^a\log(yz)}$.

b. $e^{2\ln(x)} - \ln(e^{2x}) = e^{\ln(x^2)} - \ln(e^{2x}) = x^2 - 2x$.

c. $2^{x^2} = 4^{-3x-4} \to 2^{x^2} = 2^{2(-3x-4)} \to x^2 = 2(-3x-4) \to x^2 + 6x + 8 = 0 \to x = -2,\ x = -4$.

d. $\left(\frac{1}{4}\right)^{x^2} = 64\left(2^{-2x}\right)^4 \to 2^{-2x^2} = 2^6 2^{-8x} \to 2^{-2x^2} = 2^{6-8x} \to -2x^2 = 6 - 8x \to x = 1,\ x = 3.$

e. $2\ln(x - \frac{1}{2}e) - \ln(x - \frac{3}{4}e) = 1 \to \ln(x - \frac{1}{2}e)^2 - \ln(x - \frac{3}{4}e) = \ln(e) \to \ln\frac{(x-\frac{1}{2}e)^2}{x-\frac{3}{4}e} = \ln(e) \to \frac{(x-\frac{1}{2}e)^2}{x-\frac{3}{4}e} = e \to (x - \frac{1}{2}e)^2 = e(x - \frac{3}{4}e) \to x^2 - 2ex + e^2 = 0 \to x = e.$

EXERCISE 1.4 Let $f(x) = x^2 + 1$ and $g(x) = \log(x^2 + 1)$. Determine

a. $(f + g)(x)$. b. $(fg)(x)$. c. $\left(\frac{f}{g}\right)(x)$. d. $(f \circ g)(x)$. e. $(g \circ f)(x)$.

ANSWER 1.4

a. $(f + g)(x) = x^2 + 1 + \log(x^2 + 1)$.

b. $(fg)(x) = (x^2 + 1)\log(x^2 + 1)$.

c. $\left(\frac{f}{g}\right)(x) = \frac{x^2+1}{\log(x^2+1)}$.

d. $(f \circ g)(x) = \log^2(x^2 + 1) + 1$.

e. $(g \circ f)(x) = \log((x^2 + 1)^2 + 1)$.

1.2 Compact Sets

EXERCISE 1.5 Determine the boundary and accumulation points of next sets. Motivate which of these sets is bounded, open, closed or compact, respectively.

a. $[1, 2] \cup [3, \infty)$. b. $\{1, 2\}$. c. $[1, 2) \cup (2, 3]$. d. $[1, 3) \cap (0, 2]$.

ANSWER 1.5

a. Boundary points: 1,2,3. Accumulation points: 1,2,3. Not bounded, not open, closed, not compact.

b. Boundary points: 1,2. Accumulation points: none. Bounded, not open, closed, compact.

c. Boundary points: 1,3. Accumulation points: 1,2,3. Bounded, not open, not closed, not compact.

d. $[1, 3) \cap (0, 2] = [1, 2] \to$ Boundary points: 1,2. Accumulation points: 1,2. Bounded, not open, closed, compact.

1.3 Continuous Functions

EXERCISE 1.6 Motivate which of the next functions is continuous.

a. $f(x) = \begin{cases} \log(x^2 + 1), & x \leq 0 \\ x, & x > 0 \end{cases}$.

b. $f(x) = xe^{x^2 + \ln(x^2+1)}$.

c. $f(x) = \begin{cases} \frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$.

ANSWER 1.6

a. $g(x) = x^2 + 1 \geq 1$ is continuous; $h(y) = \log(y)$ is continuous $\rightarrow h(g(x))$ is continuous; $k(x) = x$ is continuous. Notice $k(0) = h(g(0)) \rightarrow f$ is continuous.

b. $g(x) = x^2 + 1 \geq 1$ is continuous; $h(y) = \ln(y)$ is continuous $\rightarrow h(g(x))$ is continuous $\rightarrow z(x) = x^2 + h(g(x))$ is continuous; $k(y) = e^y$ is continuous $\rightarrow f(x) = xk(z(x))$ is continuous.

c. Let $g(x) = \frac{1}{x},\ x \neq 0$. Consider the sequence $t_k = \frac{1}{k}$. Then, t_k converges to zero if k goes to infinity. Furthermore, $g(t_k) = k$ goes to infinity. So, $g(t_k)$ does not converge to zero. So, g is not continuous.

EXERCISE 1.7 Motivate whether the next functions have an optimum.

a. $f(x) = \frac{1}{x},\ x > 0$.

b. $f(x) = \frac{x^2 - e^x}{\ln(x^2+2)},\ -10 \leq x \leq 12$.

ANSWER 1.7

a. The range of f is $(0, \infty)$. Furthermore $f(x) < f(y)$ if $x > y$. So, f has no optimum.

b. $D := [-10, 12]$ is compact. f is continuous on D. So, by Weierstrass' theorem, f has both a global maximum and a global minimum on D.

1.4 Differentiable Functions

EXERCISE 1.8 Determine the derivative of the next functions $f(x)$.

a. $f(x) = x^2 + {}^2\log(x),\ x > 0$.

b. $f(x) = \frac{2}{x^5} + 2^x,\ x \neq 0$.

c. $f(x) = \frac{\ln(x)}{x+1},\ x > 0$.

d. $f(x) = x^2 e^{x^2} + 1$.

e. $f(x) = \log(x^2+1) + \sqrt{x^4 + 1}$.

ANSWER 1.8

a. $f'(x) = 2x + \frac{1}{\ln 2}\frac{1}{x}$.

b. $f'(x) = -\frac{10}{x^6} + 2^x \ln 2$.

c. $f'(x) = \frac{\frac{1}{x}(x+1) - \ln(x)}{(x+1)^2}$.

d. $f'(x) = 2xe^{x^2} + 2x^3 e^{x^2}$.

e. $f'(x) = \frac{1}{\ln 10}\frac{2x}{x^2+1} + \frac{2x^3}{\sqrt{x^4+1}}$.

EXERCISE 1.9 Motivate which of the next functions is differentiable.

a. $f(x) = \frac{x^2+e^x}{\ln(x^2+2)}$.

b. $f(x) = \begin{cases} x, & x \leq 0 \\ x+1, & x > 0 \end{cases}$.

c. $f(x) = x + |x|$.

d. $f(x) = \begin{cases} x, & x \leq 0 \\ 2x, & x > 0 \end{cases}$.

e. $f(x) = x|x| + \log(e^x + 4)$.

ANSWER 1.9

a. $g(x) = x^2 + 2 \geq 2$ is differentiable; $h(y) = \ln(y)$ is differentiable $\rightarrow h(g(x))$ is differentiable, moreover $h(g(x)) \neq 0$; $k(x) = x^2 + e^x$ is differentiable $\rightarrow f = \frac{k(x)}{h(g(x))}$ is differentiable.

b. Let $g(x) := x$ and $h(x) := x + 1$. Then $g(0) = 0 \neq 1 = h(0)$. So, f is not continuous at $x = 0$. Therefore f is not differentiable at $x = 0$.

c. $f(x) = 2x$ if $x \geq 0$ and $f(x) = 0$ if $x \leq 0$. So, the tangent line to the graph of f at $x = 0$ is not uniquely determined. Therefore, f is not differentiable at $x = 0$.

d. Similar to item c. it follows that the tangent line to the graph of f at $x = 0$ is not uniquely determined. Therefore, f is not differentiable at $x = 0$.

e. $f(x) = x^2 + \log(e^x + 4)$ is differentiable if $x > 0$ and $f(x) = -x^2 + \log(e^x + 4)$ is differentiable if $x < 0$. Let $g_+(x) = x^2$ and $g_-(x) = -x^2$. Then, both g_+ and g_- are differentiable at $x = 0$ with derivative $0 \to h(x) = x|x|$ is differentiable at $x = 0$. Also, $\log(e^x + 4)$ is differentiable at $x = 0 \to f$ is differentiable at $x = 0$ too.

1.5 First-Order Criterium for an Optimum of a Scalar Function

EXERCISE 1.10 Determine the stationary points of the next functions.

a. $f(x) = xe^x$. b. $f(x) = \frac{x}{x^2+1}$. c. $f(x) = \ln(x^2 + 2x + 10)$. d. $f(x) = \frac{e^{\frac{1}{2}x^2}}{x}$, $x \neq 0$.

ANSWER 1.10

a. $f'(x) = (x + 1)e^x = 0 \to x = -1$.

b. $f'(x) = \frac{x^2+1-2x^2}{(x^2+1)^2} = 0 \to x^2 = 1 \to x = 1,\ x = -1$

c. $f'(x) = \frac{2x+2}{x^2+2x+10} = 0 \to x = -1$.

d. $f'(x) = \frac{x^2 e^{\frac{1}{2}x^2} - e^{\frac{1}{2}x^2}}{x^2} = 0 \to x^2 = 1 \to x = 1,\ x = -1$

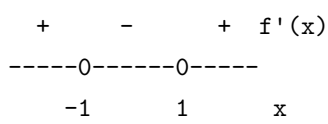
1.6 Monotonicity of Functions

EXERCISE 1.11 Determine the intervals where the next functions $f(x)$ are decreasing and increasing, respectively.

a. $f(x) = x^3 - 3x$. b. $f(x) = e^{x^2}$. c. $f(x) = \frac{1}{x^2}$, $x \neq 0$. d. $f(x) = \frac{e^x}{x^2}$, $x \neq 0$.

ANSWER 1.11

a. $f'(x) = 3x^2 - 3 = 0 \to x = 1,\ x = -1$.



$\to f$ is monotonically increasing on $(-\infty, -1] \cup [1, \infty)$;
f is monotonically decreasing on $[-1, 1]$.

b. $f'(x) = 2xe^{x^2} = 0 \to x = 0$.

```
  -         +     f'(x)
-------0--------
       0          x
```

$\rightarrow f$ is monotonically decreasing on $(-\infty, 0]$;
f is monotonically increasing on $[0, \infty)$.

c. $f'(x) = \frac{-2}{x^3},\ x \neq 0.$

```
  +         -     f'(x)
-------*--------
       0          x
```

$\rightarrow f$ is monotonically increasing on $(-\infty, 0)$;
f is monotonically decreasing on $(0, \infty)$.

d. $f'(x) = \frac{x^2e^x - 2xe^x}{x^4} = \frac{x(x-2)e^x}{x^4} = 0, \rightarrow x = 2.$

```
  +         -         +    f'(x)
-------*-------0-------
       0        2          x
```

$\rightarrow f$ is monotonically increasing on $(-\infty, 0) \cup [2, \infty)$;
f is monotonically decreasing on $(0, 2]$.

EXERCISE 1.12 Determine the optima of the next functions, using Theorem 1.34.

a. $f(x) = x^3 - 3x^2 + 3x + 1.$

b. $f(x) = x^7 - \frac{7}{5}x^5.$

c. $f(x) = \begin{cases} -(x+1)^2, & x \leq -1; \\ 0, & x \in (-1, 0); \\ x^2, & x \geq 0. \end{cases}$

ANSWER 1.12

a. $f'(x) = 3x^2 - 6x + 3 = 0 \rightarrow x = 1.$

```
  +         +     f'(x)
-------0--------
       1          x
```

$\rightarrow f$ has a saddle point at $x = 1$.

b. $f'(x) = 7x^6 - 7x^4 = 0 \rightarrow 7x^4(x^2 - 1) = 0 \rightarrow x = -1,\ x = 0,\ x = 1.$

```
  +         -         -         +    f'(x)
-------0--------0--------0-------
      -1         0         1          x
```

$\rightarrow f$ has a maximum at $x = -1$, a saddle point at $x = 0$ and a minimum at $x = 1$. As $f(100.000) > f(-1)$, the maximum is not global; as $f(-100.000) < f(1)$, the minimum is not global.

c. Notice that f is differentiable for all $x \neq \{-1, 0\}$. Furthermore, $g(x) = -(x+1)^2$ is differentiable at $x = -1$ with $g^{'}(-1) = 0$ and $h(x) = 0$ is differentiable at $x = -1$ with $h^{'}(-1) = 0$. As $g(-1) = h(-1)$ it follows that f is differentiable at $x = -1$ too. Similarly it follows that f is also differentiable at $x = 0$.
$\rightarrow f^{'}(x) = -2(x+1),\ x \leq -1$; $f^{'}(x) = 0,\ x \in (0,1)$ and $f^{'}(x) = 2x,\ x \geq 0$.

```
  +         0         +    f'(x)
-------0-------0-------
       -1        0          x
```

$\rightarrow f$ has a local maximum at $x \in [-1, 0)$ and a local minimum at $x \in (-1, 0]$.

EXERCISE 1.13 Determine the optima of the next functions $f(x)$. Indicate whether the optima are local or global optima.

a. $f(x) = x^3 - 3x + 1$.
b. $f(x) = 3x^4 - 4x^3 + 5$.
c. $f(x) = xe^{-\frac{1}{2}x^2},\ -2 \leq x \leq 2$.
d. $f(x) = \frac{x}{1+x^2}$.

ANSWER 1.13

a. $f^{'}(x) = 3x^2 - 3 \rightarrow$ Stationary points: $x = 1,\ x = -1$.

```
  +         -         +    f'(x)
-------0-------0-------
       -1        1          x
```

$\rightarrow f$ has a local maximum at $x = -1$ and a local minimum at $x = 1$. As $f(100.000) > f(-1)$, the maximum is not global; as $f(-100.000) < f(1)$, the minimum is not global.

b. $f^{'}(x) = 12x^3 - 12x^2 = 12x^2(x-1) \rightarrow$ Stationary points: $x = 0,\ x = 1$.

```
  -         -         +    f'(x)
-------0-------0-------
       0         1          x
```

$\rightarrow f$ has a saddle point at $x = 0$ and a global minimum at $x = 1$.

c. $f^{'}(x) = e^{-\frac{1}{2}x^2} - x^2e^{-\frac{1}{2}x^2} = (1-x^2)e^{-\frac{1}{2}x^2} \rightarrow$ Stationary points: $x = 1,\ x = -1$.

```
     -         +         -    f'(x)
|-------0-------0-------|
-2      -1        1        2   x
```

$\rightarrow f$ has a local minimum at $x = -1$ and $x = 2$; a local maximum at $x = -2$ and $x = 1$.
Note that the domain is compact and f is continuous. Therefore, by Weierstrass' theorem, f attains both a global maximum and a global minimum. Since $f(-1) < f(2)$, f has a global minimum at $x = -1$. Similarly, since $f(-2) < f(1)$, f has a global maximum at $x = 1$.

d. $f^{'}(x) = \frac{1-x^2}{(1+x^2)^2} = 0 \rightarrow x^2 = 1 \rightarrow x = 1,\ x = -1$

```
  -       +         -   f'(x)
------0-------0-------
     -1       1         x
```

$\rightarrow f$ has a local minimum at $x = -1$ and a local maximum at $x = 1$. Note $f(x) \rightarrow 0$ if $x \rightarrow \pm\infty$, $f(1) > 0$ and $f(-1) < 0$. Consequently, f has a global minimum at $x = -1$ and a global maximum at $x = 1$.

1.7 Convexity and Second-Order Criteria for an Optimum of a Scalar Function

EXERCISE 1.14 Indicate the intervals where the next functions are convex and concave, respectively.
a. $f(x) = e^x$. b. $f(x) = \ln(x),\ x > 0$. c. $f(x) = e^{-\frac{1}{2}x^2}$. d. $f(x) = e^x - \ln(x),\ x > 0$.

ANSWER 1.14
a. $f^{'}(x) = e^x \rightarrow f^{"}(x) = e^x$. $f^{"}(x) > 0 \rightarrow f$ is convex on $(-\infty, \infty)$.
b. $f^{'}(x) = \frac{1}{x} \rightarrow f^{"}(x) = \frac{-1}{x^2}$. $f^{"}(x) < 0 \rightarrow f$ is concave on $(0, \infty)$.
c. $f^{'}(x) = -xe^{-\frac{1}{2}x^2} \rightarrow f^{"}(x) = (x^2 - 1)e^{-\frac{1}{2}x^2} \rightarrow f^{"}(x) = 0$ at $x = 1$ and $x = -1$

```
  +       -         +   f"(x)
------0-------0-------
     -1       1         x
```

$\rightarrow f$ is convex on $(-\infty, -1] \cup [1, \infty)$ and f is concave on $[-1, 1]$.
d. $f^{'}(x) = e^x - \frac{1}{x} \rightarrow f^{"}(x) = e^x + \frac{1}{x^2} \rightarrow f^{"}(x) > 0 \rightarrow f$ is convex on $(0, \infty)$.

EXERCISE 1.15 Verify the correctness of the result stated in Remark 1.44 by checking the advertized conditions for the examples considered in Example 1.43.

ANSWER 1.15
a. $f(x) = x^4 \rightarrow f^{'}(0) = f^{"}(0) = f^{(3)}(0) = 0$; $f^{(4)}(0) = 24 > 0 \rightarrow$ item i. of Remark 1.44 applies $\rightarrow f$ has local minimum at $x = 0$.
b. $f(x) = -x^4 \rightarrow f^{'}(0) = f^{"}(0) = f^{(3)}(0) = 0$; $f^{(4)}(0) = -24 < 0 \rightarrow$ item ii. of Remark 1.44 applies $\rightarrow f$ has local maximum at $x = 0$.
c. $f(x) = x^3 \rightarrow f^{'}(0) = f^{"}(0) = 0$; $f^{(3)}(0) = 6 \neq 0 \rightarrow$ item iii. of Remark 1.44 applies $\rightarrow f$ has saddle point at $x = 0$.

EXERCISE 1.16 Consider the function $f(x) = 1 + (x^2 - 1)^3$. Determine all extremal points of f and indicate whether they are local or global optima.

ANSWER 1.16

$f^{'}(x) = 6x(x^2-1)^2 \rightarrow$ Stationary points: $x = 0$, $x = 1$ and $x = -1$.

```
   -        +        -        +    f'(x)
 -----0-------0-------0------
      -1       0        1        x
```

$\rightarrow f$ has a local minimum at $x = -1$ and $x = 1$; a local maximum at $x = 0$. $f(-1) = f(1) \rightarrow f$ has a global minimum at $x = -1$ and $x = 1$; $f(100) > f(0) \rightarrow f$ has no global maximum.

EXERCISE 1.17 Show that the function $f(x) = x^8 - 4x^2$ has a local maximum at $x = 0$. Is this also a global maximum? Determine all extremal points of f on $[-1.1, 1.1]$.

ANSWER 1.17

a. $f^{'}(x) = 8x^7 - 8x \rightarrow f^{"}(x) = 56x^6 - 8 \rightarrow f^{'}(0) = 0$ and $\rightarrow f^{"}(0) = -8 < 0 \rightarrow f$ has local maximum at $x = 0$.
   $f(100) > f(0) \rightarrow f$ has not a global maximum at $x = 0$.

b. $f^{'}(x) = 8x(x^6 - 1) \rightarrow$ stationary points at $x = 0$, $x = -1$ and $x = 1$

```
        -        +        -        +     f'(x)
     |------0-------0-------0------|
  -1.1    -1        0        1      1.1   x
```

$\rightarrow f$ has a local minimum at $x = -1$ and $x = 1$; a local maximum at $x = 0$. $f(-1.1) = f(1.1) > f(0)$ $\rightarrow f$ has a global maximum at $x = -1.1$ and $x = 1.1$; $f(-1) = f(1) \rightarrow f$ has global minimum at $x = 1$ and $x = -1$.

1.8 Sexi Functions

EXERCISE 1.18 Consider the function $f(x) = x^2 \sin \frac{1}{x}$, $x \neq 0$. Let $f(0) = 0$.

a. Use the definition of differentiability to show that $f(x)$ is differentiable at $x = 0$.

b. Show that $f^{'}(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x},\ x \neq 0$.

c. Show that $f^{'}(x)$ is not continuous at $x = 0$.

ANSWER 1.18

a. From the definition of differentiability it follows that we have to show that next difference quotient has a limit if Δx approaches zero.

$$\frac{f(0+\Delta x) - f(0)}{\Delta x} = \frac{(\Delta x)^2 \sin \frac{1}{\Delta x} - 0}{\Delta x} = \Delta x \sin \frac{1}{\Delta x}.$$

Note that $|\Delta x \sin \frac{1}{\Delta x}| \leq |\Delta x|$. Therefore, $|\Delta x \sin \frac{1}{\Delta x}|$ approaches zero if $|\Delta x|$ approaches zero.

b. For $x \neq 0$, f is differentiable. Using the standard calculation rules we have $f^{'}(x) = 2x \sin \frac{1}{x} + x^2 \frac{-1}{x^2} \cos \frac{1}{x}$.

c. Note that if x approaches zero $2x \sin \frac{1}{x}$ approaches zero (see item a.). On the other hand $\cos \frac{1}{x}$ fluctuates between -1 and 1 if x approaches zero. So, $f^{'}(x)$ does not approach a limiting value if x approaches zero. Therefore $f^{'}(x)$ is not continuous at $x = 0$.

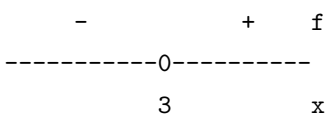
Mixed Exercises

EXERCISE 1.19 Determine for the functions f below all stationary points. Indicate whether f attains at these points a local or global minimum/maximum.

a. $f(x) = x^2 - 6x + 8$.
b. $f(x) = x^4 - 2x^2 + 2$.
c. $f(x) = \frac{1}{3}x^3 - \frac{1}{2}x^2 - 2x - 1$.
d. $f(x) = x^5 + 1$.
e. $f(x) = \frac{x}{e^x}$.
f. $f(x) = \frac{4x+3}{x^2+1}$.
g. $f(x) = \frac{e^x}{1+x^2}$.
h. $f(x) = \frac{2x^2+4}{\frac{1}{4}x^4+\frac{1}{2}x^2+1}$.

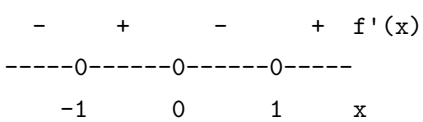
ANSWER 1.19

a. $f^{'}(x) = 2x - 6 \rightarrow$ Stationary point: $x = 3$.



$\rightarrow f$ has a global minimum at $x = 3$.

b. $f^{'}(x) = 4x^3 - 4x = 4x(x^2 - 1) \rightarrow$ Stationary points: $x = 0$, $x = 1$ and $x = -1$.

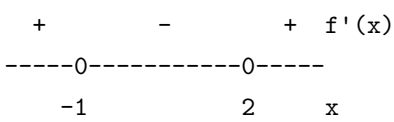


$\rightarrow f$ has a local minimum at $x = -1$ and $x = 1$; f has a local maximum at $x = 0$.

$f(100) > f(0) \rightarrow f$ has no global maximum.

$f(-1) = f(1) \rightarrow f$ has global minimum at $x = -1$ and $x = 1$.

c. $f^{'}(x) = x^2 - x - 2 = (x - 2)(x + 1) \rightarrow$ Stationary points: $x = -1$ and $x = 2$.



$\rightarrow f$ has a local maximum at $x = -1$; f has a local minimum at $x = 2$.

$f(100) > f(-1) \rightarrow f$ has no global maximum.

$f(-100) < f(2) \rightarrow f$ has no global minimum.

d. $f^{'}(x) = 5x^4 \rightarrow$ Stationary point: $x = 0$.

```
  +          -    f'(x)
--------0-------
         0         x
```

$\rightarrow f$ has a global maximum at $x = 0$.

e. $f'(x) = \frac{e^x - xe^x}{e^{2x}} \rightarrow$ Stationary point: $x = 1$.

```
  +         -     f'(x)
-------0------
        1         x
```

$\rightarrow f$ has a global maximum at $x = 1$.

f. $f'(x) = \frac{4(x^2+1)-2x(4x+3)}{(x^2+1)^2} = \frac{-4x^2-6x+4}{(x^2+1)^2} \rightarrow$ Stationary points: $x = -2$ and $x = \frac{1}{2}$.

```
  -       +          -    f'(x)
-----0--------0-------
    -2        0.5       x
```

$\rightarrow f$ has a local minimum at $x = -2$; f has a local maximum at $x = \frac{1}{2}$.

Note $f(x) \rightarrow 0$ if $x \rightarrow \pm\infty$, $f(-2) < 0$ and $f(\frac{1}{2}) > 0$. Consequently, f has a global minimum at $x = -2$ and a global maximum at $x = \frac{1}{2}$.

g. $f'(x) = \frac{e^x(x^2+1)-2xe^x}{(x^2+1)^2} = \frac{(x^2-2x+1)e^x}{(x^2+1)^2} \rightarrow$ Stationary point: $x = 1$.

```
    +           +     f'(x)
---------0----------
          1           x
```

$\rightarrow f$ has a saddle point at $x = 1$.

h. $f'(x) = \frac{4x(\frac{1}{4}x^4+\frac{1}{2}x^2+1)-(2x^2+4)(x^3+x)}{(\frac{1}{4}x^4+\frac{1}{2}x^2+1)^2} = \frac{-x^3(x^2+4)}{(\frac{1}{4}x^4+\frac{1}{2}x^2+1)^2} \rightarrow$ Stationary point: $x = 0$

```
      +             -     f'(x)
-----------0-----------
           0            x
```

$\rightarrow f$ has a global maximum at $x = 0$.

EXERCISE 1.20 Determine for the functions g below all stationary points. Indicate whether g attains at these points a local or global minimum/maximum.

a. $g(x) = \log(1 + x^2)$. b. $g(x) = e^{x^2+2x}$. c. $g(x) = \sqrt{\frac{1}{4}x^4 + x^2 + 1}$. d. $g(x) = xe^{-\frac{1}{2}x^2}$.

ANSWER 1.20

a. $g'(x) = \frac{2x}{\ln(10)(1+x^2)} \rightarrow$ Stationary point: $x = 0$.

```
      -             +     g'(x)
-----------0----------
           0            x
```

$\rightarrow g$ has a global minimum at $x = 0$.

b. $g^{'}(x) = (2x + 2)e^{x^2+2x} \rightarrow$ Stationary point: $x = -1$.

```
      -           +     g'(x)
----------0----------
          -1          x
```

$\rightarrow g$ has a global minimum at $x = 0$.

c. $g^{'}(x) = \frac{x^3+2x}{2\sqrt{\frac{1}{4}x^4+x^2+1}} = \frac{x(x^2+2)}{2\sqrt{\frac{1}{4}x^4+x^2+1}} \rightarrow$ Stationary point: $x = 0$

```
      -           +     g'(x)
----------0----------
          0           x
```

$\rightarrow g$ has a global minimum at $x = 0$.

d. $g^{'}(x) = e^{-\frac{1}{2}x^2} - x^2e^{-\frac{1}{2}x^2} = (1 - x^2)e^{-\frac{1}{2}x^2} \rightarrow$ Stationary points: $x = -1$ and $x = 1$.

```
 -          +          -   g'(x)
-----0-----------0-----
    -1           1      x
```

$\rightarrow g$ has a local minimum at $x = -1$; g has a local maximum at $x = 1$.

Note $g(x) \rightarrow 0$ if $x \rightarrow \pm\infty$, $g(1) > 0$ and $g(-1) < 0$. Consequently, g has a global minimum at $x = -1$ and a global maximum at $x = 1$.

EXERCISE 1.21 Determine for the functions h below all stationary points. Indicate whether h attains at these points a local or global minimum/maximum.

a. $h(x) = (e^{x^2} + 1)^2$. b. $h(x) = \frac{x\sqrt{x}}{e^{-\frac{3}{2}x+1}},\ x \geq 0$. c. $h(x) = (x^2 + 1)^{\frac{1}{3}}(x^2 + 2)^{\frac{2}{3}}$.

d. $h(x) = \ln(\frac{1}{1+x^2}) + \ln(\sqrt{1 + x^2})$.

ANSWER 1.21

a. $h^{'}(x) = 4x(e^{x^2} + 1)e^{x^2} \rightarrow$ Stationary point: $x = 0$.

```
      -           +     g'(x)
----------0----------
          0           x
```

$\rightarrow h$ has a global minimum at $x = 0$.

b. $h^{'}(x) = \frac{\frac{3}{2}x^{\frac{1}{2}}e^{-\frac{3}{2}x+1} + \frac{3}{2}e^{-\frac{3}{2}x+1}x^{\frac{3}{2}}}{(e^{-\frac{3}{2}x+1})^2} = \frac{\frac{3}{2}x^{\frac{1}{2}}e^{-\frac{3}{2}x+1}(1+x)}{(e^{-\frac{3}{2}x+1})^2} \rightarrow$ No stationary points

```
          +          g'(x)
|---------------------
0                    x
```

$\rightarrow h$ has a global minimum at $x = 0$.

c. $h^{'}(x) = \frac{1}{3}(x^2+1)^{-\frac{2}{3}}2x(x^2+2)^{\frac{2}{3}} + (x^2+1)^{\frac{1}{3}}\frac{2}{3}(x^2+2)^{-\frac{1}{3}}2x$
$= (x^2+1)^{-\frac{2}{3}}(x^2+2)^{-\frac{1}{3}}2x(\frac{1}{3}(x^2+2) + \frac{2}{3}(x^2+1)) \rightarrow$ Stationary point: $x = 0$.

```
      -           +     g'(x)
----------0----------
          0           x
```

$\rightarrow h$ has a global minimum at $x = 0$.

d. $h^{'}(x) = -\frac{2x}{1+x^2} + \frac{x}{1+x^2} = -\frac{x}{1+x^2} \rightarrow$ Stationary point: $x = 0$.

```
      +           -     g'(x)
----------0----------
          0           x
```

$\rightarrow h$ has a global maximum at $x = 0$.

EXERCISE 1.22 Determine for the functions f below all extremal points. Indicate whether f attains at these points a local or global minimum/maximum.

a. $f(x) = x^2 - 4x + 3,\ x \in [0, 5]$.
b. $f(x) = \frac{1}{3}x^3 - 2x^2 + 3x,\ x \in [0, 4]$.
c. $f(x) = \frac{1}{3}x^3 - x + 1,\ x \in [-2, 2)$.
d. $f(x) = \ln(1+x),\ x \in [0, 10]$.
e. $f(x) = \frac{x^2}{1-x},\ x \in [\frac{3}{2}, 5]$.
f. $f(x) = \frac{x^2}{1+x},\ x \in [-4, -1) \cup (-1, 4]$.
g. $f(x) = \frac{1}{1+x^2},\ x \in [-1, \infty)$.
h. $f(x) = \frac{x}{1+x^2},\ x \in [-\frac{1}{2}, \infty)$.

ANSWER 1.22

a. $f^{'}(x) = 2x - 4 \rightarrow$ Stationary point: $x = 2$.

```
      -           +       g'(x)
|----------0----------|
0           2          5   x
```

$\rightarrow f$ has a global minimum at $x = 2$.

$f(0) < f(5) \rightarrow f$ has a global maximum at $x = 5$ and a local maximum at $x = 0$.

b. $f^{'}(x) = x^2 - 4x + 3 \rightarrow$ Stationary points: $x = 2$.

```
   +       -       +       g'(x)
|-----0--------0-----|
0     1        3     4    x
```

$\rightarrow f$ has a local maximum at $x = 1$ and $x = 4$; f has a local minimum at $x = 0$ and $x = 3$.

$f(1) = f(4) \rightarrow f$ has a global maximum at $x = 1$ and $x = 4$;

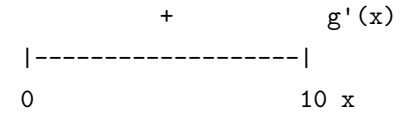
$f(0) = f(3) \rightarrow f$ has a global minimum at $x = 0$ and $x = 3$.

c. $f^{'}(x) = x^2 - 1 \rightarrow$ Stationary points: $x = 1$ and $x = -1$.

```
   +       -       +       g'(x)
|------0--------0-----)
-2     -1        1     2    x
```

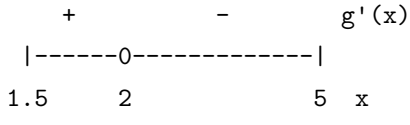
$\rightarrow f$ has a local maximum at $x = -1$; f has a local minimum at $x = -2$ and $x = 1$.
$f(-1) = f(2) \rightarrow f$ has a global maximum at $x = -1$;
$f(-2) = f(1) \rightarrow f$ has a global minimum at $x = -2$ and $x = 1$.

d. $f'(x) = \frac{1}{1+x} \rightarrow$ No stationary points.



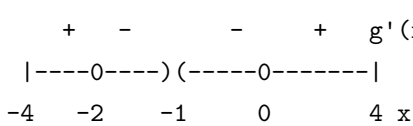
$\rightarrow f$ has a global minimum at $x = 0$; f has a global maximum at $x = 10$.

e. $f'(x) = \frac{2x(1-x)+x^2}{(1-x)^2} = \frac{x(2-x)}{(1-x)^2} \rightarrow$ Stationary points: $x = 2$.



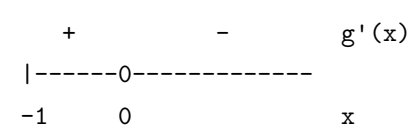
$\rightarrow f$ has a global maximum at $x = 2$; f has a local minimum at $x = \frac{3}{2}$ and $x = 5$.
$f(\frac{3}{2}) > f(5) \rightarrow f$ has a global minimum at $x = 5$.

f. $f'(x) = \frac{2x(1+x)-x^2}{(1+x)^2} = \frac{x(2+x)}{(1+x)^2} \rightarrow$ Stationary points: $x = -2$ and $x = 0$.



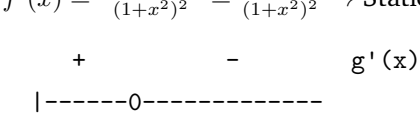
$\rightarrow f$ has a local maximum at $x = -2$ and $x = 4$; f has a local minimum at $x = -4$ and $x = 0$.
$f(-1.01) < f(0)$ and $f(-1.01) < f(-4) \rightarrow f$ has no global minimum.
$f(-0.99) > f(-2)$ and $f(-0.99) > f(4) \rightarrow f$ has no global maximum.

g. $f'(x) = -\frac{2x}{(1+x^2)^2} =\rightarrow$ Stationary points: $x = 0$.



$\rightarrow f$ has a global maximum at $x = 0$; f has a local minimum at $x = -1$.
$f(5) < f(-1) \rightarrow f$ has no global minimum.

h. $f'(x) = \frac{1+x^2-2x^2}{(1+x^2)^2} = \frac{1-x^2}{(1+x^2)^2} \rightarrow$ Stationary point: $x = 1$.



$\rightarrow f$ has a global maximum at $x = 1$; f has a local minimum at $x = -0.5$.
Note $f(x) \rightarrow 0$ if $x \rightarrow \infty$, $f(-0.5) < 0$. Consequently, f has a global minimum at $x = -0.5$.

EXERCISE 1.23 Solve the next equations.

a. $2 * 2^x + 4 * 2^x = 24$.

b. $2^{2x} + 2^{x+2} = 5$.

c. $\frac{1}{4} * 2^{2x} + 4^x + \frac{3}{4^{1-x}} = 32$.

d. $\ln(e^{x^2}) + e^{\ln(x)} = 2$.

e. $\ln(5x-7) - \ln(x-1) = \ln(x+1),\ x > \frac{7}{5}$.

f. $\ln(x^2 - 2x + 1) = \frac{2}{\ln(x-1)},\ x > 1$.

ANSWER 1.23

a. $2 * 2^x + 4 * 2^x = 24 \Leftrightarrow 6 * 2^x = 24 \Leftrightarrow 2^x = 4 \Leftrightarrow 2^x = 2^2 \Leftrightarrow x = 2$.

b. $2^{2x} + 2^{x+2} = 5 \Leftrightarrow (2^x)^2 + 4 * 2^x - 5 = 0$, or, with $y := 2^x$, $y^2 + 4y - 5 = 0 \Leftrightarrow y = -5$ or $y = 1 \Leftrightarrow 2^x = -5$ or $2^x = 1 \Leftrightarrow x = 0$.

c. $\frac{1}{4} * 2^{2x} + 4^x + \frac{3}{4^{1-x}} = 32 \Leftrightarrow 4^{-1} * 4^x + 4^x + \frac{3}{4} * 4^x = 32 \Leftrightarrow 2 * 4^x = 32 \Leftrightarrow 4^x = 16 \Leftrightarrow x = 2$

d. $\ln(e^{x^2}) + e^{\ln(x)} = 2 \Leftrightarrow x^2 + x = 2 \Leftrightarrow x = 1$ or $x = -2$.

e. $\ln(5x-7) - \ln(x-1) = \ln(x+1) \Leftrightarrow \ln\frac{5x-7}{x-1} = \ln(x+1) \Leftrightarrow \frac{5x-7}{x-1} = x+1 \Leftrightarrow 5x - 7 = x^2 - 1 \Leftrightarrow x^2 - 5x + 6 = 0 \Leftrightarrow x = 3$ or $x = 2$.

f. $\ln(x^2 - 2x + 1) = \frac{2}{\ln(x-1)} \Leftrightarrow \ln((x-1)^2) = \frac{2}{\ln(x-1)} \Leftrightarrow 2\ln(x-1) = \frac{2}{\ln(x-1)} \Leftrightarrow$ with $y := \ln(x-1)$, $2y^2 = 2 \Leftrightarrow y = 1$ or $y = -1 \Leftrightarrow \ln(x-1) = 1$ or $\ln(x-1) = -1 \Leftrightarrow x - 1 = e$ or $x - 1 = \frac{1}{e} \Leftrightarrow x = 1 + e$ or $x = 1 + \frac{1}{e}$

EXERCISE 1.24 Determine the boundary and accumulation points of next sets. Motivate whether these sets are

bounded/open/closed or compact, respectively

a. $(-1, 2] \cup (3, 4)$.

b. $[-1, 2) \cup (2, 4]$.

c. $[-1, 2] \cup [3, 4]$.

d. $(-1, 2) \cup (2, \infty)$.

ANSWER 1.24

a. Boundary points: 2

Accumulation points: -1, 2, 3, 4

Set is bounded: $|x| < 10$ for all $x \in S$.

Set is not open: every interval $(2 - \epsilon_1, 2 + \epsilon_2)$ contains an element not belonging to S ($\epsilon_i > 0$).

Set is not closed: every interval $(-1 - \epsilon_1, -1 + \epsilon_2)$ contains an element belonging to S ($\epsilon_i > 0$). Or, $t_k = -1 + \frac{1}{k} \to -1$, with $t_k \in S$ for all $k > 1$, whereas its limit $-1 \notin S$.

Set is not compact: it is not closed.

b. Boundary points: -1, 4

Accumulation points: -1, 2, 4

Set is bounded: $|x| < 10$ for all $x \in S$.

Set is not open: every interval $(-1 - \epsilon_1, -1 + \epsilon_2)$ contains an element not belonging to S ($\epsilon_i > 0$).

Set is not closed: every interval $(2 - \epsilon_1, 2 + \epsilon_2)$ contains an element belonging to S ($\epsilon_i > 0$). Or, $t_k = 2 + \frac{1}{k} \to 2$, with $t_k \in S$ for all $k > 1$, whereas its limit $2 \notin S$.

Set is not compact: it is not closed.

c. Boundary points and accumulation points: -1, 2, 3, 4

Set is bounded: $|x| < 10$ for all $x \in S$.

Set is not open: every interval $(-1-\epsilon_1, -1+\epsilon_2)$ contains an element not belonging to S ($\epsilon_i > 0$).
Set is closed: complement set is open. Proof: construct appropriate intervals for the cases: $x < -1$, $2 < x < 3$ and $x > 4$, respectively.
Set is compact: it is both closed and bounded.

d. Boundary points: none
Accumulation points: -1, 2
Set is not bounded: Assume there exists an m such that $|x| < m$ for all $x \in S$. Choose $x = m + 1$, clearly $x \in S$ and $|x| > m$.
Set is open: construct an appropriate interval for the cases $-1 < x < 2$ and $x > 2$, respectively.
Set is not closed: every interval $(2 - \epsilon_1, 2 + \epsilon_2)$ contains an element belonging to S ($\epsilon_i > 0$). Or, $t_k = 2 + \frac{1}{k} \to 2$, with $t_k \in S$ for all $k > 1$, whereas its limit $2 \notin S$.
Set is not compact: it is neither closed nor bounded.

EXERCISE 1.25 Determine the largest domain for next functions g where they are continuous, respectively differentiable. Motivate your answer.

a. $g(x) = x + \ln(1 + x^2)$.
b. $g(x) = |x + 1| + e^{x-x^2}$.
c. $g(x) = \frac{\sqrt{x-1}}{x}$.
d. $g(x) = \frac{\ln(1+x)}{x}$.
e. $g(x) = \begin{cases} \frac{x^2-1}{x+1}, & x \neq 1; \\ 0, & x = 1. \end{cases}$
f. $g(x) = \begin{cases} x^2 + 1, & x \leq 0; \\ |x + 1|, & x > 0. \end{cases}$

ANSWER 1.25

a. g is defined on $x \in \mathbb{R}$.
g is continuous on $x \in \mathbb{R}$ (both $g_1(x) = x$ and $g_2(x) = \ln(1 + x^2)$ are continuous on this set).
g is differentiable on $x \in \mathbb{R}$ (both $g_1(x) = x$ and $g_2(x) = \ln(1 + x^2)$ are differentiable on this set).

b. g is defined on $x \in \mathbb{R}$.
g is continuous on $x \in \mathbb{R}$ (both $g_1(x) = |x + 1|$ and $g_2(x) = e^{x-x^2}$ are continuous on this set).
g is differentiable on $x \in \mathbb{R}\backslash\{-1\}$ ($g_1(x) = |x+1|$ is differentiable for all $x \neq -1$ and $g_2(x) = e^{x-x^2}$ is differentiable everywhere).

c. g is defined on $x \geq 1$.
g is continuous on $x \geq 1$ (both $g_1(x) = \sqrt{x-1}$ and $g_2(x) = \frac{1}{x}$ are continuous on this set).
g is differentiable on $x > 1$ ($g_1(x) = \sqrt{x-1}$ is differentiable for all $x > 1$ and $g_2(x) = \frac{1}{x}$ is differentiable for all $x \geq 0$).

d. g is defined on $D := \{x \mid x > -1,\ x \neq 0\}$.
g is continuous on D ($g_1(x) = \ln x + 1$ is continuous for all $x > -1$ and $g_2(x) = \frac{1}{x}$ is continuous for all $x \neq 0$).
g is differentiable on D ($g_1(x) = \ln x + 1$ is differentiable for all $x > -1$ and $g_2(x) = \frac{1}{x}$ is differentiable for all $x \neq 0$).

e. g is defined on $D := \{x \mid x \neq -1\}$.
Note that at $x = 1$, $g_1(x) := \frac{x^2-1}{x+1}$ equals 0. So, g is continuous on D.
g is differentiable on D, since g_1 is differentiable on D.

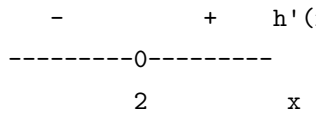
f. g is defined on $\mathbb{R}$.
Note that $g_1(x) := x^2 + 1$ is continuous for all $x \leq 0$ and $g_2(x) := |x + 1| = x + 1$, for all $x > 0$, is continuous. As $g_1(0) = g_2(0)$, g is continuous on $\mathbb{R}$.
g_1 is differentiable for all x, with $g_1^{'} = 2x$. g_2 is differentiable for all $x > -1$, with $g_2^{'} = 1$. So, $g_1^{'}(0) \neq g_2^{'}(0)$. Therefore, g is differentiable for all $x \neq 0$.

EXERCISE 1.26 Determine the intervals where the next functions h are decreasing and increasing, respectively.

a. $h(x) = x^2 - 4x + 4$.
b. $h(x) = -x^3 + 3x^2 - 3x + 1$.
c. $h(x) = x^4 - \frac{8}{3}x^3 + 2x^2 + 1$.
d. $h(x) = e^{-x^2+2x}$.
e. $h(x) = \ln(1 + x),\ x > 0$.

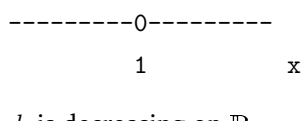
ANSWER 1.26

a. $h^{'}(x) = 2x - 4 \rightarrow h^{'}(x) = 0$ at $x = 2$



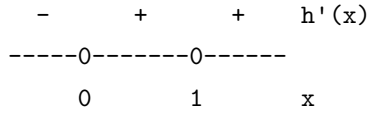
$\rightarrow h$ is increasing on $[2, \infty)$ and decreasing on $(-\infty, 2]$.

b. $h^{'}(x) = -3x^2 + 6x - 3 = -3(x - 1)^2 \rightarrow h^{'}(x) = 0$ at $x = 1$



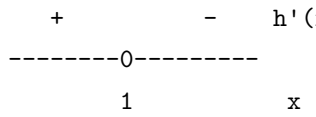
$\rightarrow h$ is decreasing on $\mathbb{R}$.

c. $h^{'}(x) = 4x^3 - 8x^2 + 4x = 4x(x - 1)^2 \rightarrow h^{'}(x) = 0$ at $x = 0$ and $x = 1$



$\rightarrow h$ is increasing on $[0, \infty)$ and decreasing on $(-\infty, 0]$.

d. $h^{'}(x) = (-2x + 2)e^{-x^2+2x} \rightarrow h^{'}(x) = 0$ at $x = 1$



$\rightarrow h$ is increasing on $(-\infty, 1]$ and decreasing on $[1, \infty)$.

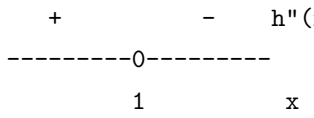
e. $h^{'}(x) = \frac{1}{1+x} \rightarrow h^{'}(x) > 0$ for all $x > 0 \rightarrow h$ is increasing on $(0, \infty)$.

EXERCISE 1.27 Determine for the functions h in Exercise 1.26 the intervals where they are convex and concave, respectively.

ANSWER 1.27

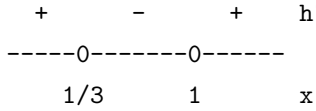
a. $h^{"}(x) = 2 \rightarrow h$ is convex everywhere.

b. $h^{"}(x) = -6x + 6 \rightarrow h^{"}(x) = 0$ at $x = 1$.



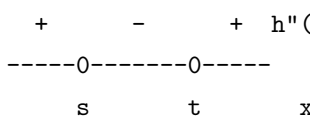
$\rightarrow h$ is convex on $(-\infty, 1]$ and concave on $[1, \infty)$.

c. $h^{"}(x) = 12x^2 - 16x + 4 \rightarrow h^{"}(x) = 0$ at $x = \frac{1}{3}$ and $x = 1$



$\rightarrow h$ is convex on $(-\infty, \frac{1}{3}]$ and $[1, \infty)$ and concave on $[\frac{1}{3}, 1]$.

d. $h^{"}(x) = (-2 + (-2x + 2)^2)e^{-x^2+2x} = 2(2x^2 - 4x + 1)e^{-x^2+2x} \rightarrow h^{"}(x) = 0$ at $s = \frac{4-\sqrt{8}}{2}$ and $t = \frac{4+\sqrt{8}}{2}$



$\rightarrow h$ is convex on $(-\infty, s]$ and $[t, \infty)$ and concave on $[s, t]$.

e. $h^{"}(x) = -\frac{1}{(1+x)^2} \rightarrow h^{"}(x) < 0$ for all $x > 0 \rightarrow h$ is concave for all $x > 0$.

Advanced, Deepening and Broadening Exercises

EXERCISE 1.28 Consider the intervals below in $\mathbb{R}$. Show that next statements are true.

a. The intervals (a, b), (a, ∞) and $(-\infty, b)$ are open.

b. The intervals $[a, b]$, $[a, \infty)$ and $(-\infty, b]$ are closed.

c. The intervals $[a, b)$ and $(a, b]$ are neither open or closed .

d. $\mathbb{R}$, is as well open as closed.

ANSWER 1.28

a. Let $x \in (a, b)$. Consider, e.g., $a_x = a$ and $b_x = b$. Then $(a_x, b_x) \subset (a, b)$.
Let $x \in (a, \infty)$. Consider, e.g., $a_x = a$ and $b_x = x + 1$. Then $(a, x, b_x) \subset (a, \infty)$.
Let $x \in (-\infty, b)$. Consider, e.g., $a_x = x - 1$ and $b_x = b$. Then $(a, x, b_x) \subset (-\infty, b)$.

b. The complement set of $[a, b]$ is $(-\infty, a) \cup (b, \infty)$, which is the union of two open sets (see item a.). By Proposition 1.10, the union of open sets is also open. A similar argument applies to the other intervals.

Another way to justify these statements is to show that the set of accumulation points of this interval coincides with its set of boundary points, namely $\{a, b\}$.
Consider the sequence $\{s_n\}_{n=1}^{\infty}$ defined by $s_n = a + \frac{b-a}{n}$. Then s_n converges to a, and both s_n and a lie in the interval $[a, b]$. Similarly, consider the sequence $\{t_n\}_{n=1}^{\infty}$ defined by $t_n = b + \frac{a-b}{n}$. Then t_n converges to b, and both t_n and b belong to $[a, b]$.
It is straightforward to verify that $[a, b]$ has no accumulation points other than $\{a, b\}$.

c. The interval $[a, b)$ is not open. Indeed, for any interval (a_x, b_x) containing a, there exists a point - such as $a - \frac{a-a_x}{2}$ - that lies in (a_x, b_x) but not in $[a, b)$.
To see that $[a, b)$ is not closed, consider its complement set $(-\infty, a) \cup [b, \infty)$. This set is not open; by an argument similar to the one above, one can show that $[b, \infty)$ is not open.

d. Let $x \in \mathbb{R}$. Define $a_x := x - 1$ and $b_x := x + 1$. Then $(a_x, b_x) \subset \mathbb{R}$, which shows that $\mathbb{R}$ is open.
Moreover, for any sequence that converges to a limit s, we have $s \in \mathbb{R}$. Therefore, if $\mathbb{R}$ has any boundary points, they must lie in $\mathbb{R}$.

EXERCISE 1.29 Which of the next statements are true. Motivate your answer.

a. Every differentiable function is continuous.
b. There exist continuous functions which are not differentiable.
c. Every continuous function attains both a global maximum and a global minimum on $[0, 2]$.
d. There exist continuous functions that attain both a global maximum and a global minimum on $(0, 2)$.
e. A bounded set which contains all its boundary points is compact.
f. Every strictly increasing function is convex.
g. Every boundary point is an accumulation point.

ANSWER 1.29

a. True, see paragraph above Theorem 1.23.
b. True, see Figure 1.12.
c. True, $[0, 2]$ is compact, so by Weierstrass' Theorem 1.19 this result holds
d. True, consider $f(x) = x,\ 0 < x < 1,\ f(x) = -3x + 4,\ 1 \leq x < \frac{3}{2},\ f(x) = x - 2, \frac{3}{2} \leq x < 2$. Then f has a global maximum at $x = 1$ and a global minimum at $x = \frac{3}{2}$.
e. False, by definition a boundary point always belongs to the set. $S = (0, 2]$ has one boundary point (which belongs to S), but S is not closed (it does not contain all its accumulation points, i.e., 0 is an accumulation point of S but does not belong to S).
f. False, consider $h(x) = \ln(1 + x)$ for $x > 0$ (see Exercises 26e,27e).
g. True. If b is a boundary point, by definition $b \in S$. So, every interval (b_l, b_r) containing b, contains $b \in S$ and, again by definition of b being a boundary point, a point outside S. That is, b satisfies the definition of accumulation point.

EXERCISE 1.30 *(uniqueness inverse function)*
Assume $f : D \to R$ is invertible. Show, using the definition, that the inverse function is uniquely determined. Hint: assume f has two inverse functions. Show they have to coincide.

ANSWER 1.30
Assume that both $g(y)$ and $h(y)$ are inverse functions of f. By definition, they satisfy

$$\text{i) } (fog)(y) = y, \ (foh)(y) = y, \text{ for all } y \in R, \text{ and}$$
$$\text{ii) } (gof)(x) = x, \ (hof)(x) = x, \text{ for all } x \in D.$$

Using i) and ii), we obtain

$$g(y) = g(foh(y)) = (gof)(h(y)) = h(y) \text{ for all } y \in R.$$

EXERCISE 1.31 The polynomial $p(x)$ of degree 3 has zeros at $x = 1$ and $x = 2$. Furthermore, $p(3) = 4$. Determine $p(x)$. Hint: use the property that if a polynomial $p_n(x)$ of degree n has a zero at $x = a$, $p_n(x) = (x - a)p_{n-1}(x)$, where p_{n-1} is a polynomial of degree $n - 1$.

ANSWER 1.31
$p(x) = (x - 1)(x - 2)(x - a)$. As $p(3) = 4$ it follows $4 = 2 * 1 * (3 - a) \to a = 1$.

EXERCISE 1.32 *(Limits and continuity)*
The intuition that a sequence $s_n, \ n = 1, 2, 3, \cdots$ approaches the number s (or: the limit of the sequence $\{s_n\}_{n=1}^{\infty}$ is s) is formal defined as: for every $\epsilon > 0$ (however small that may be) we can find a number N such that, for all $n > N$ the distance between s_n and s, that is $|s_n - s|$, is smaller than this ϵ.
Use this definition to show that
a. $f(x) = 1$ is continuous on $\mathbb{R}$.
b. $f(x) = x$ is continuous on $\mathbb{R}$.
c. $f(x)$ is discontinuous at $x = 1$ if $f(x) = 2$ for $x \in [0, 1)$ and $f(x) = 3$ for $x \in [1, 2)$.

ANSWER 1.32
a. Let $x_0 \in \mathbb{R}$ be arbitrary, and suppose that s_n approaches x_0. Let $\epsilon > 0$ be given. By definition, there exists an N_ϵ such that for all $n > N_\epsilon$, we have $|s_n - x_0| < \epsilon$.
Now consider the distance between $f(s_n)$ and $f(x_0)$. For all $n > N_\epsilon$, we have $|f(s_n) - f(x_0)| = |1 - 1| = 0$. Hence, $f(s_n)$ approaches $f(x_0)$ whenever s_n approaches x_0.
b. Let $x_0 \in \mathbb{R}$ be an arbitrary number, and assume s_n approaches x_0. Choose an arbitrary $\epsilon > 0$. By definition, we can find an N_ϵ such that for all $n > N_\epsilon$, $|s_n - x_0| < \epsilon$. Next consider the distance between $f(s_n)$ and $f(x_0)$. Note that for all $n > N_\epsilon$, $|f(s_n) - f(x_0)| = |s_n - x_0| < \epsilon$. So, whenever

the distance between s_n and x_0 becomes small, the distance between $f(s_n)$ and $f(x_0)$ becomes small too.

c. Consider $s_n = 1 - \frac{1}{n}$. Then s_n approaches 1. But, $f(s_n) = 2$ for all n, whereas $f(1) = 3$. So, $f(s_n)$ does not approach $f(1)$. Therefore f is discontinuous at 1.

EXERCISE 1.33 *(Differentiability implies continuity)*
Formal a function f is called differentiable at x_0 if the difference quotient at this point approaches some number if the change in the function argument, Δx, approaches zero. Show that if f is differentiable at x_0, this implies that $f(x_0 + \Delta x)$ approaches $f(x_0)$ if Δx approaches zero. Or, stated differently, f is continuous at x_0.

ANSWER 1.33
By definition, f is differentiable at x_0 if

$$\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

approaches some number s if Δx approaches zero. Equivalently,

$$\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} - s \text{ approaches zero if } \Delta x \text{ approaches zero,}$$

or

$$\frac{f(x_0 + \Delta x) - f(x_0) - s\Delta x}{\Delta x} \text{ approaches zero if } \Delta x \text{ approaches zero.}$$

Since the denumerator tends to zero as Δx approaches zero, this fraction can only converge to zero if the numerator also tends to zero. That is,

$$f(x_0 + \Delta x) - f(x_0) - s\Delta x \text{approaches zero if } \Delta x \text{ tends to zero.}$$

This implies that

$$f(x_0 + \Delta x) - f(x_0) \text{ must approach zero if } \Delta x \text{ approaches zero,}$$

which shows that f is continuous at x_0.

EXERCISE 1.34 *(First-order condition optimum)*
Show, using the definition of differentiability (see Exercise 1.32), that if f is differentiable on an open interval I, and f attains a maximum at $x_0 \in I$, $f'(x_0) = 0$. Hint, show this by contradiction. That is, assume either $f'(x_0) > 0$ or $f'(x_0) < 0$ and show that under those assumptions f cannot attain a maximum at x_0.

ANSWER 1.34

Assume, for example, that $f'(x_0) > 0$. Since f is differentiable, the difference quotient

$$\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

approaches the number $s := f'(x_0)$ as Δx approaches zero.

Now consider approaching x_0 from the right, i.e. $\Delta x > 0$. Then the difference quotient is close to s when Δx is sufficiently small. Since $s > 0$, there must exist some $\overline{\Delta x}$ such that for all $\Delta x \in (0, \overline{\Delta x})$,

$$\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} > 0.$$

This implies that

$$f(x_0 + \Delta x) - f(x_0) > 0 \text{ for all } \Delta x \in (0, \overline{\Delta x}),$$

or equivalently,

$$f(x) > f(x_0) \text{ for all } x \in (x_0, x_0 + \overline{\Delta x}).$$

Thus, f cannot have a maximum at x_0.

A similar argument shows that the assumption $f'(x_0) < 0$ leads to a contradiction when approaching x_0 from the left.

EXERCISE 1.35 *(Rolle's theorem)*

Consider the continuous function f on the compact interval $[a, b]$. Assume $f(x)$ is differentiable on the open interval (a, b) and $f(a) = f(b)$. Below you are asked to show that there exists a point $c \in (a, b)$ for which $f'(c) = 0$.

a. Motivate why f attains both a maximum M and a minimum m on $[a, b]$.

b. Assume $f(a) = M$ and $f(b) = m$. Show that there exists a point $c \in (a, b)$ for which $f'(c) = 0$.

c. Assume either M or m is not attained at an endpoint of the interval. Show that there exists a point $c \in (a, b)$ for which $f'(c) = 0$.

d. Assume $h(0) = h(1) = 0$ and $h(x)$ is differentiable on $(0, 1)$. Consider the function $f(x) = e^{-x^3 - x^2 - x} h(x)$ on $[0, 1]$. Motivate, using Rolle's Theorem, why there exists a point $c \in (0, 1)$ such that $f'(c) = 0$.

ANSWER 1.35

a. Follows by Weierstrass' Theorem.

b. As $f(a) = f(b) = M = m$ it follows that $f(x)$ is constant. So, $f'(x) = 0$ on (a, b).

c. Assume the maximum M is not attained at an endpoint. Then at some point $c \in (a, b)$, $f(c) = M$. As f is differentiable on (a, b), it follows that $f'(c) = 0$.

d. Note: $f(0) = f(1)$. Furthermore, f is differentiable on $(0,1)$ as both $e^{-x^3-x^2-x}$ and $h(x)$ are differentiable on $(0,1)$. So, the conditions of Rolle's theorem are satisfied and it follows by this theorem that there exists a point $c \in (0,1)$ such that $f^{'}(c) = 0$.

EXERCISE 1.36 *(Mean Value Theorem)*
Consider the continuous function f on the compact interval $[a,b]$. Assume $f(x)$ is differentiable on the open interval (a,b). Below, in item c., you are asked to show that there exists a point $c \in (a,b)$ for which $f^{'}(c) = \frac{f(b)-f(a)}{b-a}$.
a. Consider $g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a)$. Show that $g(a) = g(b)$.
b. Use Rolle's Theorem (Exercise 1.34) to conclude that there exists a point $c \in (a,b)$ for which $g^{'}(c) = 0$.
c. Show, using item b., that there exists a point $c \in (a,b)$ for which $f^{'}(c) = \frac{f(b)-f(a)}{b-a}$.
d. Assume $f(x)$ is continuous on $[0,1]$, differentiable on $(0,1)$ and $f^{'}(x) \in [-2,2]$. Show that $-2 + f(0) \leq f(1) \leq 2 + f(0)$.

ANSWER 1.36
a. $g(a) = f(a) - \frac{f(b)-f(a)}{b-a}(a-a) = f(a)$; $g(b) = f(b) - \frac{f(b)-f(a)}{b-a}(b-a) = f(a)$.
b. $g(a) = g(b)$ and g is differentiable on (a,b). So by Rolle's theorem there exists a $c \in (a,b)$ such that $g^{'}(c) = 0$.
c. Note $g^{'}(c) = f^{'}(c) - \frac{f(b)-f(a)}{b-a}$. The result follows then by item b.
d. f satisfies the conditions of the mean value theorem on $[0,1] \rightarrow$ there exists a $c \in (0,1)$ such that $f^{'}(c) = \frac{f(1)-f(0)}{1-0} = f(1) - f(0)$. So, $f(1) = f(0) + f^{'}(0)$. As $f^{'}(x) \in [-2,2]$ it follows that $-2 + f(0) \leq f(1) \leq 2 + f(0)$.

EXERCISE 1.37 *(Monotonicity and derivative)*
Assume f is differentiable on an open interval I. Show, using the definition of differentiability (see Exercise 1.32), that f is increasing on I if and only if $f^{'}(x) \geq 0$. Hint: use the mean value theorem (Exercise 1.35) to estimate the average change in f between any two points in I.

ANSWER 1.37
" $\Leftarrow$ " Let a and b be any two points in I with $a < b$. Consider $\frac{f(b)-f(a)}{b-a}$.
Note that f is continuous on $[a,b]$ and differentiable on (a,b). By the Mean Value Theorem, there exists a $c \in (a,b)$ such that

$$\frac{f(b) - f(a)}{b - a} = f^{'}(c).$$

By assumption, $f^{'}(c) \geq 0$. Since $b - a > 0$, it follows that $f(b) - f(a) \geq 0$, i.e., $f(b) \geq f(a)$.
" $\Rightarrow$ " Let $x_0 \in I$ and take $b > x_0$. Since f is monotonically increasing, $f(b) \geq f(x_0)$, and hence $\frac{f(b)-f(x_0)}{b-x_0} \geq 0$.
Now consider $b < x_0$. Then $f(b) \leq f(x_0)$, so again $\frac{f(b)-f(x_0)}{b-x_0} \geq 0$.

Thus,

$$\text{for every } \Delta x \neq 0, \ \frac{f(x_0+\Delta x)-f(x_0)}{\Delta x} \geq 0.$$

Since f is differentiable on I, it follows that when Δx approaches zero, $\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}$ approaches $f^{'}(x_0)$. Since for every Δx the difference quotient is nonnegative, we conclude that $f^{'}(x_0)$ must be nonnegative too.

EXERCISE 1.38 *(First-order characterization convex function)*
Assume f is differentiable on an open interval I. Below, in items c.and e., respectively, you are asked to show that f is convex on I if and only if $f(y) \geq f(x) + f^{'}(x)(y-x)$ for all $x, y \in I$.

a. Consider x_0 and $y = x_0 + \Delta x \in I$. Use the definition of convexity to conclude that $(1-\lambda)f(x_0+\Delta x) \geq f(x_0+(1-\lambda)\Delta x_0) - f(x_0) + (1-\lambda)f(x_0)$, for each $\lambda \in [0,1]$.

b. Let $t := 1-\lambda$. Show that we obtain the next inequality from item a.: $f(x_0+\Delta x) \geq f(x_0) + \frac{f(x_0+t\Delta x_0)-f(x_0)}{t\Delta x_0}\Delta x_0$.

c. Use the definition of differentiability (see Exercise 1.32) to show from item b., that if f is convex, $f(y) \geq f(x_0) + f^{'}(x_0)(y-x_0)$.

d. Assume $f(y) \geq f(x) + f^{'}(x)(y-x)$ for all $x, y \in I$. Let $z := \lambda x + (1-\lambda y)$. Show that (i) $f(x) \geq f(z) + f^{'}(z)(x-z)$ and (ii) $f(y) \geq f(z) + f^{'}(z)(y-z)$.

e. By multiplication of (i) by λ and (ii) by $(1-\lambda)$ in item d., and adding both resulting inequalities show that if $f(y) \geq f(x) + f^{'}(x)(y-x)$ for all $x, y \in I$, f is convex on I.

ANSWER 1.38

a. By the definition of convexity

$$f(\lambda x_0 + (1-\lambda)y) \leq \lambda f(x_0) + (1-\lambda)f(y).$$

If we take $y = x_0 + \Delta x$, this becomes

$$f(x_0+(1-\lambda)\Delta x) \leq \lambda f(x_0) + (1-\lambda)f(x_0+\Delta x_0) + f(x_0) - f(x_0).$$

Rearranging, we get

$$(1-\lambda)f(x_0+\Delta x) \geq f(x_0+(1-\lambda)\Delta x_0) - f(x_0) + (1-\lambda)f(x_0).$$

b. Setting $t = (1-\lambda)$, the inequality from item a. becomes:

$$tf(x_0+\Delta x) \geq f(x_0+t\Delta x_0) - f(x_0) + tf(x_0),$$

which can be rewritten as

$$f(x_0+\Delta x) \geq \frac{f(x_0+t\Delta x_0)-f(x_0)}{t} + f(x_0) = \frac{f(x_0+t\Delta x_0)-f(x_0)}{t\Delta x_0}\Delta x_0 + f(x_0).$$

c. Since f is differentiable on I,

$$\frac{f(x_0 + t\Delta x_0) - f(x_0)}{t\Delta x_0} \text{ approaches } f^{'}(x_0) \text{ if } t \text{ approaches zero.}$$

Because the inequality in item b. holds for all $t \in [0,1]$, we can in particular consider t approaching zero to obtain

$$f(x_0 + \Delta x) \geq f^{'}(x_0)\Delta x_0 + f(x_0)$$

or, equivalently, using $y = x_0 + \Delta x$, $f(y) \geq f(x_0) + f^{'}(x_0)(y - x_0)$.

d. For $x, y \in I$, also $z := \lambda x + (1 - \lambda y) \in I$. Using the definition of convexity with respect to the pairs (x, z) and (y, z), respectively, we obtain both inequalities.

e. From (i):

$$\lambda f(x) \geq \lambda f(z) + \lambda f^{'}(z)(x - z),$$

and from (ii):

$$(1 - \lambda)f(y) \geq (1 - \lambda)f(z) + (1 - \lambda)f^{'}(z)(y - z).$$

Adding these two inequalities gives:

$$\lambda f(x)+(1-\lambda)f(y) \geq \lambda f(z)+\lambda f^{'}(z)(x-z)+(1-\lambda)f(z)+(1-\lambda)f^{'}(z)(y-z) = f(z) = f(\lambda x+(1-\lambda)y).$$

EXERCISE 1.39 *(Taylor's theorem)*

Assume $f(x)$ is n times differentiable on the open interval I and $a \in I$. Below, in item d., you are asked to show that for any $x \in I$ there exists a point ξ located between x and a such that

$$f(x) = p_{n-1}(x) + \frac{f^{(n)}(\xi)}{n!}(x - a)^n, \tag{1.1}$$

where $p_{n-1}(x)$ is given by next polynomial of degree $n - 1$

$$p_{n-1}(x) := f(a) + \frac{f^{'}(a)}{1!}(x - a) + \frac{f^{"}(a)}{2!}(x - a)^2 + \frac{f^{(3)}(a)}{3!}(x - a)^3 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(x - a)^{(n-1)}.$$

a. Let $F(y) := f(x) - \left(f(y) + \frac{f^{'}(y)}{1!}(x - y) + \frac{f^{"}(y)}{2!}(x - y)^2 + \cdots + \frac{f^{(n-1)}(y)}{(n-1)!}(x - y)^{(n-1)} \right)$.
Consider $g(y) := F(y) - \left(\frac{x-y}{x-a}\right)^n F(a)$. Show that $g(a) = g(x) = 0$.

b. Use Rolle's Theorem (Exercise 1.34) to conclude that there exists a point ξ between x and a such that $F^{'}(\xi) + \frac{n(x-\xi)^{n-1}}{(x-a)^n}F(a) = 0$.

c. Show that $F^{'}(y) = -\frac{(x-y)^{n-1}}{(n-1)!}f^{(n)}(y)$.

d. Calculate $F(a)$ from item b. Next show that there exists a point ξ between x and a such that (1.1) holds.

ANSWER 1.39

a. From the definition of $g(y$ we have

$$g(a) = F(a) - \left(\frac{x-a}{x-a}\right)^n F(a) = F(a) - F(a) = 0,$$

and

$$g(x) = F(x) - \left(\frac{x-x}{x-a}\right)^n F(a) = F(x).$$

From the definition of $F(y)$ we have

$$\begin{aligned} F(x) &= f(x) - \left(f(x) + \frac{f^{'}(x)}{1!}(x-x) + \frac{f^{"}(x)}{2!}(x-x)^2 + \cdots + \frac{f^{(n-1)}(x)}{(n-1)!}(x-x)^{(n-1)}\right) \\ &= f(x) - f(x) = 0. \end{aligned}$$

Thus, $g(x) = 0$.

b. Without loss of generality, assume $x > a$. By item a., we have $g(a) = g(x)$. The function g is continuous on $[a, x]$ differentiable on (a, x). Therefore, by Rolle's theorem, there exists some $\xi \in (a, x)$ such that

$$g^{'}(\xi) = 0.$$

This implies that

$$F^{'}(\xi) + n\frac{(x-\xi)^{n-1}}{(x-a)^n}F(a) = 0.$$

c. By differentiating $F(y)$ we obtain

$$\begin{aligned} F^{'}(y) &= -(f^{'}(y) - f^{'}(y) + \frac{f^{"}(y)}{1!}(x-y) - 2\frac{f^{"}(y)}{2!}(x-y) + \cdots + \\ &\quad \frac{f^{(n-1)}(y)}{(n-2)!}(x-y)^{n-2} - (n-1)\frac{f^{(n-1)}(y)}{(n-1)!}(x-y)^{n-2} + \frac{(x-y)^{n-1}}{(n-1)!}f^{(n)}(y)) \\ &= -\frac{(x-y)^{n-1}}{(n-1)!}f^{(n)}(y). \end{aligned}$$

d. From item b, we obtain

$$F(a) = -\frac{F^{'}(\xi)}{n\frac{(x-\xi)^{n-1}}{(x-a)^n}} = -\frac{-\frac{(x-\xi)^{n-1}}{(n-1)!}f^{(n)}(\xi)}{n\frac{(x-\xi)^{n-1}}{(x-a)^n}} = \frac{f^{(n)}(\xi)}{n!}(x-a)^n.$$

Using the definition of $F(a)$ from item a, this yields $f(x) - p_{n-1}(x) = \frac{f^{(n)}(\xi)}{n!}(x-a)^n$.

EXERCISE 1.40 *(Brouwer's Fixed Point Theorem)*

Consider a continuous function f on the compact interval $[a, b]$. Assume $f(x) \in [a, b]$ for all $x \in [a, b]$. Then Brouwer's Fixed Point Theorem states that under these assumptions, there exists a point $c \in [a, b]$

for which $f(c) = c$. (Such a point is called a *fixed point* of f). Below you are asked to prove this theorem.

a. Consider the function $g(x) = f(x) - x$. Motivate that g is continuous.
b. Show that in case neither $f(a) = a$ or $f(b) = b$ there exists a point $c_1 \in [a, b]$ where $g(c_1) > 0$ and a point $c_2 \in [a, b]$ where $g(c_2) < 0$.
c. Use items a and b to conclude that there must exist a point $c \in [a, b]$ for which $f(c) = c$.
d. Consider the function $f(x) = e^{-x^3-x^2-x}$ on $[0, 1]$. Motivate why there exists a point $c \in [0, 1]$ such that $f(c) = c$.

ANSWER 1.40

a. The function $f(x)$ is continuous, and $f_2(x) = x$ is also continuous. Therefore, $g = f(x) - x$ is continuous.
b. We have
$$g(a) = f(a) - a > a - a = 0, \text{ and } g(b) = f(b) - b < b - b = 0.$$
c. From item b., either $f(a) = a$ or $f(b) = b$, or there exist points $c_1,\ c_2 \in [a, b]$ such that $g(c_1) > 0$ and $g(c_2) < 0$. Since g is continuous it follows in the latter case that there exists a point c between c_1 and c_2 such that $g(c) = 0$, that is, $f(c) = c$.
d. Note that $-x^3 - x^2 - x \leq 0$ for all $x \in [0, 1]$. Hence,
$$e^{-x^3-x^2-x} \in [0, 1] \text{ for all} x \in [0, 1].$$
The function f is continuous on the compact interval $[0, 1]$. Therefore, by Brouwer's Fixed Point Theorem, there exists a point $c \in [0, 1]$ such that $e^{-c^3-c^2-c} = c$.

EXERCISE 1.41 *(Fixed Point Contraction Mapping Theorem)*
Consider the function f on the closed interval I. Assume $f(x) \in I$ for all $x \in I$ and, moreover, there exists a constant $0 \leq c < 1$ such that for all $x, y \in I$, $|f(x) - f(y)| \leq c|x - y|$. Then the Contraction Mapping Theorem states that under these conditions f has a *unique* fixed point $\bar{x} \in I$. Below you are asked to show the correctness of this statement under the simplifying condition that f is continuous and the interval is compact.

a. Assume f is continuous and I is compact. Show that f has a fixed point on I.
b. Show that f has a unique fixed point on I. Hint: assume there is another one.
c. Let $x \in I$. Consider the sequence defined by: $s_1 = x$ and $s_{n+1} = f(s_n),\ n = 2, 3, ...$ Show that $|f(s_n) - s_n| \leq c^n|s_2 - s_1|$. (Using this, one can show that this implies that f must have a fixed point, i.e., that the Contraction Mapping Theorem applies also without assuming continuity of f and the assumption that the interval is compact.)
d. Consider the function $f(x) = \frac{1}{7}e^{-x^3-x^2-x}$ on $[0, 1]$. Show that f has a unique fixed point on $[0, 1]$. Hint: use the Mean value Theorem to estimate $f(x) - f(y)$.

e. Consider $f(x) = \frac{1}{x^2+1}$. Show that $f(x)$ has a unique fixed point. Hint: use the Mean value Theorem to estimate $f(x) - f(y)$.

ANSWER 1.41

a. Let $I = [a, b]$. Note that i.) $f(x)$ is continuous $[a, b]$; ii.) for all $x \in [a, b]$, we have $f(x) \in [a, b]$. Therefore, by Brouwer's Fixed Point Theorem, f has a fixed point in $[a, b]$.

b. Assume that f has two fixed points $\bar{x}_1$ and $\bar{x_2}$. Then,

$$|\bar{x}_1 - \bar{x}_2| = |f(\bar{x}_1) - f(\bar{x}_2)| \leq c|\bar{x}_1 - \bar{x}_2| < |\bar{x}_1 - \bar{x}_2|.$$

This is a contradiction. Therefore, the assumption is false, and it follows that f has a unique fixed point.

c. Since $s_{n+1} = f(s_n)$,

$$|f(s_n) - s_n| = |s_{n+1} - s_n| = |f(s_n) - f(s_{n-1})| \leq c|s_n - s_{n-1}|$$

Repeating this argument, we obtain

$$|s_{n+1} - s_n| \leq c|s_n - s_{n-1}| \leq c^2|s_{n-1} - s_{n-2}| \leq \cdots \leq c^n|s_2 - s_1|.$$

d. Note (see the previous exercise) that $f(x) \in [0, 1]$ for all $x \in [0, 1]$.
Furthermore, f is differentiable on $[0, 1]$. Therefore, by the Mean Value Theorem, there exists a $\xi \in (0, 1)$ such that

$$f(x) - f(y) = f^{'}(\xi)(x - y).$$

We have $f^{'}(x) = \frac{1}{7}(-3x^2 - 2x - 1)e^{-x^3-x^2-x}$, which implies that $|f^{'}(x)| < \frac{6}{7}$ for all $x \in (0, 1)$. Hence,

$$|f(x) - f(y)| = |f^{'}(\xi)||(x - y)| < \frac{6}{7}|(x - y)|.$$

Therefore, by the Contraction Mapping Theorem, f has a unique fixed point on $[0, 1]$.

e. The function f is differentiable on $\mathbb{R}$. Therefore, by the Mean value Theorem, there exists a $\xi \in (x, y)$ such that

$$f(x) - f(y) = f^{'}(\xi)(x - y).$$

We have $f^{'}(x) = \frac{-2x}{(1+x^2)^2}$. To determine its extrema, we compute the second derivative of f: $f^{"}(x) = \frac{6x^4+4x^2-2}{(1+x^2)^2}$.
The stationary points of $f^{'}$ satisfy $6x^4 + 4x^2 - 2 = 0$. Straightforward calculations show that at $x^2 = \frac{1}{3}$ $f^{'}$ has a stationary point. An elementary analysis of $f^{'}$ shows that $f^{'}(x)$ attains a global maximum at $x = -\sqrt{\frac{1}{3}}$ and a global minimum at $x = \sqrt{\frac{1}{3}}$. Moreover, $|f^{'}(-\sqrt{\frac{1}{3}})| = |f^{'}(\sqrt{\frac{1}{3}})|$. Hence,

$$|f^{'}(x)| = \frac{2|x|}{(1 + x^2)^2} \leq \frac{9\sqrt{\frac{1}{3}}}{8} < 1.$$

It follows that

$$|f(x) - f(y)| = |f^{'}(\xi)||(x-y)| < \frac{9\sqrt{\frac{1}{3}}}{8}|(x-y)|.$$

Therefore, by the Contraction Mapping Theorem, f has a unique fixed point.

EXERCISE 1.42 *(Stationary point of a convex function is a global minimum)*
Taylor's theorem (Exercise 1.39) states that if a function is twice differentiable we can find for every x and a, a point ζ that is located somewhere between x and a such that,

$$f(x) = f(a) + f^{'}(a)(x-a) + \frac{1}{2}f^{"}(\zeta)(x-a)^2. \tag{1.2}$$

We will use this property to show that if a function f is convex, twice differentiable and has a stationary point at $x = a$, $f(a)$ is a global minimum.

a. Show that if f has a stationary point at $x = a$ there exists a point ζ, that is located somewhere between x and a, such that $f(x) = f(a) + \frac{1}{2}f^{"}(\zeta)(x-a)^2$.
b. Use item a to show that if f is convex, $f(x) \geq f(a)$.
c. Show that $f(x) = x^2 - 2x + e^{-x} + 1$ has a global minimum. Hint: consider the sign scheme of the derivative to conclude that f has a stationary point.

ANSWER 1.42

a. If f has a stationary point at $x = a$, $f^{'}(a) = 0$. So, by (1.2), $f(x) = f(a) + \frac{1}{2}f^{"}(\zeta)(x-a)^2$.
b. Since f is twice differentiable, it is convex iff. $f^{"}(x) \geq 0$. In particular, $f^{"}(\zeta) \geq 0$. Consequently,

$$f(x) = f(a) + \frac{1}{2}f^{"}(\zeta)(x-a)^2 \geq f(a) \text{ for every } x.$$

Hence f attains a global minimum at $x = a$.
c. We have $f^{'}(x) = 2x - 2 - e^{-x}$, so $f^{'}(-100) > 0$ and $f^{'}(100) < 0$. Furthermore,

$$f^{"}(x) = 2 + e^{-x} \geq 0 \text{ for all } x,$$

which shows that f is convex. Therefore, by item b., f has a global minimum.

EXERCISE 1.43 Make a Taylor series expansion (1.2) for the functions f below, with $a = 0$ and $a = 1$, respectively.

a. $f(x) = e^x$. b. $f(x) = \frac{1}{1+x}$. c. $f(x) = x^2 + x + 2$.

ANSWER 1.43

a. $e^x = 1 + x + \frac{1}{2}e^{\zeta}x^2$, where ζ is located between x and 0.
$e^x = e + e(x-1) + \frac{1}{2}e^{\zeta}(x-1)^2$, where ζ is located between x and 1.

b. $\frac{1}{1+x} = 1 - x + \frac{1}{2}\frac{1}{2(1+\zeta)^3}x^2$, where ζ is located between x and 0.
$\frac{1}{1+x} = \frac{1}{2} - \frac{1}{4}(x-1) + \frac{1}{2}\frac{1}{2(1+\zeta)^3}(x-1)^2$, where ζ is located between x and 1.
c. $x^2 + x + 2 = 2 + x + \frac{1}{2} * 2 * x^2$.
$x^2 + x + 2 = 4 + 3(x-1) + \frac{1}{2} * 2 * (x-1)^2$.

EXERCISE 1.44 *(Newton-Raphson's algorithm to find a root of an equation)*
Assume a function f is twice differentiable and has a root at $x = a$. In this exercise we describe an algorithm, introduced by Newton and Raphson, to produce successively better approximations of this root a.
To that end recall from Taylor's theorem (Exercise 1.39) that for every x we can find a point ζ, located somewhere between x and a, such that:

$$f(a) = f(x) + f^{'}(x)(a-x) + \frac{1}{2}f^{"}(\zeta)(a-x)^2.$$

a. Show that if x_n is an initial guess for the root a there exists a point ξ_n, that is located somewhere between x_n and a, such that $0 = f(x_n) + f^{'}(x_n)(a-x_n) + \frac{1}{2}f^{"}(\xi_n)(a-x_n)^2$.
b. Assume our initial guess of the root, x_n, is already close to a, $f^{'}(x_n) \neq 0$ and $f^{"}(x)$ is bounded by some number M. Show that $a \approx x_{n+1} := x_n - \frac{f(x_n)}{f^{'}(x_n)}$.
c. Consider next successive (so-called Newton-Raphson) approximations of the root: $x_{n+1} := x_n - \frac{f(x_n)}{f^{'}(x_n)}$, with $x_0 = \bar{x}$. Interpret the new guess x_{n+1} geometrically. Illustrate geometrically x_1 and x_2 under the assumptions of item b.
d. Consider the error between the nth approximation x_n and the root a, $\epsilon_n := x_n - a$. Show that $\epsilon_{n+1} = \frac{1}{2}\frac{f^{"}(\xi_n)}{f^{'}(x_n)}\epsilon_n^2$, with $\epsilon_0 = \bar{x} - a$.
e. Consider a twice differentiable function f on the interval $[c,d]$, where $|f^{'}(x)| \geq m > 0$ and $|f^{"}(x)| \leq M$. Assume f has a root on this interval. Let $L := \frac{M}{2m}$.
Show that with $\tilde{\epsilon}_n := L\epsilon_n$, $\tilde{\epsilon}_{n+1} \leq \tilde{\epsilon}_n^2$. Use this inequality to conclude that $\tilde{\epsilon}_{n+1} \leq \tilde{\epsilon}_0^{2^n}$.
Assume $L(d-c) < 1$. Show that the Newton-Raphson algorithm converges to the unique root in the interval (c,d).
f. Assume f is 3 times differentiable on the interval (c,d) and has an extremal point in this interval. Use the Newton-Raphson procedure to derive a numerical algorithm to find this extremal point. Formulate sufficient conditions under which the algorithm will be guaranteed successful in finding this extremal point.
g. Formulate a recursive Newton-Raphson procedure that will provide the location x at which the function $g(x) = e^x + x^2$ attains the value 3.

ANSWER 1.44
a. This follows directly from Taylor's theorem with $x := x_n$, as $f(a) = 0$.

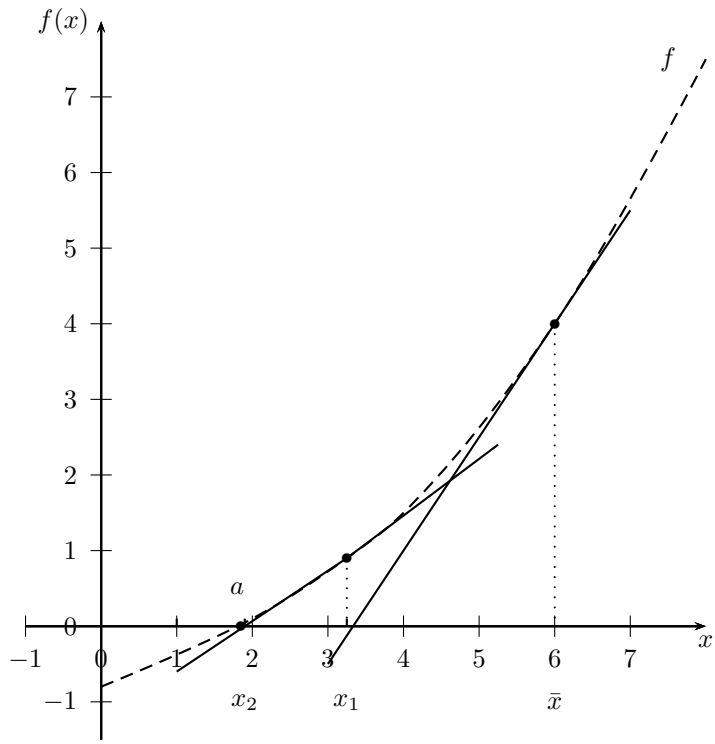


FIGURE 1.1 – *Geometric visualization Newton-Raphson algorithm.*

b. If x_n is already a good approximate of the root, it follows that $(a - x_n)^2$ is very small. So, also $(a - x_n)^2 f"(\xi)$ is a small number. This implies by item a that $0 \approx f(x_n) + f^{'}(x_n)(a - x_n)$. That is, $a \approx x_n - \frac{f(x_n)}{f^{'}(x_n)}$.

c. Note that the tangent line to the graph of f at x_n is given by

$$y(x) = f(x_n) + f^{'}(x_n)(x - x_n).$$

Thus, $y(x_{n+1}) = 0$, meaning that x_{n+1} is the point where the tangent line to the graph of f at x_n intersects the horizontal axis.

The recursive Newton-Raphson procedure is illustrated in Figure 1.1.

d. By assumption, the updates satisfy $x_{n+1} := x_n - \frac{f(x_n)}{f^{'}(x_n)}$. Equivalently,

$$x_{n+1} - a := x_n - a - \frac{f(x_n)}{f^{'}(x_n)}.$$

That is:

$$\epsilon_{n+1} = \epsilon_n - \frac{f(x_n)}{f^{'}(x_n)} \text{ (i)}.$$

By item a, we have $0 = f(x_n) + f^{'}(x_n)(a - x_n) + \frac{1}{2}f^{"}(\xi_n)(a - x_n)^2$, which implies

$$f(x_n) = f^{'}(x_n)\epsilon_n - \frac{1}{2}f^{"}(\xi_n)\epsilon_n^2.$$

Substituting this into (i) yields the advertized result.

e. By item d we have $\epsilon_{n+1} = \frac{1}{2}\frac{f^{"}(\xi_n)}{f^{'}(x_n)}\epsilon_n^2$. Thus,

$$|L\epsilon_{n+1}| = L\frac{1}{2}\frac{|f^{"}(\xi_n)|}{|f^{'}(x_n)|}\epsilon_n^2 \leq L^2\epsilon_n^2.$$

Equivalently,

$$|\tilde{\epsilon}_{n+1}| \leq \tilde{\epsilon}_n^2, \text{ for all } n = 0, 1, 2, ...$$

Applying this formula recursively yields: $|\tilde{\epsilon}_{n+1}| \leq \tilde{\epsilon}_0^{2^n}$. Hence, $|\epsilon_{n+1}| \leq \frac{(L\epsilon_0)^{2^n}}{L}$.

Therefore, if $L|\epsilon_0| < 1$, the sequence ϵ_n converges exponentially to zero. Note that $|\epsilon_0| \leq (d - c)$. Consequently, if $L(d - c) < 1$, the error converges exponentially to zero. In other words, the algorithm produces approximations that converge exponentially to the root of the function on the interval $[c, d]$.

The uniqueness of the root follows from the assumptions, which imply that the sign of $f^{'}$ does not change on the interval (c, d).

f. Consider the recursion

$$x_{n+1} := x_n - \frac{f^{'}(x_n)}{f^{"}(x_n)}, \; x_0 = \bar{x}.$$

Under the assumptions that $|f^{'''}(x)| \leq M$, $|f^{"}(x)| \geq m$ and $\frac{M}{m}(d - c) < 1$, this iterative procedure converges to the unique stationary point of the function on the interval $[c, d]$.

g. Consider the function $f(x) = e^x + x^2 - 3$. We aim to find a root of this function f. Consider the interval $[\frac{1}{2}, 1]$. Then $f(\frac{1}{2}) < 0$ and $f(1) > 0$, so a root lies within this interval. Moreover,

$$f^{'}(x) = e^x + 2x > \sqrt{e} + 1 =: m \text{ and } f^{"}(x) = e^x + 2 < e + 2 =: M.$$

Note that $\frac{1}{2}\frac{M}{m}\frac{1}{2} = \frac{e+2}{4(\sqrt{e}+1)} < 1$. Therefore, according to the previous analysis, the iterative scheme

$$x_{n+1} := x_n - \frac{e^{x_n} + x_n^2 - 3}{e^{x_n} + 2x_n} \text{ with, for example } x_0 = \frac{3}{4},$$

converges to the point where $f(x) = 0$, or, $g(x) = 3$.

EXERCISE 1.45 *(Some important limiting behaviour of standard functions)*

a. Make a Taylor series expansion for $f(x) = \ln(1 + x)$, with $a = 0$, to show that $\frac{\ln(1+x)}{x}$ approaches 1 if x approaches zero.

b. Use item a. to conclude that $(1 + x)^{\frac{1}{x}}$ approaches e if x approaches zero. (Hint: use the fact that $f(x) = e^x$ is a continuous function).

c. Make a Taylor series expansion for $f(x) = e^x$, with $a = 0$, to show that $\frac{e^x-1}{x}$ approaches 1 if x approaches zero.

d. Make a Taylor series expansion for $f(x) = e^x$, with $a = 0$, to show that $\frac{x^2}{e^x}$ approaches 0 if x goes to infinity. What about $\frac{x^n}{e^x}$ if x goes to infinity, if n is an arbitrary number?

ANSWER 1.45

a. $\ln(1+x) = x - \frac{1}{2(1+\zeta)^2}x^2$, where $0 < \zeta < x$. It follows that

$$\frac{\ln(1+x)}{x} = \frac{x - \frac{1}{2(1+\zeta)^2}x^2}{x} = 1 - \frac{1}{2(1+\zeta)^2}x.$$

Hence, $\frac{\ln(1+x)}{x}$ approaches 1 as x approaches zero (and consequently ζ too).

b. Assume that x_n approaches zero. By item a., $\frac{\ln(1+x_n)}{x_n}$ approaches 1. Equivalently, defining $s_n := \ln(1+x_n)^{\frac{1}{x_n}}$, we have s_n approaches 1 as x_n approaches zero.
Since the exponential function e^x is continuous, it follows that e^{s_n} approaches e^1 as x_n approaches zero. Therefore

$$(1+x_n)^{\frac{1}{x_n}} \text{ approaches } e \text{ as } x_n \text{ approaches zero.}$$

c. $e^x = 1 + x + \frac{1}{2}e^\zeta x^2$, where ζ lies between x and 0. It follows that

$$\frac{e^x - 1}{x} = \frac{1 + x + \frac{1}{2}e^\zeta x^2 - 1}{x} = 1 + \frac{1}{2}e^\zeta x.$$

Hence, $\frac{e^x-1}{x}$ approaches 1 as x approaches zero (and consequently ζ).

d. $e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}e^\zeta x^3$, where ζ lies between x and 0. It follows that

$$\frac{x^2}{e^x} = \frac{x^2}{1 + x + \frac{1}{2}x^2 + \frac{1}{6}e^\zeta x^3} = \frac{1}{\frac{1}{x^2} + \frac{1}{x} + \frac{1}{2} + \frac{1}{6}e^\zeta x}.$$

Since $0 < \zeta < x$, we have $e^\zeta > 1$, and thus $\frac{1}{6}e^\zeta x > \frac{1}{6}x$. Therefore,

$$\frac{1}{\frac{1}{x^2} + \frac{1}{x} + \frac{1}{2} + \frac{1}{6}e^\zeta x} \text{ approaches zero as } x \text{ goes to infinity.}$$

By considering a Taylor expansion of order $k > n$, one can similarly show that $\frac{x^n}{e^x}$ approaches zero, as x goes to infinity.

EXERCISE 1.46 *(Derivative inverse function)*

Assume f is a differentiable function that is, moreover, invertible. Denote this inverse function by f^{inv}.

a. Show that $f^{inv} \circ f(x) = x$.

b. Show that $(f^{inv})'(y) = \frac{1}{f'(x)}$, where $y = f(x)$.

c. $f(x) = e^{2x}$. Calculate $(f^{inv})'(1)$.

ANSWER 1.46

a. Follows by definition of f^{inv}.

b. From item a., we have $\frac{df^{inv}\circ f(x)}{dx} = 1$, which implies $f^{inv'}(f(x))f'(x) = 1$. Therefore,

$$f^{inv'}(y) = \frac{1}{f'(x)}, \text{ where } y = f(x).$$

c. First, note that the value of x for which $1 = e^{2x}$ is $x = 0$. Therefore, $(f^{inv})'(1) = \frac{1}{f'(0)} = \frac{1}{2}$.

EXERCISE 1.47 Assume f and g are continuous functions on $\mathbb{R}$. Consider the function $m(x) = \text{minimum}\{f(x), g(x)\}$ for all $x \in \mathbb{R}$. Motivate why $m(x)$ attains both a global maximum and a global minimum on $[-1, 1]$.

ANSWER 1.47 $m(x)$ is a continuous function on $[-1, 1]$. So, by Weierstrass' Theorem, it has both a global maximum and a global minimum on the compact set $[-1, 1]$.

EXERCISE 1.48 *(characterization extremal points if derivatives up to order $n-1$ are zero)*
Assume all derivatives of f up to some order $n-1$ are zero at $x = a$, i.e. $f^{(k)}(a) = 0,\ k = 1, \cdots, n-1$, but $f^{(n)}(a) \neq 0$ and, moreover, $f^{(n)}(x)$ is continuous.

a. Use Taylor's Theorem to conclude that for all x there exists a ξ between x and a such that

$$f(x) = f(a) + \frac{f^{(n)}(\xi)}{n!}(x-a)^n.$$

b. Assume $f^{(n)}(a) > 0$. Motivate why there is an interval I, with $a \in I$, such that for all $x \in I$, $f^{(n)}(x) > 0$.

c. Conclude from items a. and b. that if n is even and $f^{(n)}(a) > 0$, f has a minimum at $x = a$.

d. Proceed similarly as in item c. to conclude that if n is even and $f^{(n)}(a) < 0$, f has a maximum at $x = a$.

e. Proceed similarly as in item c. to conclude that if n is odd and $f^{(n)}(a) \neq 0$, f has neither a maximum nor a minimum at $x = a$.

ANSWER 1.48

a. Follows directly from (1.1).

b. Since $f^{(n)}(x)$ is continuous it follows that if $f^{(n)}(a) > 0$, there must be an interval around this point where $f^{(n)}(x) > 0$ too.

c. Assume that $f^{(n)}(x) > 0$ for all $x \in I$, where I contains a. By item a., for every $x \in I$, there exists a ξ between x and a (and therefore $\xi \in I$) such that

$$f(x) = f(a) + \frac{f^{(n)}(\xi)}{n!}(x-a)^n.$$

Since $\xi \in I$, we have $f^{(n)}(\xi) > 0$. Thus, if n is even, it follows that

$$f(x) = f(a) + \frac{f^{(n)}(\xi)}{n!}(x-a)^n \geq f(a).$$

Therefore, for every $x \in I$, $f(x) \geq f(a)$, which shows that f has a local minimum at $x = a$.

d. Assume that $f^{(n)}(x) < 0$ for all $x \in I$, where I contains a. By item a., for every $x \in I$, there exists a ξ between x and a (and therefore $\xi \in I$) such that

$$f(x) = f(a) + \frac{f^{(n)}(\xi)}{n!}(x-a)^n.$$

Since $\xi \in I$, we have $f^{(n)}(\xi) < 0$. Thus, if n is even, it follows that

$$f(x) = f(a) + \frac{f^{(n)}(\xi)}{n!}(x-a)^n \leq f(a).$$

Therefore, for every $x \in I$, $f(x) \leq f(a)$, which shows that f has a local maximum at $x = a$.

e. Assume that $f^{(n)}(x) > 0$ for all $x \in I$, where I contains a. By item a., for every $x \in I$, there exists a ξ between x and a (and therefore $\xi \in I$) such that

$$f(x) = f(a) + \frac{f^{(n)}(\xi)}{n!}(x-a)^n.$$

Since $\xi \in I$, we have $f^{(n)}(\xi) > 0$. Moreover, if $x < a$, then $x - a < 0$, and if n is odd, $(x-a)^n < 0$ as well. Hence, for $x < a$,

$$f(x) = f(a) + \frac{f^{(n)}(\xi)}{n!}(x-a)^n < f(a).$$

Similarly, for $x > a$,

$$f(x) = f(a) + \frac{f^{(n)}(\xi)}{n!}(x-a)^n > f(a).$$

Therefore, f has a saddle point at $x = a$.

By the same reasoning, f also has a saddle point at $x = a$ if $f^{(n)}(x) < 0$ for all $x \in I$, where I contains a.

EXERCISE 1.49 *(a sufficient condition under which a function is locally convex at minimum location)*
Assume $f(x)$ has a minimum at $x = a$; all derivatives of f up to some order $n-1$ are zero at $x = a$, i.e. $f^{(k)}(a) = 0,\ k = 1, \cdots, n-1$; $f^{(n)}(a) \neq 0$; and $f^{(n)}(x)$ is continuous.

a. Use Exercise 1.48 to conclude that n must be even and $f^{(n)}(a) > 0$.

b. Use Taylor's Theorem to conclude that for all x there exists a ξ between x and a such that

$$f^{"}(x) = \frac{f^{(n)}(\xi)}{(n-2)!}(x-a)^{n-2}.$$

c. Assume $f^{(n)}(a) > 0$. Motivate why there is an interval I, with $a \in I$, such that for all $x \in I$, $f^{(n)}(x) > 0$.
d. Use the previous items to conclude that $f^{"}(x) > 0$ on an open interval I containing $x = a$.
e. Use item d. to conclude that f is locally convex at $x = a$.
f. Compare this result with Example 1.50. Motivate which property does probably not apply in that example.

ANSWER 1.49

a. Since $f^{(k)}(a) = 0$ for $k = 1, \cdots, n-1$, but $f^{(n)}(a) \neq 0$, and moreover $f^{(n)}(x)$ is continuous, the conditions for applying the results from Exercise 1.48 are satisfied. Therefore, we can conclude from that exercise that n must be even and $f^{(n)}(a) > 0$, since f has a local minimum at $x = a$.
b. According Taylor's theorem (see (1.1) for an arbitrary function $g(x)$ that is $n-2$ times differentiable on an open interval I containing $x = a$ we have

$$g(x) = p_{n-3}(x) + \frac{g^{(n-2)}(\xi)}{(n-2)!}(x-a)^{(n-2)},$$

where $p_{n-3}(x)$ is given by next polynomial of degree $n-3$

$$g(a) + \frac{g^{'}(a)}{1!}(x-a) + \frac{g^{"}(a)}{2!}(x-a)^2 + \frac{g^{(3)}(a)}{3!}(x-a)^3 + \cdots + \frac{g^{(n-3)}(a)}{(n-1)!}(x-a)^{(n-3)}.$$

So, with $g(x) := f^{"}(x)$ we get:

$$f^{"}(x) = p_{n-3}(x) + \frac{f^{(n)}(\xi)}{(n-2)!}(x-a)^{(n-2)},$$

where $p_{n-3}(x)$ is given by next polynomial of degree $n-3$

$$f^{"}(a) + \frac{f^{(3)}(a)}{1!}(x-a) + \frac{f^{(4)}(a)}{2!}(x-a)^2 + \frac{f^{(5)}(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{(n-3)}.$$

As $f^{(k)}(a) = 0,\ k = 1, \cdots, n-1$, we conclude that $p_{n-3}(x) = 0$. Which yields the advertized result.
c. See Exercise 1.48, item b.
d. Consider the interval I introduced in item c. For all $x \in I$, we conclude that there exists a $\xi \in I$ such that $f^{"}(x) = \frac{f^{(n)}(\xi)}{(n-2)!}(x-a)^{(n-2)} \geq 0$.
e. As $f^{"}(x) \geq 0$ on the interval I containing $x = a$, it follows that f is convex on this interval.
f. In the example it is noted that f is infinitely many times differentiable at $x = 0$. So, since according the example f is not convex at $x = 0$, all derivatives of any order of f at $x = 0$ must be zero.

Exercises with an Economic Content

EXERCISE 1.50 The yield of corn C (in 10 kilograms) as a function of the amount of fertilizer x (in kilograms) used is given by: $R(x) = 1 + 2x - 0.025x^2$. Assume the price of corn is 2 and the price of fertilizer is $\frac{1}{2}$. Calculate the maximal profit and optimal amount of fertilizer for the farmer.

ANSWER 1.50

The profit function is

$$\Pi(x) = 2 * R(x) - \frac{1}{2}x = -0.05x^2 + \frac{7}{2}x + 2.$$

Differentiating, we get $\Pi^{'}(x) = -0.1x + \frac{7}{2}$. Setting $\Pi^{'} = 0$ gives the stationary point $x^* = \frac{2}{70}$.
Since $\Pi^{"}(x) < 0$, this point corresponds to a maximum. Therefore, x^* yields the maximum profit $\Pi(x^*)$.

EXERCISE 1.51 The total cost of producing q PC's are $0.01q^2 + 100q + 100$. Determine the number of PC's for which the average cost are minimized.

ANSWER 1.51

Average cost $AC(q) = \frac{0.01q^2+100q+100}{q}$. Therefore, $AC^{'}(q) = \frac{(0.02q+100)q-0.01q^2-100q-100}{q^2} = \frac{0.01q^2-100}{q^2}$
By solving $AC^{'}(q) = 0$ we obtain the stationary point $q = 100$ (assuming $q > 0$). Next, consider the sign scheme of the derivative.

```
          -         +    AC'(q)
(---------0--------
0         100         q
```

From above sign scheme we conclude that AC has a global minimum at $q = 100$.

EXERCISE 1.52 *(marginal output rule under perfect competition)*
Below you are asked to show that under perfect competition (that is, price p for the good is given for every firm) the amount of produced products q^* is found as that amount where the marginal revenues (i.e. the derivative of the revenue function) equals this price p. Furthermore, at q^* also the average cost and marginal cost coincide.
Assume the total cost of producing q units of a product are $f(q)$ and the market price of the product is p.

a. Show that profits are maximized at q^*, where $p = f^{'}(q^*)$.
b. Let $AC(q)$ be the average cost of producing q units of the product. Show that the number of units for which $AC(q)$ is minimal is also found as the number of units for which marginal cost and average cost coincide.

ANSWER 1.52

a. Profit of a firm is the difference between its revenues and cost. So, if the price p for the product is fixed, profit $\Pi(q) = pq - f(q)$. Therefore, its derivative is $\Pi^{'}(x) = p - f^{'}(q)$. From this we observe that the production level q^* where profits are maximized satisfies: $p = f^{'}(q^*)$.

b. Average cost $AC(q) = \frac{f(q)}{q}$. Therefore, $AC^{'}(q) = \frac{f^{'}(q)*q-f(q)}{q^2}$. Thus, $AC^{'}(q) = 0$ only if $f^{'}(q) * q - f(q) = 0$, or, $f^{'}(q^*) = \frac{f(q^*)}{q^*}$. That is, at q^* where marginal cost=average cost.

EXERCISE 1.53 Suppose that the market price of fish depends on the quantity of fish q supplied to the fish market according to $p(q) = 100e^{-q}$. The fisherman incurs a fixed cost c. Determine the value of c at which the fisherman will stop fishing.

ANSWER 1.53

Profit $\Pi(q) = p(q)*q-c = 100e^{-q}q-c$. The derivative of this function is $\Pi^{'}(q) = -100e^{-q}q+100e^{-q} = 100e^{-q}(-q+1)$. Solving $\Pi^{'}(q) = 0$, shows there is one stationary point: $q^* = 1$. The second-order derivative is $\Pi^{"}(q) = -100e^{-q}(-q+1) - 100e^{-q} = 100e^{-q}(q-2)$. Thus, $\Pi^{"}(1) < 0$. Therefore, profit is maximized at $q^* = 1$. Maximum profit equals $\Pi(1) = 100e^{-1} - c$. Thus, the fisherman stops fishing if $\Pi(1) < 0$. That is no fishing occurs if $c > \frac{100}{e}$.

EXERCISE 1.54 The number of shirts produced by a textile company is $50\sqrt{L}$, where L is the amount of labour used. The price the company gets per shirt is 10 euro, the price per unit of labor is 5 euro.

a. Determine the number of employed labour for which profits are maximal.

b. For which price of labour will the company produce no shirts anymore?

ANSWER 1.54

a. The profit function of the company is $\Pi(L) = 50\sqrt{L} * 10 - 5L$. Therefore, $\Pi^{'}(L) = 250\frac{1}{\sqrt{L}} - 5$. Thus the profit function has a stationary point at $L^* = 2500$. Note that $\Pi^{"}(L) = -125\frac{1}{L\sqrt{L}}$. Thus, $\Pi^{"}(2500) < 0$. Consequently profit is maximized at $L^* = 2500$. Maximum profit is $\Pi(2500) = 5 * 2500$.

b. Assume the price of labour is ω. Then, the profit function is $\Pi(L) = 50\sqrt{L} * 10 - \omega * L$. Therefore, $\Pi^{'}(L) = 250\frac{1}{\sqrt{L}} - \omega$. This gives the stationary point: $L^* = \frac{250^2}{\omega^2}$. Note that $\Pi^{"}(L) = -125\frac{1}{L\sqrt{L}}$. Consequently, $\Pi^{"}(L^*) < 0$. Thus profit is maximized at L^*. Maximum profit is $\Pi(L^*) = \frac{250^2}{\omega} > 0$ for all ω. Hence, the company will continue production under all circumstances.

EXERCISE 1.55 Government wants to set a tax rate, x, which maximizes total utility from taxes given by: $U(T) = T + \frac{1}{T}$. The tax, T, a person pays on gross income, W, equals $T = x * (W - 20.000)$. Determine the optimal tax rate.

ANSWER 1.55

The utility function is $U(T(x)) = T(x) + \frac{1}{T(x)}$. Therefore,

$$\frac{dU(T(x))}{dx} = \frac{dU(T)}{dT}\frac{dT(x)}{dx} = (1 - \frac{1}{T^2})(W - 20.000).$$

Thus, $U_x^{'} = 0$ at $T = 1$. Since $U^{"}(T^*) < 0$ it follows that utility is maximized at $T = 1$, resulting in $x^* = \frac{1}{W-20.000}$.

EXERCISE 1.56 The cost involved for raising a dyke are $f(x) = c_1 + c_2x + c_3e^{-\alpha x}$. Here $c_1 + c_2x$ are the cost for raising the height by x metres and $c_3e^{-\alpha x}$ is an estimate of the cost when flooding occurs. Assume $\frac{c_2}{\alpha c_3} < 1$. Determine the height that minimizes the cost.

ANSWER 1.56

Since the cost function is $f(x)) = c_1 + c_2x + c_3e^{-\alpha x}$, it follows that $f^{'}(x) = c_2 - \alpha c_3e^{-\alpha x}$. The stationary point of the cost function is obtained by solving $f_x^{'} = 0$. This yields $x = x^* := -\frac{1}{\alpha}\ln\frac{c_2}{\alpha c_3}$. Since $f^{"}(x) > 0$ it follows that at x^* the cost is minimal.

EXERCISE 1.57 Assume the cost of producing y units of a product are $y^3 - 18y^2 + 27y$ and the price per sold unit is p. Determine the supply curve of this producer as a function of p.

ANSWER 1.57

The profit function is $\Pi(y) = p * y - (y^3 - 18y^2 + 27y)$. Therefore, $\Pi^{'}(y) = p - (3y^2 - 36y + 27)$, from which we obtain the stationary points of the profit function by solving $3y^2 - 36y + 27 - p = 0$. They are: $y_{1,2} = 6 \pm \sqrt{27 + \frac{p}{36}}$. Next, consider the sign scheme of the derivative

```
          -        +        -    Π'(x)
|-----0-------0-----
0       y1       y2         x
```

From this we observe that profit is maximized at $y = y_2$ provided $\Pi(y_2) \geq \Pi(0) = 0$. Therefore, the supply curve is

$$y = 6 + \sqrt{27 + \frac{p}{3}} \text{ if } p * y - (y^3 - 18y^2 + 27y) \geq 0, \text{ and } y = 0 \text{ otherwise.}$$

EXERCISE 1.58 Consider the Cobb-Douglas production function $Y(t) = K^{\alpha}(t)L^{1-\alpha}(t)$. Here $Y(t)$ is the quantity of the product produced using a capital input of $K(t)$ and labour input of $L(t)$, all at time t. Assume capital decreases over time and labour increases over time according the next formulae: $K^{'}(t) = -\delta K(t)$ and $L^{'}(t) = \eta L(t)$. Let $y(t) := \log(\frac{Y(t)}{L(t)})$. Show that $y^{'}(t) = -\frac{1}{\ln 10}\alpha(\eta + \delta)$.

ANSWER 1.58

Direct Method: The derivative of $y(t)$ is

$$y^{'}(t) := \frac{1}{\ln 10}\frac{L(t)}{Y(t)}\frac{Y^{'}(t)L(t) - Y(t)L^{'}(t)}{L^2(t)},$$

where

$$\begin{aligned} Y^{'}(t) &= \alpha K^{\alpha-1}(t)K^{'}(t)L^{1-\alpha}(t) + (1-\alpha)K^{\alpha}(t)L^{-\alpha}(t)L^{'}(t) \\ &= -\delta\alpha K^{\alpha-1}(t)K(t)L^{1-\alpha}(t) + \eta(1-\alpha)K^{\alpha}(t)L^{-\alpha}(t)L(t) \\ &= -\delta\alpha K^{\alpha}(t)L^{1-\alpha}(t) + \eta(1-\alpha)K^{\alpha}(t)L^{1-\alpha}(t) \\ &= (\eta(1-\alpha) - \delta\alpha)Y(t). \end{aligned}$$

Substituting this into $y^{'}(t)$ gives

$$\begin{aligned} y^{'}(t) &= \frac{1}{\ln 10}\frac{L(t)}{Y(t)}\frac{(\eta(1-\alpha) - \delta\alpha)Y(t)L(t) - \eta Y(t)L(t)}{L^2(t)} \\ &= \frac{1}{\ln 10}((\eta(1-\alpha) - \delta\alpha) - \eta) = -\frac{1}{\ln 10}(\eta + \delta)\alpha. \end{aligned}$$

Indirect Method:

$$\begin{aligned} y(t) &= \log(\frac{Y(t)}{L(t)}) = \log Y(t) - \log L(t) = \log(K^{\alpha}(t)L^{1-\alpha}(t)) - \log L(t) \\ &= \alpha\log K(t) + (1-\alpha)\log L(t) - \log L(t) \\ &= \alpha\log K(t) - \alpha\log L(t) \end{aligned}$$

Therefore,

$$y^{'}(t) = \frac{1}{\ln 10}\alpha(\frac{1}{K(t)}K^{'} - \frac{1}{L(t)}L^{'}(t)) = \frac{1}{\ln 10}\alpha(-\delta - \eta).$$

EXERCISE 1.59 A producer has next cost function $C(x) = \min\{x^2, x\}$ on $[0, 2]$, for the production of x units of a product.

a. Motivate why C is continuous.

b. Draw the graph of C.

c. Is C differentiable?

d. Calculate the maximum profit if the price per sold unit is p.

e. For which price will the producer stop the production of this product?

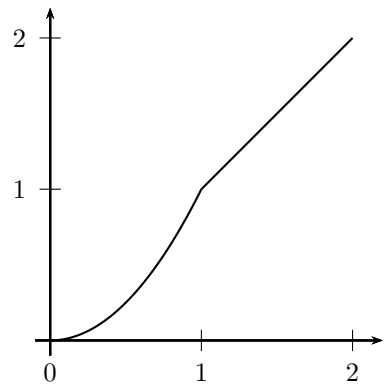


FIGURE 1.2 – *Exercise 59, item b.:* $C(x) = \min\{x^2, x\}$.

ANSWER 1.59

a. $C_1(x) = x^2$, if $x^2 \leq x$, is continuous; $C_2(x) = x$, if $x \leq x^2$, is continuous. If $x^2 = x$, $C_1(x) = C_2(x)$. Therefore, $C(x)$ is continuous.

b. See Figure 1.2

c. At $x = 1$, $C(x)$ is not differentiable. $C_1^{'}(1) = 2$ whereas $C_2^{'}(1) = 1$.

d. Profit $\Pi(x) = px - C(x) = px - x = (p-1)x$, if $x \in [1, 2]$ and $\Pi(x) = px - x^2$, if $0 \leq x \leq 1$. Note that on $[1, 2]$ $\Pi(x)$ is a linear function of x. So, the question where $\Pi(x)$ is maximized on $[1, 2]$ depends on the sign of $p-1$. If $p < 1$, it is optimal not to produce. If $p > 1$ it is optimal to produce $x = 2$, yielding a profit of $2(p-1)$.

On $[0, 1]$ profit is maximized at either $x = 0$ or $x = 1$. Note $\Pi(0) = 0$ and $\Pi(1) = p - 1$. So, if $p < 1$ it is optimal not to produce. If $p > 1$ it is optimal to produce $x = 1$.

Combining both cases we find: if $p < 1$ don't produce, if $p > 1$ produce $x = 2$ (since $2(p-1) > (p-1)$).

e. For $p = 1$.

EXERCISE 1.60 *(properties per capita instantaneous rate of growth)*

The *per capita instantaneous rate of growth* of a differentiable function $N(x)$ is defined as $r(N(x)) := \frac{N^{'}(x)}{N(x)}$.

a. Show that $r(x) = \frac{d \ln(N(x))}{dx}$.

b. Show that $r(N(x)M(x)) = r(N(x)) + r(M(x))$.

ANSWER 1.60

a. $\frac{d \ln(N(x))}{dx} = \frac{1}{N(x)} N^{'}(x) = r(x)$

b. $r(N(x)M(x)) = \frac{d \ln(N(x)M(x))}{dx} = \frac{d \ln(N(x)) + \ln(M(x))}{dx} = \frac{d \ln(N(x))}{dx} + \frac{d \ln(M(x))}{dx} = r(N(x)) + r(M(x))$.

EXERCISE 1.61 Assume the value of a tree (in terms of wood production) is approximately given by $R(t) = e^{\sqrt{t}}$ (t in years).

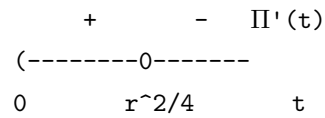
a. Assuming a discount rate of r, calculate its present value.

b. Calculate the year from now when it is optimal to cut the tree.

ANSWER 1.61

a. Its present value is: $\Pi(t) := e^{-rt}R(t) = e^{-rt+\sqrt{t}}$.

b. Note that $\Pi^{'}(t) = (-r + \frac{1}{2\sqrt{t}})e^{-rt+\sqrt{t}}$. Therefore, $\Pi(t)$ has a stationary point at $t^* = \frac{r^2}{4}$. Next, consider the sign scheme of the derivative.



From this it is clear that Π has a global maximum at t^*.

CHAPTER 2

Static Optimization: The Multi-Variable Case

2.1 Functions of more than one variable

EXERCISE 2.1 Determine in the exercises below a set D as large as possible such that the rule f defines a function from $D \to \mathbb{R}$. Calculate the range of the corresponding function.

a. $f(x, y) = x + y$.

b. $f(x, y) = x^2 + y^2$.

c. $f(x, y, z) = e^{x-y} + \frac{1}{z^2}$.

d. $f(x, y) = \{(x, y) \mid x^2 + y^2 = 1 \text{ and } y \geq 0\}$.

ANSWER 2.1

a. $D = \mathbb{R}^2$; $f(D) = \mathbb{R}$.

b. $D = \mathbb{R}^2$; $f(D) = \{y \mid y \geq 0\}$.

c. $D = \mathbb{R}^3 \setminus \{(x, y, z) \mid z = 0\}$; $f(D) = \{y \mid y > 0\}$

d. $D = \{x \mid x \in [-1, 1]\}$; $f(D) = [0, 1]$.

EXERCISE 2.2 Classify which of the next functions is linear, affine, quadratic or a quadratic form.

a. $f(x, y) = 2x + y + 1$.

b. $f(x, y) = x^2 + y^2$.

c. $f(x, y, z) = 1 + x + y + z + xz$.

d. $f(x, y, z) = x + 2y - 3z$.

ANSWER 2.2

a. affine.

b. quadratic form.

c. quadratic.

d. linear.

EXERCISE 2.3 For the functions $f(x,y)$, below, calculate the section curve at the point $(3,4)$ in the x-direction and y-direction, respectively. Also, calculate the level curve corresponding with a level of 4. Plot all these curves in the (x,y)-plane.

a. $f(x,y) = 3x - 4y + 1$.
b. $f(x,y) = x^2 + y^2$.
c. $f(x,y) = 2^{xy}$.
d. $f(x,y) = -\ln(x+y),\ x+y>0$.
e. $f(x,y) = x^{\frac{1}{2}}y^{\frac{1}{2}},\ x \geq 0, y \geq 0$.

ANSWER 2.3

a. $h(x) = 3x - 15$; $g(y) = -4y + 10$. Level curve: $y = \frac{3}{4}x - \frac{3}{4}$.
b. $h(x) = x^2 + 16$; $g(y) = y^2 + 9$. Level curve: $x^2 + y^2 = 4$, circle with radius 2, center $(0,0)$.
c. $h(x) = 2^{4x} = 16^x$; $g(y) = 2^{3y} = 8^y$; Exponential functions. Level curve: $2^{xy} = 2^2 \rightarrow xy = 2 \rightarrow y = \frac{2}{x},\ x \neq 0$, hyperbola.
d. $h(x) = -\ln(x+4),\ x > -4$; $g(y) = -\ln(y+3),\ y > -3$. Level curve: $-\ln(x+y) = 4,\ x+y>0 \rightarrow \ln(x+y) = -4,\ x+y>0 \rightarrow y = e^{-4} - x$.
e. $h(x) = 2\sqrt{x},\ x > 0$; $g(y) = \sqrt{3}\sqrt{y},\ y > 0$. Level curve: $\sqrt{x}\sqrt{y} = 4,\ x>0,\ y>0 \rightarrow xy = 16,\ x>0,\ y>0 \rightarrow y = \frac{16}{x},\ x>0$, positive part hyperbola.

2.2 Distance, Bounded Sets and Compact Sets

EXERCISE 2.4 Determine the boundary and accumulation points of next sets. Motivate which of these sets is bounded, open, closed or compact, respectively.

a. $\{(x,y) \in \mathbb{R}^2 \mid x \geq 2\}$.
b. $\{(x,y) \in \mathbb{R}^2 \mid y \geq x \text{ and } y \leq x+1\}$.
c. $\{(x,y) \in \mathbb{R}^2 \mid |x| \leq 2,\ |y| < 2\}$.
d. $\{(1,2)\}$.
e. $\{(x,y) \in \mathbb{R}^2 \mid x < 1,\ x \geq -1,\ y \leq -x,\ y \geq x\}$.

ANSWER 2.4

a. Boundary points: $\{(x,y) \in \mathbb{R}^2 \mid x = 2\}$. Accumulation points: $\{(x,y) \in \mathbb{R}^2 \mid x = 2\}$. Not bounded, not open, closed, not compact.
b. Boundary points: $\{(x,y) \in \mathbb{R}^2 \mid y = x \text{ or } y = x+1\}$. Accumulation points: $\{(x,y) \in \mathbb{R}^2 \mid y = x \text{ or } y = x+1\}$. Not bounded, not open, closed, not compact.
c. Boundary points: $\{(x,y) \in \mathbb{R}^2 \mid (2,y) \text{ where } y \in (-2,2) \text{ or } (-2,y) \text{ where } y \in (-2,2)\}$. Accumulation points: $\{(x,y) \in \mathbb{R}^2 \mid |x| = 2 \text{ where } y \in [-2,2] \text{ or } |y| = 2 \text{ where } x \in [-2,2]\}$ Bounded, not open, not closed, not compact.
d. Boundary point: $(1,2)$. Accumulation point: none. Bounded, not open, closed, compact.
e. Boundary points:
$\{(x,y) \in \mathbb{R}^2 \mid (-1,y) \text{ where } y \in [-1,1] \text{ or } (x,x) \text{ where } x \in [-1,0] \text{ or } (x,-x) \text{ where } x \in [-1,0]\}$.
Accumulation points: coincides with set of boundary points. Bounded, not open, closed, compact.

2.3 Continuous Functions

EXERCISE 2.5 Consider the function $f(x,y) = \begin{cases} 0, & \text{if } x \geq 0 \\ 1, & \text{if } x < 0 \end{cases}$.

a. Show that the sequence $s_n = (\frac{-1}{n}, 0),\ n = 1, 2, \cdots$ converges to $(0,0)$.

b. Show that f is not continuous in $(0,0)$.

ANSWER 2.5

a. $x_n := \frac{-1}{n}$ converges to zero as n goes to infinity. Hence, s_n converges to $(0,0)$ as n goes to infinity.

b. $f(s_n) = 1$ for every integer n. So $f(s_n)$ converges to 1 as n goes to infinity. However, $f(0,0) = 0$. Hence, $f(s_n)$ does not converge to $f(0,0)$, though s_n converges to $(0,0)$. Therefore, f is not continuous at $(0,0)$.

EXERCISE 2.6 Show that the function $f(x,y)$ defined by $f(x,y) = \begin{cases} \frac{1}{x+y}, & \text{if } x+y \neq 0 \\ 0, & \text{if } x+y = 0 \end{cases}$, is not continuous in $(0,0)$.

ANSWER 2.6

Consider the sequence $s_n = (\frac{1}{n}, 0)$. s_n converges to $(0,0)$ as n goes to infinity. $f(s_n) = n$; $f(0,0) = 0$. So, $f(s_n)$ does not converge to $f(0,0)$ as n goes to infinity. Therefore, f is not continuous at $(0,0)$.

EXERCISE 2.7 Motivate which of the next functions is continuous.

a. $f(x,y) = \begin{cases} \log(x^2+y^2+1), & x \leq 0; \\ \log(xy^2+y^2+1), & x > 0. \end{cases}$

b. $f(x,y,z) = ze^{x+\ln(y^2+1)}$.

c. $f(x,y) = \begin{cases} \frac{1}{x^2+y^2}, & (x,y) \neq (0,0); \\ 0, & (x,y) = (0,0). \end{cases}$

ANSWER 2.7

a. $g_1(x,y) = x^2+y^2+1$ is continuous on $\mathbb{R}^2$. Furthermore $g_1(x,y) \geq 1$. $\tilde{f}(z) = \log(z)$ is continuous on $\{z \mid z > 0\}$. Consequently, $F_1(x,y) := \tilde{f}og_1(x,y)$ is continuous on $\mathbb{R}^2$. And thus in particular on $S_1 = \{(x,y) \mid x \leq 0\}$.

$g_2(x,y) = xy^2+y^2+1$ is continuous on $\mathbb{R}^2$. Furthermore $g_2(x,y) \geq 1$ if $x > 0$. $\tilde{f}(z) = \log(z)$ is continuous on $\{z \mid z > 0\}$. Consequently, $F_2(x,y) := \tilde{f}og_2(x,y)$ is continuous on $S_2 = \{(x,y) \mid x > 0\}$.

Furthermore, $F_1(0,y) = F_2(0,y)$. $\rightarrow f$ is continuous.

b. $g(x,y,z) = x + \ln(y^2+1)$ is continuous on $\mathbb{R}^3$. $\tilde{f}(s) = e^s$ is continuous on $\mathbb{R} \rightarrow \tilde{f}og(x,y,z)$ is continuous on $\mathbb{R}^3 \rightarrow f(x,y,z) = z * \tilde{f}og(x,y,z)$ is continuous on $\mathbb{R}^3$.

c. Consider $s_n = (\frac{1}{n}, 0)$. Then s_n approaches $(0,0)$ if n goes to infinity. $f(s_n) = n^2$ does not approach $f(0,0) = 0$. So, f is not continuous at $(0,0)$.

EXERCISE 2.8 Motivate whether the next function has an optimum on S.

a. $f(x,y) = x^2 + y^2,\ S = \{(x,y) \mid x > -1,\ y > -1\}$.

b. $f(x,y,z) = \frac{xyz^2 - e^x}{\ln(x^2+2y+z)},\ S = \{(x,y) \mid -10 \leq x \leq 12,\ 1 \leq y \leq 2,\ 2 \leq z \leq 3\}$.

ANSWER 2.8

a. $f(x,y) \geq 0$; $f(0,0) = 0 \rightarrow f$ has a global minimum at $(0,0)$.

b. 1. $g(x,y,z) = x^2 + 2y + z$ is continuous on $\mathbb{R}^3$. Furthermore, $g(x,y,z) \geq 4$ on S. $\tilde{f}(s) = \ln(s)$ is continuous on $\{s \mid s > 0\} \rightarrow \tilde{f}og(x,y,z)$ is continuous on S. Note $\tilde{f}og(x,y,z) > 0$ on $S \rightarrow \frac{1}{\tilde{f}og(x,y,z)}$ is continuous on S. $f_2(x,y,z) := xyz^2 - e^x$ is continuous on $\mathbb{R}^3 \rightarrow \frac{f_2(x,y,z)}{\tilde{f}og(x,y,z)}$ is continuous on S.
2. S is bounded and closed $\rightarrow S$ is compact.
By Weierstrass' theorem it follows from 1. and 2. that f has both a global maximum and a global minimum on S.

EXERCISE 2.9 Motivate why the next set S is compact. $S = \{(x,y) \mid 1 \leq x^2 + y^2 \leq 4 \text{ and } xy \geq 1\}$.

ANSWER 2.9

$f(x,y) = x^2 + y^2$ is continuous on $\mathbb{R}^2$. So the set $S_1 := \{(x,y) \mid 1 \leq x^2 + y^2\}$ is closed and $S_2 := \{(x,y) \mid x^2 + y^2 \leq 4\}$ is closed.
$f(x,y) = xy$ is continuous on $\mathbb{R}^2$. So the set $S_3 := \{(x,y) \mid xy \geq 1\}$ is closed.
So, $S = S_1 \cap S_2 \cap S_3$ is closed.
Note S_2 is bounded $\rightarrow S$ is bounded.
So S is bounded and closed, i.e., S is compact.

2.4 Differentiable Functions

EXERCISE 2.10 Calculate next partial derivatives.

a. $\frac{\partial f}{\partial x}(1,0)$ and $\frac{\partial f}{\partial y}(1,0)$ if $f(x,y) = xy$.

b. $\frac{\partial f}{\partial x}(1,2)$ and $\frac{\partial f}{\partial y}(1,2)$ if $f(x,y) = \frac{y}{x}$.

ANSWER 2.10

a. $\frac{\partial f}{\partial x} = y$ and $\frac{\partial f}{\partial y} = x \rightarrow Df(1,0) = [0\ 1]$.

b. $\frac{\partial f}{\partial x} = -\frac{y}{x^2}$; $\frac{\partial f}{\partial y} = \frac{1}{x} \rightarrow Df(1,2) = [-2\ 1]$.

EXERCISE 2.11 Use Theorem 2.26 to show that next functions are differentiable. Determine their derivative.

a. $f(x,y) = 1$.

b. $f(x,y) = x$.

c. $f(x,y) = y$.

ANSWER 2.11

a. $Df(x,y) = [0\ 0]$. Both partial derivatives are continuous on $\mathbb{R}^2 \to f$ is differentiable on $\mathbb{R}^2$ and $f' = [0\ 0]$.

b. $Df(x,y) = [1\ 0]$. Both partial derivatives are continuous on $\mathbb{R}^2 \to f$ is differentiable on $\mathbb{R}^2$ and $f' = [1\ 0]$.

c. $Df(x,y) = [0\ 1]$. Both partial derivatives are continuous on $\mathbb{R}^2 \to f$ is differentiable on $\mathbb{R}^2$ and $f' = [0\ 1]$.

EXERCISE 2.12 Motivate why the following functions f are differentiable. Furthermore, determine their derivatives and the tangent plane to the graph of f at the indicated point.

a. $f(x,y) = x^{\alpha}y^{1-\alpha}$, at $(x,y) = (1,1)$, where α is some scalar in the interval $(0,1)$.

b. $f(x,y) = \frac{1}{2}x^2e^{y^2}$, at $(x,y) = (0,0)$.

c. $f(x,y) = \frac{x-y}{x+y}$, at $(x,y) = (1,1)$.

d. $f(x,y) = (\alpha x^2 + \beta y)^4$ at $(x,y) = (0,1)$, where $\alpha,\ \beta$ are some positive numbers.

e. $f(x,y,z) = \sqrt{xy} + \frac{z}{x}$ at $(x,y,z) = (1,1,0)$.

ANSWER 2.12

a. $f_1(x,y) = x^{\alpha}$ and $f_2(x,y) = y^{1-\alpha}$ are both differentiable at $(x,y) = (1,1) \to f(x,y)$ is differentiable at $(x,y) = (1,1)$. $f'(x,y) = [\alpha x^{\alpha-1}y^{1-\alpha}\ \ (1-\alpha)x^{\alpha}y^{-\alpha}]$ around $(x,y) = (1,1) \to f'(1,1) = [\alpha\ \ (1-\alpha)]$. Tangent plane: $z(x,y) = 1 + \alpha(x-1) + (1-\alpha)(y-1)$.

b. $f_1(x,y) = x^2$ and $f_2(x,y) = e^{y^2}$ are both differentiable at $(x,y) = (0,0) \to f(x,y)$ is differentiable at $(x,y) = (0,0)$. $f'(x,y) = [xe^{y^2}\ yx^2e^{y^2}]$ around $(x,y) = (0,0) \to f'(0,0) = [0\ 0]$. Tangent plane: $z(x,y) = 0$.

c. $f_1(x,y) = x - y$ and $f_2(x,y) = x + y$ are both differentiable at $(x,y) = (1,1)$, $f_2(1,1) \neq 0 \to f(x,y)$ is differentiable at $(x,y) = (1,1)$. $f'(x,y) = [\frac{2y}{(x+y)^2}\ \ \frac{-2x}{(x+y)^2}]$ around $(x,y) = (1,1) \to f'(1,1) = [\frac{1}{2}\ \ -\frac{1}{2}]$. Tangent plane: $z(x,y) = \frac{1}{2}(x-1) - \frac{1}{2}(y-1)$.

d. $f_1(x,y) = \alpha x^2$, $f_2(x,y) = \beta y$ and $f_3(z) = z^4$ are differentiable at $\mathbb{R}^2$ and $\mathbb{R}$, respectively. $\to f(x,y) = f_3 \circ (f_1 + f_2)(x,y)$ is differentiable at $(x,y) = (0,1)$. $\to$ around $(0,1)$

$$f'(x,y) = [8\alpha x(\alpha x^2 + \beta y)^3\ \ 4\beta(\alpha x^2 + \beta y)^3]$$

$\to f'(0,1) = [0\ \ 4\beta^4]$. Tangent plane: $z(x,y) = \beta^4 + 4\beta^4(y-1)$.

e. $f_1(x,y,z) = xy$ and $f_2(s) = \sqrt{s}$ are both differentiable at $(x,y,z) = (1,1,0)$ and $s = 1$, respectively. $f_3(x,y,z) = z$ and $f_4(x,y,z) = x$ are both differentiable on $\mathbb{R}^3$, $f_4(1,1,0) \neq 0 \to f(x,y,z) = f_2 \circ f_1(x,y,z) + \frac{f_3}{f_4}(x,y,z)$ is differentiable at $(x,y,z) = (1,1,0)$. $\to$ around $(1,1,0)$

$$f'(x,y,z) = [\frac{1}{2}y(xy)^{-\frac{1}{2}} - \frac{z}{x^2}\ \ \frac{1}{2}x(xy)^{-\frac{1}{2}}\ \ \frac{1}{x}]$$

Hence, $f'(1,1,0) = [\frac{1}{2}\ \ \frac{1}{2}\ \ 1]$, and the tangent plane at $(1,1,0)$ is $h(x,y,z) = 1 + \frac{1}{2}(x-1) + \frac{1}{2}(y-1) + z$.

2.5 First-Order Condition for an Optimum of a Multi-variable Function

EXERCISE 2.13 Determine the stationary points of the functions defined by the following formulas:

a. $x^2 - y^2$.
b. $\ln(1 + x^2 + y^2)$.
c. $x^3 - y^3 + 9xy$.
d. $\frac{x^2+2x}{1+y^2}$.
e. $(x - y)^2 + (z - y)^2 + y^4 - 2y^2 + 1$.
f. $y^2 e^{x^2}$.

ANSWER 2.13

a. $Df(x, y) = [2x \quad -2y] \rightarrow$ Stationary points: $Df(x, y) = [0\ 0] \rightarrow x = 0, y = 0 \rightarrow$ Stationary point: $(0, 0)$.
b. $Df(x, y) = [\frac{2x}{1+x^2+y^2} \quad \frac{2y}{1+x^2+y^2}] \rightarrow$ Stationary points: $Df(x, y) = [0\ 0] \rightarrow x = 0, y = 0 \rightarrow$ Stationary point: $(0, 0)$.
c. $Df(x, y) = [3x^2+9y \quad -3y^2+9x] \rightarrow$ Stationary points: $Df(x, y) = [0\ 0] \rightarrow x = \frac{1}{3}y^2$ and $3x^2+9y = 0 \rightarrow \frac{1}{3}y^4 + 9y = 0 \rightarrow y(\frac{1}{3}y^3 + 9) = 0 \rightarrow y = 0$ or $y = -3 \rightarrow$ Stationary points: $(0, 0)$ and $(3, -3)$.
d. $Df(x, y) = [\frac{2x+2}{1+y^2} \quad \frac{-2y(x^2+2x)}{(1+y^2)^2}] \rightarrow$ Stationary points: $Df(x, y) = [0\ 0] \rightarrow 2x + 2 = 0$ and $-2y(x^2 + 2x) = 0 \rightarrow x = -1$ and $(y = 0$ or $x = 0$ or $x = -2) \rightarrow$ Stationary point: $(-1, 0)$.
e. $Df(x, y, z) = [2(x-y) \quad -2(x-y)-2(z-y)+4y^3-4y \quad 2(z-y)] \rightarrow$ Stationary points: $Df(x, y) = [0\ 0\ 0] \rightarrow x = y = z$ and $4y^3 - 4y = 0 \rightarrow x = y = z$ and $4y(y^2 - 1) = 0 \rightarrow x = y = z$ and $(y = 0$ or $y = 1$ or $y = -1) \rightarrow$ Stationary points: $(0, 0, 0)$, $(1, 1, 1)$ and $(-1, -1, -1)$.
f. $Df(x, y) = [2xy^2e^{x^2} \quad 2ye^{x^2}] \rightarrow$ Stationary points: $Df(x, y) = [0\ 0] \rightarrow 2xy^2 = 0$ and $2y = 0 \rightarrow (x = 0$ or $y = 0)$ and $y = 0 \rightarrow (0, 0)$ or $(x, 0)$ with x arbitrary $\rightarrow$ Stationary points: $(x, 0)$, with $x \in \mathbb{R}$.

2.6 Convexity and Second-Order Condition for an Optimum of Multi-Variable Functions

EXERCISE 2.14

a. Consider $f(x, y) = -x^2 - y^2$. Determine the tangent plane to the graph of $f(x, y)$ at (a, b).
b. Show that $f(x, y) \leq f(a, b) + Df(a, b)[x - a\ y - b]^T$ for all $(x, y) \in \mathbb{R}^2$.
c. Verify that $f(x, y)$ is concave on $\mathbb{R}^2$.

ANSWER 2.14

a. $Df(x, y) = [-2x \quad -2y] \rightarrow$ Tangent plane: $z(x, y) = -a^2 - b^2 - 2a(x - a) - 2b(y - b)$.
b. $f(x, y) - z(x, y) = -x^2 - y^2 + a^2 + b^2 + 2a(x - a) + 2b(y - b) = -x^2 - y^2 + 2ax + 2by - a^2 - b^2 = -(x - a)^2 - (y - b)^2 \leq 0$.

c. $\lambda[x\ y]^T + (1-\lambda)[a\ b]^T = [\lambda x + (1-\lambda)a \ \ \lambda y + (1-\lambda)b]^T \rightarrow$ Using the definition of concavity, we have

$$
\begin{aligned}
&f(\lambda[x\ y]^T + (1-\lambda)[a\ b]^T) - \lambda f([x\ y]^T) - (1-\lambda) f([a\ b]^T) = \\
&-(\lambda x + (1-\lambda)a)^2 - (\lambda y + (1-\lambda)b)^2 + \lambda(x^2+y^2) + (1-\lambda)(a^2+b^2) = \\
&-\lambda^2 x^2 - 2\lambda(1-\lambda)ax - (1-\lambda)^2 a^2 - \lambda^2 y^2 - 2\lambda(1-\lambda)by - (1-\lambda)^2 b^2 + \lambda(x^2+y^2) + \\
&(1-\lambda)(a^2+b^2) = \\
&\lambda(1-\lambda)(x^2 - 2ax + a^2 + y^2 - 2by + b^2) = \lambda(1-\lambda)((x-a)^2 + (y-b)^2) \geq 0.
\end{aligned}
$$

EXERCISE 2.15 Calculate all second-order partial derivatives of next functions f.

a. $f(x,y) = xy^2$. b. $f(x,y) = e^{xy} + x^3$. c. $f(x,y) = \frac{x}{y},\ y \neq 0$. d. $f(x,y,z) = xy + yz$.

ANSWER 2.15

a. $Df(x,y) = [y^2 \ \ 2xy] \rightarrow D_{11}f(x,y) = 0$, $D_{12}f(x,y) = 2y$, $D_{21}f(x,y) = 2y$, $D_{22}f(x,y) = 2x$.

b. $Df(x,y) = [ye^{xy} + 3x^2 \ \ xe^{xy}] \rightarrow D_{11}f(x,y) = y^2 e^{xy} + 6x$, $D_{12}f(x,y) = (1+xy)e^{xy}$, $D_{21}f(x,y) = (1+xy)e^{xy}$, $D_{22}f(x,y) = x^2 e^{xy}$.

c. $Df(x,y) = [\frac{1}{y} \ \ -\frac{x}{y^2}] \rightarrow D_{11}f(x,y) = 0$, $D_{12}f(x,y) = -\frac{1}{y^2}$, $D_{21}f(x,y) = -\frac{1}{y^2}$, $D_{22}f(x,y) = \frac{2x}{y^3}$.

d. $Df(x,y) = [y \ \ x+z \ \ y] \rightarrow D_{11}f(x,y,z) = 0$, $D_{12}f(x,y,z) = 1$, $D_{13}f(x,y,z) = 0$, $D_{21}f(x,y) = 1$, $D_{22}f(x,y) = 0$, $D_{23}f(x,y,z) = 1$, $D_{31}f(x,y,z) = 0$, $D_{32}f(x,y,z) = 1$, $D_{33}f(x,y,z) = 0$.

EXERCISE 2.16 Calculate Hf for the next functions f.

a. $f(x,y) = x^\alpha y^{1-\alpha}$. b. $f(x,y,z) = (x^2 y - y^2 x)z$.

ANSWER 2.16

a. $Df(x,y) = [\alpha x^{\alpha-1} y^{1-\alpha} \ \ (1-\alpha)x^\alpha y^{-\alpha}] \rightarrow Hf(x,y) = \begin{bmatrix} \alpha(\alpha-1)x^{\alpha-2}y^{1-\alpha} & \alpha(1-\alpha)x^{\alpha-1}y^{-\alpha} \\ \alpha(1-\alpha)x^{\alpha-1}y^{-\alpha} & -\alpha(1-\alpha)x^\alpha y^{-1-\alpha} \end{bmatrix}$.

b. $Df(x,y,z) = [(2xy - y^2)z \ \ (x^2 - 2xy)z \ \ x^2 y - y^2 x] \rightarrow$

$$
Hf(x,y,z) = \begin{bmatrix} 2yz & (2x-2y)z & 2xy - y^2 \\ (2x-2y)z & -2xz & x^2 - 2xy \\ 2xy - y^2 & x^2 - 2xy & 0 \end{bmatrix}.
$$

EXERCISE 2.17 Rewrite the quadratic forms, below, as $s^T As$, with A a symmetric matrix.

a. $f(x,y) = 2x^2 + 2xy + 3y^2$. b. $f(x,y,z) = xy - 2yz + 4xz$. c. $f(x,y,z) = 2x^2 + 2yz$.

ANSWER 2.17

a. $f(x,y) = [x\ y] \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} [x\ y]^T$.

b. $f(x,y,z) = [x\ y\ z]\begin{bmatrix} 0 & \frac{1}{2} & 2 \\ \frac{1}{2} & 0 & -1 \\ 2 & -1 & 0 \end{bmatrix}[x\ y\ z]^T$.

c. $f(x,y,z) = [x\ y\ z]\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}[x\ y\ z]^T$.

EXERCISE 2.18 Calculate all principal matrices and principal minors of $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

ANSWER 2.18

The principal matrices of A are the matrices

$$[1];\ [5];\ [9];\ \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix};\ \begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix};\ \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} \text{ and } A.$$

Therefore, the principal minors are: $1,\ 5,\ 9,\ -3,\ -12,\ -3$ and 0.

EXERCISE 2.19 Verify whether the next functions f are either convex, concave or neither of them on $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively.

a. $f(x,y) = x^2 + y^2$. b. $f(x,y) = x^2 - y^2$. c. $f(x,y,z) = -x^2 - y^2 - z$.

ANSWER 2.19

a. $Df(x,y) = [2x\ \ 2y] \to Hf(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \to$ The principal matrices of Hf are

$$[2];\ [2];\ \text{and } Hf.$$

Therefore, the principal minors are: $2,\ 2$ and 4. $\to Hf$ is positive definite $\to f$ is convex.

b. $Df(x,y) = [2x\ \ -2y] \to Hf(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \to$ The principal matrices of Hf are

$$[2];\ [-2];\ \text{and } Hf.$$

Therefore, the principal minors are: $2,\ -2$ and -4. $\to Hf$ is not positive semi-definite. Similarly it follows that $B := -Hf$ is neither positive semi-definite (principal minors are: $-2,\ 2$ and -4). $\to f$ is neither convex nor concave.

c. $Df(x,y) = [-2x \;\; -2y \;\; 1] \rightarrow Hf(x,y) = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow$ The principal matrices of $B := -Hf$ are

$$[2];\ [2];\ [0];\ \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix};\ \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix};\ \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } B.$$

Therefore, the principal minors of B are: $2,\ 2,\ 0,\ 4,\ 0,\ 0$ and 0. $\rightarrow B$ is positive semi-definite. $\rightarrow Hf$ is negative semi-definite. $\rightarrow f$ is concave.

EXERCISE 2.20 Verify whether the next functions f are either convex, concave or neither of them on $S = \{(x,y) \mid x > 0,\ y > 0\}$.

a. $f(x,y) = e^{xy}$.

b. $f(x,y) = \ln(x+y+1)$.

c. $f(x,y) = e^{-x^2-y^2}$.

d. $f(x,y) = (1+x+y)^{\alpha}, \alpha \in (0,1)$.

ANSWER 2.20

a. $Df(x,y) = [ye^{xy} \;\; xe^{xy}] \rightarrow Hf(x,y) = \begin{bmatrix} y^2e^{xy} & (1+xy)e^{xy} \\ (1+xy)e^{xy} & x^2e^{xy} \end{bmatrix} \rightarrow$ The principal matrices of Hf are

$$[y^2e^{xy}];\ [x^2e^{xy}];\ \text{and } Hf.$$

Therefore, the principal minors are: $y^2e^{xy},\ x^2e^{xy}$ and $(x^2y^2 - (1+xy)^2)e^{2xy} = (-1-2xy)e^{2xy}$. Note: $(-1-2xy)e^{2xy} < 0$ on $S \rightarrow Hf$ is not positive semi-definite on $S \rightarrow f$ is not convex on S. The principal minors of $B := -Hf$ are: $-y^2e^{xy},\ -x^2e^{xy}$ and $(x^2y^2 - (1+xy)^2)e^{2xy} = (-1-2xy)e^{2xy}$. Note: $-y^2e^{2xy} < 0$ on $S \rightarrow B$ is not positive semi-definite on $S \rightarrow f$ is not concave on S.

b. $Df(x,y) = [\frac{1}{1+x+y} \;\; \frac{1}{1+x+y}] \rightarrow Hf(x,y) = \frac{-1}{(1+x+y)^2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \rightarrow$ The principal matrices of $B := -Hf$ are

$$[\frac{1}{(1+x+y)^2}];\ [\frac{-1}{(1+x+y)^2}];\ \text{and } B.$$

Therefore, the principal minors are: $\frac{1}{(1+x+y)^2},\ \frac{1}{(1+x+y)^2}$ and 0. Note: $\frac{1}{(1+x+y)^2} > 0$ on $S \rightarrow B$ is positive semi-definite on $S \rightarrow f$ is concave on S.

c. $Df(x,y) = [-2xe^{-x^2-y^2} \;\; -2ye^{-x^2-y^2}] \rightarrow Hf(x,y) = \begin{bmatrix} (4x^2-2)e^{-x^2-y^2} & 4xye^{-x^2-y^2} \\ 4xye^{-x^2-y^2} & (4y^2-2)e^{-x^2-y^2} \end{bmatrix}$ $\rightarrow$ The principal matrices of Hf are

$$[(4x^2-2)e^{-x^2-y^2}];\ [(4y^2-2)e^{-x^2-y^2}];\ \text{and } Hf.$$

Therefore, the principal minors are: $(4x^2-2)e^{-x^2-y^2}$, $(4y^2-2)e^{-x^2-y^2}$ and $\det(Hf)$. Note: $(4x^2-2)e^{-x^2-y^2} < 0$ if $(x,y) = (\frac{1}{4},\frac{1}{4}) \rightarrow Hf$ is not positive semi-definite on $S \rightarrow f$ is not convex on S.

The principal minors of $B := -Hf$ are: $(2-4x^2)e^{-x^2-y^2}$, $(2-4y^2)e^{-x^2-y^2}$ and $\det(B)$. Note: $(2-4x^2)e^{-x^2-y^2} < 0$ if $(x,y) = (1,1) \rightarrow B$ is not positive semi-definite on $S \rightarrow f$ is not concave on S.

d. $Df(x,y) = \alpha(1+x+y)^{\alpha-1}[1\ \ 1] \rightarrow Hf(x,y) = \alpha(\alpha-1)(1+x+y)^{\alpha-2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \rightarrow$ The principal matrices of $B := -Hf$ are

$$[\alpha(1-\alpha)(1+x+y)^{\alpha-2}];\ [\alpha(1-\alpha)(1+x+y)^{\alpha-2}];\ \text{and } B.$$

Therefore, the principal minors are: $\alpha(1-\alpha)(1+x+y)^{\alpha-2}$, $\alpha(1-\alpha)(1+x+y)^{\alpha-2}$ and 0. Note: $\alpha(1-\alpha)(1+x+y)^{\alpha-2} > 0$ on $S \rightarrow B$ is positive semi-definite on $S \rightarrow f$ is concave on S.

EXERCISE 2.21 Determine all optima (location, value, kind (local/global)) of the functions

a. $f(x,y) = xy^2 + x^3y - xy$.

b. $f(x,y,z) = (x-y)^2 + (z-y)^2 + y^4 - 2y^2 + 1$.

c. $f(x,y) = e^{xy}$.

ANSWER 2.21

a. $Df(x,y) = [y^2 + 3x^2y - y \ \ 2xy + x^3 - x] \rightarrow$ Stationary points:

$y^2 + 3x^2y - y = 0$ and $2xy + x^3 - x = 0 \rightarrow$

$y(y + 3x^2 - 1) = 0$ and $x(2y + x^2 - 1) = 0 \rightarrow$

($y = 0$ or $y + 3x^2 - 1 = 0$) and ($x = 0$ or $2y + x^2 - 1 = 0$) $\rightarrow$

$(x,y) = (0,0)$ or ($y = 0$ and $2y + x^2 - 1 = 0$) or ($y + 3x^2 - 1 = 0$ and $x = 0$) or ($y + 3x^2 - 1 = 0$ and $2y + x^2 - 1 = 0$) $\rightarrow$

$(x,y) = (0,0)$ or ($y = 0$ and $x^2 = 1$) or $(x,y) = (0,1)$ or ($y = 1 - 3x^2$ and $2(1-3x^2) + x^2 - 1 = 0$) $\rightarrow$

$(x,y) = (0,0)$ or $(x,y) = (1,0)$ or $(x,y) = (-1,0)$ or $(x,y) = (0,1)$ or ($y = 1 - 3x^2$ and $5x^2 = 1$) $\rightarrow$

$(x,y) = (0,0)$ or $(x,y) = (1,0)$ or $(x,y) = (-1,0)$ or $(x,y) = (0,1)$ or $(x,y) = (\sqrt{\frac{1}{5}},\frac{2}{5})$ or $(x,y) = (-\sqrt{\frac{1}{5}},\frac{2}{5})$.

Next, consider the second-order derivative: $Hf(x,y) = \begin{bmatrix} 6xy & 2y + 3x^2 - 1 \\ 2y + 3x^2 - 1 & 2x \end{bmatrix} \rightarrow$

$Hf(0,0) = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \rightarrow Hf(0,0)$ is indefinite $\rightarrow$ At $(x,y) = (0,0)$ there is a saddle-point.

$Hf(1,0) = \begin{bmatrix} 0 & 2 \\ 2 & 2 \end{bmatrix} \rightarrow Hf(1,0)$ is indefinite $\rightarrow$ At $(x,y) = (1,0)$ there is a saddle-point.

$Hf(-1,0) = \begin{bmatrix} 0 & 2 \\ 2 & -2 \end{bmatrix} \rightarrow Hf(-1,0)$ is indefinite $\rightarrow$ At $(x,y) = (-1,0)$ there is a saddle-point.

$Hf(0,1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow Hf(0,1)$ is indefinite $\rightarrow$ At $(x,y) = (0,1)$ there is a saddle-point.

$Hf(\sqrt{\frac{1}{5}}, \frac{2}{5}) = \begin{bmatrix} \frac{24}{5}\sqrt{\frac{1}{5}} & \frac{6}{5} \\ \frac{6}{5} & 2\sqrt{\frac{1}{5}} \end{bmatrix} \rightarrow Hf(\sqrt{\frac{1}{5}}, \frac{2}{5})$ is positive definite $\rightarrow$ At $(x,y) = (\sqrt{\frac{1}{5}}, \frac{2}{5})$ there is a minimum. Local min as $f(-100,1) < f(-\sqrt{\frac{1}{5}}, \frac{2}{5})$.

$Hf(\sqrt{-\frac{1}{5}}, \frac{2}{5}) = \begin{bmatrix} -\frac{24}{5}\sqrt{\frac{1}{5}} & \frac{6}{5} \\ \frac{6}{5} & -2\sqrt{\frac{1}{5}} \end{bmatrix} \rightarrow Hf(\sqrt{\frac{1}{5}}, \frac{2}{5})$ is negative definite $\rightarrow$ At $(x,y) = (-\sqrt{\frac{1}{5}}, \frac{2}{5})$ there is a maximum. Local max as $f(100,1) > f(\sqrt{\frac{1}{5}}, \frac{2}{5})$.

b. $Df(x,y,z) = [2(x-y) \quad -2(x-y) - 2(z-y) + 4y^3 - 4y \quad 2(z-y)] \rightarrow$ Stationary points:
$x = y$ and $-2(x-y) - 2(z-y) + 4y^3 - 4y = 0$ and $z = y \rightarrow$
$x = y$ and $4y^3 - 4y = 0$ and $z = y \rightarrow$
$x = y$ and $4y(y^2 - 1) = 0$ and $z = y \rightarrow$
$(x,y,z) = (0,0,0)$ or $(x,y,z) = (1,1,1)$ or $(x,y,z) = (-1,-1,-1)$.
Next, consider the second-order derivative: $Hf(x,y,z) = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 12y^2 & -2 \\ 0 & -2 & 2 \end{bmatrix} \rightarrow$

$Hf(0,0,0) = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 0 & -2 \\ 0 & -2 & 2 \end{bmatrix} \rightarrow Hf(0,0,0)$ is indefinite $\rightarrow$
At $(x,y) = (0,0,0)$ there is a saddle-point.

$Hf(1,1,1) = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 12 & -2 \\ 0 & -2 & 2 \end{bmatrix} \rightarrow Hf(1,1,1)$ is positive definite $\rightarrow$
At $(x,y) = (1,1,1)$ there is a minimum.
Note: $f(1,1,1) = 0$, whereas $f(x,y,z) = (x-y)^2 + (z-y)^2 + (y^2-1)^2 \geq 0 \rightarrow$
Global minimum at $(x,y) = (1,1,1)$.

$Hf(-1,-1,-1) = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 12 & -2 \\ 0 & -2 & 2 \end{bmatrix} \rightarrow Hf(-1,-1,-1)$ is positive definite $\rightarrow$
At $(x,y) = (-1,-1,-1)$ there is a minimum.
Note: $f(-1,-1,-1) = 0$, whereas $f(x,y,z) = (x-y)^2 + (z-y)^2 + (y^2-1)^2 \geq 0 \rightarrow$
Global minimum at $(x,y) = (-1,-1,-1)$.

c. $Df(x,y,z) = [ye^{xy} \quad xe^{xy}] \rightarrow$ Stationary point: $(x,y) = (0,0)$.
The second-order derivative is: $Hf(x,y) = \begin{bmatrix} y^2e^{xy} & (1+xy)e^{xy} \\ (1+xy)e^{xy} & x^2e^{xy} \end{bmatrix} \rightarrow$

$Hf(0,0) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow Hf(0,0)$ is indefinite $\rightarrow$ At $(x,y) = (0,0)$ there is a saddle-point.

EXERCISE 2.22 Consider the set $D = \{(x, y) \mid x \geq 0,\ y \geq 0, x \leq 2, y \leq 2\}$. Determine all global optima (location, value, kind) of the next functions on D

a. $f(x, y) = x - y + 1$. b. $f(x, y) = x^2 + 2y^2 - 2x - 4y + 1$.

ANSWER 2.22

a. $Df(x, y) = [1 \ \ -1] \rightarrow$ Stationary points: none. Note, f is continuous and D is compact $\rightarrow f$ attains global max and min at its set of boundary points $\rightarrow f$ attains global max and min at either
   i. $x = 0$, $0 \leq y \leq 2 \rightarrow f(x, y) = -y + 1 \rightarrow$ min $= -1$ at $(x, y) = (0, 2)$, max $= 1$ at $(x, y) = (0, 0)$
   ii. $x = 2$, $0 \leq y \leq 2 \rightarrow f(x, y) = -y + 3 \rightarrow$ min $= 1$ at $(x, y) = (2, 2)$, max $= 3$ at $(x, y) = (2, 0)$
   iii. $y = 0$, $0 \leq x \leq 2 \rightarrow f(x, y) = x + 1 \rightarrow$ min $= 1$ at $(x, y) = (0, 0)$, max $= 3$ at $(x, y) = (2, 0)$
   iv. $y = 2$, $0 \leq x \leq 2 \rightarrow f(x, y) = x - 1 \rightarrow$ min $= -1$ at $(x, y) = (0, 2)$, max $= 1$ at $(x, y) = (2, 2)$
   $\rightarrow f$ attains global max at $(x, y) = (2, 0)$.
   $\rightarrow f$ attains global min at $(x, y) = (0, 2)$.

b. $Df(x, y) = [2x - 2 \ \ 4y - 4] \rightarrow$ Stationary point: $(1, 1)$ with $f(1, 1) = -2$. Note, f is continuous and D is compact $\rightarrow f$ attains global max and min at either $(1, 1)$ or its set of boundary points $\rightarrow$ f attains global max and min at either
   i. $x = 0$, $0 \leq y \leq 2 \rightarrow f(x, y) = 2y^2 - 4y + 1 \rightarrow$ Stationary point: $4y - 4 = 0 \rightarrow y = 1 \rightarrow$ Extremal points: $f(0, 0) = 1$, $f(0, 1) = -1$, $f(0, 2) = 1$.
   ii. $x = 2$, $0 \leq y \leq 2 \rightarrow f(x, y) = 2y^2 - 4y + 1 \rightarrow$ Stationary point: $4y - 4 = 0 \rightarrow y = 1 \rightarrow$ Extremal points: $f(2, 0) = 1$, $f(2, 1) = -1$, $f(2, 2) = 1$.
   iii. $y = 0$, $0 \leq x \leq 2 \rightarrow f(x, y) = x^2 - 2x + 1 \rightarrow$ Stationary point: $2x - 2 = 0 \rightarrow x = 1 \rightarrow$ Extremal points: $f(0, 0) = 1$, $f(1, 0) = 0$, $f(2, 0) = 1$.
   iv. $y = 2$, $0 \leq x \leq 2 \rightarrow f(x, y) = x^2 - 2x + 1 \rightarrow$ Stationary point: $2x - 2 = 0 \rightarrow x = 1 \rightarrow$ Extremal points: $f(0, 2) = 1$, $f(1, 2) = 0$, $f(2, 2) = 1$.
   $\rightarrow f$ attains global max at $(x, y) = (2, 0)$, $(x, y) = (0, 2)$, $(x, y) = (2, 2)$ and $(x, y) = (0, 0)$.
   $\rightarrow f$ attains global min at $(x, y) = (1, 1)$.

EXERCISE 2.23 Consider a monopolist who produces two products which are imperfect substitutes. The inverse demand q_i for product $i,\ i = 1, 2$, is given by $p_1 = 60 - q_1 - \frac{1}{2}q_2$ and $p_2 = 60 - q_2 - \frac{1}{2}q_1$, respectively. Here, p_i is the price for good i.

a. Assume that the monopolist can produce both goods at zero marginal cost. Determine the optimal quantities the profit maximizing monopolist will produce. How high are his profits?

b. Assume authorities in market 1 impose a tax of t per unit sold in market 1. Discuss the consequences.

ANSWER 2.23

a. Profit function is: $\Pi(q_1, q_2) = p_1q_1 + p_2q_2 - C = (60 - q_1 - \frac{1}{2}q_2)q_1 + (60 - q_2 - \frac{1}{2}q_1)q_2 - C \rightarrow$
   $\Pi' = [60 - 2q_1 - q_2 \ \ 60 - 2q_2 - q_1] \rightarrow$ Stationary point: $\Pi' = [0 \ \ 0] \rightarrow$
   $q_1 = 30 - \frac{1}{2}q_2$ and $60 - 2q_2 - (30 - \frac{1}{2}q_2) = 0 \rightarrow$

$q_1 = 30 - \frac{1}{2}q_2$ and $q_2 = 20 \rightarrow (q_1^*, q_2^*) = (20, 20)$.
The second-order derivative is: $\Pi^{"} = \begin{bmatrix} -2 & -1 \\ -1 & -2 \end{bmatrix} \rightarrow \Pi^{"}(q_1^*, q_2^*)$ is negative definite $\rightarrow$
profits are maximal at $(q_1^*, q_2^*) = (20, 20)$.
$\Pi(20, 20) = 1200 - C \rightarrow$ Production if fixed cost C are smaller than 1200.

b. Profit function is: $\Pi(q_1, q_2) = p_1q_1 + p_2q_2 - C - tq_1 = (60 - q_1 - \frac{1}{2}q_2)q_1 + (60 - q_2 - \frac{1}{2}q_1)q_2 - C - tq_1$
$\rightarrow \Pi^{'} = [60 - 2q_1 - q_2 - t \;\; 60 - 2q_2 - q_1] \rightarrow$ Stationary point: $\Pi^{'} = [0 \;\; 0] \rightarrow$
$q_1 = 30 - \frac{1}{2}q_2 - \frac{1}{2}t$ and $60 - 2q_2 - (30 - \frac{1}{2}q_2 - \frac{1}{2}t) = 0 \rightarrow$
$q_1 = 30 - \frac{1}{2}q_2 - \frac{1}{2}t$ and $q_2 = 20 + \frac{1}{3}t \rightarrow (q_1^*, q_2^*) = (24 - \frac{2}{3}t, 20 + \frac{1}{3}t) \rightarrow$
Production if fixed cost C and tax t are such that $\Pi(q_1^*, q_2^*) \geq 0$.

Mixed Exercises

EXERCISE 2.24 Determine for the functions f below all stationary points. Indicate whether f attains at these points a local or global minimum/maximum.

a. $f(x, y) = x^2 + 2x + y^2 - 4y$.
b. $f(x, y, z) = (x - 1)(y - 1) + z^3$.
c. $f(x, y, z) = -x^2 - 4y^2 - z^3 + 6x + 2xy + 12z$.
d. $f(x, y) = \frac{1}{2}x^4 - 2x^2 + 2xy - y^2$.
e. $f(x, y) = x^2 - 2xy + \frac{1}{2}y^4$.
f. $f(x, y) = \frac{1}{2}x^4 - x^2 + y^2 - 2y + \frac{3}{2}$.
g. $f(x, y) = x^3 + 3xy + \frac{3}{2}y^2$.
h. $f(x, y) = x^2y + xy + x$.
i. $f(x, y) = (x + y^2)e^{-\frac{1}{2}x^2}$.
j. $f(x, y) = \frac{1}{3}x^3 - \frac{1}{2}x^2 + \frac{1}{3}y^3 + \frac{1}{2}y^2$.

ANSWER 2.24

a. $Df(x, y) = [2x + 2 \;\; 2y - 4] \rightarrow$ Stationary point: $(x, y) = (-1, 2)$.
$Hf(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow Hf(-1, 2) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow Hf(-1, 2)$ is positive definite $\rightarrow$
At $(x, y) = (-1, 2)$ there is a minimum.
Note: $f(-1, 2) = -5$, whereas $f(x, y) = (x + 1)^2 + (y - 2)^2 - 5 \geq -5 \rightarrow$
Global minimum at $(x, y) = (-1, 2)$.

b. $Df(x, y, z) = [y - 1 \;\; x - 1 \;\; 3z^2] \rightarrow$ Stationary point: $(x, y, z) = (1, 1, 0)$.
$Hf(x, y, z) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 6z \end{bmatrix} \rightarrow Hf(1, 1, 0) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. From this we have:
$[1\ 1\ 0]Hf(1, 1, 0)[1\ 1\ 0]^T = 2$ whereas $[1 \;\; -1\ 0]Hf(1, 1, 0)[1 \;\; -1\ 0]^T = -2 \rightarrow$
$Hf(1, 1, 0)$ is indefinite $\rightarrow$
At $(x, y, z) = (1, 1, 0)$ there is a saddle point.

c. $Df(x, y, z) = [-2x + 6 + 2y \;\; -8y + 2x \;\; -3z^2 + 12] \rightarrow$ Stationary points:
$-2x + 6 + 2y = 0$ and $x = 4y$ and $z^2 = 4 \rightarrow$
$-8y + 6 + 2y = 0$ and $x = 4y$ and ($z = 2$ or $z = -2$) $\rightarrow$

$(x, y, z) = (4, 1, 2)$ and $(x, y, z) = (4, 1, -2)$.

The second-order derivative is: $Hf(x, y, z) = \begin{bmatrix} -2 & 2 & 0 \\ 2 & -8 & 0 \\ 0 & 0 & -6z \end{bmatrix} \rightarrow$

1. $Hf(4, 1, 2) = \begin{bmatrix} -2 & 2 & 0 \\ 2 & -8 & 0 \\ 0 & 0 & -12 \end{bmatrix} \rightarrow Hf(4, 1, 2)$ is negative definite $\rightarrow$

   At $(x, y, z) = (4, 1, 2)$ there is a maximum.

   $f(0, 0, -100) > f(4, 1, 2) \rightarrow$ local maximum.

2. $Hf(4, 1, -2) = \begin{bmatrix} -2 & 2 & 0 \\ 2 & -8 & 0 \\ 0 & 0 & 12 \end{bmatrix} \rightarrow Hf(4, 1, -2)$ is indefinite $\rightarrow$

   At $(x, y, z) = (4, 1, -2)$ there is a saddle point.

d. $Df(x, y) = [2x^3 - 4x + 2y \;\; 2x - 2y] \rightarrow$ Stationary points:

$2x^3 - 4x + 2y = 0$ and $y = x \rightarrow 2x^3 - 2x = 0$ and $y = x \rightarrow$

$2x(x^2 - 1) = 0$ and $y = x \rightarrow (x, y) = (0, 0)$, $(x, y) = (1, 1)$ and $(x, y) = (-1, -1)$.

The second-order derivative is: $Hf(x, y) = \begin{bmatrix} 6x^2 - 4 & 2 \\ 2 & -2 \end{bmatrix} \rightarrow$

1. $Hf(0, 0) = \begin{bmatrix} -4 & 2 \\ 2 & -2 \end{bmatrix} \rightarrow Hf(0, 0)$ is negative definite $\rightarrow$

   At $(x, y) = (0, 0)$ there is a maximum.

   Note: $f(0, 0) = 0$, whereas $f(100, 0) > 0 \rightarrow$ local maximum at $(x, y) = (0, 0)$.

2. $Hf(1, 1) = \begin{bmatrix} 2 & 2 \\ 2 & -2 \end{bmatrix} \rightarrow Hf(1, 1)$ is indefinite $\rightarrow$

   At $(x, y) = (1, 1)$ there is a saddle point.

3. $Hf(-1, -1) = \begin{bmatrix} 2 & 2 \\ 2 & -2 \end{bmatrix} \rightarrow Hf(-1, -1)$ is indefinite $\rightarrow$

   At $(x, y) = (-1, -1)$ there is a saddle point.

e. $Df(x, y) = [2x - 2y \;\; -2x + 2y^3] \rightarrow$ Stationary points:

$y = x$ and $-2x + 2y^3 = 0 \rightarrow y = x$ and $2y(y^2 - 1) = 0 \rightarrow$

$(x, y) = (0, 0)$ and $(x, y) = (1, 1)$ and $(x, y) = (-1, -1)$.

The second-order derivative is: $Hf(x, y) = \begin{bmatrix} 2 & -2 \\ -2 & 6y^2 \end{bmatrix} \rightarrow$

1. $Hf(0, 0) = \begin{bmatrix} 2 & -2 \\ -2 & 0 \end{bmatrix} \rightarrow Hf(0, 0)$ is indefinite $\rightarrow$

   At $(x, y) = (0, 0)$ there is a saddle point.

2. $Hf(1, 1) = \begin{bmatrix} 2 & -2 \\ -2 & 6 \end{bmatrix} \rightarrow Hf(1, 1)$ is positive definite $\rightarrow$

   At $(x, y) = (1, 1)$ there is a minimum.

Note: $f(1,1) = -\frac{1}{2}$, whereas $f(x,y) = (x-y)^2 - y^2 + \frac{1}{2}y^4 = (x-y)^2 + \frac{1}{2}(y^4 - 2y^2) = (x-y)^2 + \frac{1}{2}(y^2-1)^2 - \frac{1}{2} \geq -\frac{1}{2} \rightarrow$ global minimum at $(x,y) = (1,1)$.

3. $Hf(-1,-1) = \begin{bmatrix} 2 & -2 \\ -2 & 6 \end{bmatrix} \rightarrow Hf(-1,-1)$ is positive definite $\rightarrow$

   At $(x,y) = (-1,-1)$ there is a minimum.

   Note: $f(-1,-1) = -\frac{1}{2}$, whereas $f(x,y) = (x-y)^2 - y^2 + \frac{1}{2}y^4 = (x-y)^2 + \frac{1}{2}(y^4 - 2y^2) = (x-y)^2 + \frac{1}{2}(y^2-1)^2 - \frac{1}{2} \geq -\frac{1}{2} \rightarrow$ global minimum at $(x,y) = (-1,-1)$.

f. $Df(x,y) = [2x^3 - 2x \;\; 2y - 2] \rightarrow$ Stationary points:

   $2x(x^2-1) = 0$ and $y = 1 \rightarrow (x,y) = (0,1)$ and $(x,y) = (1,1)$ and $(x,y) = (-1,1)$.

   The second-order derivative is: $Hf(x,y) = \begin{bmatrix} 6x^2 - 2 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow$

1. $Hf(0,1) = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow Hf(0,1)$ is indefinite $\rightarrow$

   At $(x,y) = (0,1)$ there is a saddle point.

2. $Hf(1,1) = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow Hf(1,1)$ is positive definite $\rightarrow$

   At $(x,y) = (1,1)$ there is a minimum.

   Note: $f(1,1) = 0$, whereas $f(x,y) = \frac{1}{2}(x^4 - 2x^2) + (y-1)^2 + \frac{1}{2} = \frac{1}{2}(x^2-1)^2 + (y-1)^2 \geq 0 \rightarrow$ global minimum at $(x,y) = (1,1)$.

3. $Hf(-1,1) = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow Hf(-1,1)$ is positive definite $\rightarrow$

   At $(x,y) = (-1,1)$ there is a minimum.

   Note: $f(-1,1) = 0$, whereas $f(x,y) = \frac{1}{2}(x^4 - 2x^2) + (y-1)^2 + \frac{1}{2} = \frac{1}{2}(x^2-1)^2 + (y-1)^2 \geq 0$ $\rightarrow$ global minimum at $(x,y) = (-1,1)$.

g. $Df(x,y) = [3x^2 + 3y \;\; 3x + 3y] \rightarrow$ Stationary points:

   $3x^2 + 3y = 0$ and $y = -x \rightarrow$

   $3x^2 - 3x = 0$ and $y = -x \rightarrow 3x(x-1) = 0$ and $y = -x \rightarrow (x,y) = (0,0)$ and $(x,y) = (1,-1)$.

   The second-order derivative is: $Hf(x,y) = \begin{bmatrix} 6x & 3 \\ 3 & 3 \end{bmatrix} \rightarrow$

1. $Hf(0,0) = \begin{bmatrix} 0 & 3 \\ 3 & 3 \end{bmatrix} \rightarrow Hf(0,0)$ is indefinite $\rightarrow$

   At $(x,y) = (0,0)$ there is a saddle point.

2. $Hf(1,-1) = \begin{bmatrix} 6 & 3 \\ 3 & 3 \end{bmatrix} \rightarrow Hf(1,-1)$ is positive definite $\rightarrow$

   At $(x,y) = (1,-1)$ there is a minimum.

   Note: $f(1,-1) = -2.5$, whereas $f(-100,0) < -2.5 \rightarrow$ local minimum.

h. $Df(x,y) = [2xy + y + 1 \;\; x^2 + x] \rightarrow$ Stationary points:

   $2xy + y + 1 = 0$ and $x(x+1) = 0 \rightarrow 2xy + y + 1 = 0$ and ($x = 0$ or $x = -1$) $\rightarrow (x,y) = (0,-1)$

and $(x, y) = (-1, 1)$.

The second-order derivative is: $Hf(x, y) = \begin{bmatrix} 6y & 2x+1 \\ 2x+1 & 0 \end{bmatrix} \rightarrow$

1. $Hf(0, -1) = \begin{bmatrix} -6 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow Hf(0, -1)$ is indefinite $\rightarrow$
   At $(x, y) = (0, -1)$ there is a saddle point.
2. $Hf(-1, 1) = \begin{bmatrix} 6 & -1 \\ -1 & 0 \end{bmatrix} \rightarrow Hf(-1, 1)$ is indefinite $\rightarrow$
   At $(x, y) = (-1, 1)$ there is a saddle point.

i. $Df(x, y) = [(1 - x(x + y^2))e^{-\frac{1}{2}x^2} \;\; 2ye^{-\frac{1}{2}x^2}] \rightarrow$ Stationary points:
$(1 - x(x + y^2)) = 0$ and $y = 0 \rightarrow 1 - x^2 = 0$ and $y = 0 \rightarrow (x, y) = (1, 0)$ and $(x, y) = (-1, 0)$.
The second-order derivative is: $Hf(x, y) = \begin{bmatrix} -(2x + y^2 + x(1 - x(x + y^2)))e^{-\frac{1}{2}x^2} & -2xye^{-\frac{1}{2}x^2} \\ -2xye^{-\frac{1}{2}x^2} & 2e^{-\frac{1}{2}x^2} \end{bmatrix}$
$\rightarrow$

1. $Hf(1, 0) = \begin{bmatrix} -2e^{-\frac{1}{2}} & 0 \\ 0 & 2e^{-\frac{1}{2}} \end{bmatrix} \rightarrow Hf(1, 0)$ is indefinite $\rightarrow$
   At $(x, y) = (1, 0)$ there is a saddle point.
2. $Hf(-1, 0) = \begin{bmatrix} 2e^{-\frac{1}{2}} & 0 \\ 0 & 2e^{-\frac{1}{2}} \end{bmatrix} \rightarrow Hf(-1, 0)$ is positive definite $\rightarrow$
   At $(x, y) = (-1, 0)$ there is a minimum.
   Note: $f(-1, 0) = -e^{-\frac{1}{2}}$ and $f(x, y) = (x + y^2)e^{-\frac{1}{2}x^2} = xe^{-\frac{1}{2}x^2} + y^2e^{-\frac{1}{2}x^2} \geq xe^{-\frac{1}{2}x^2} + 0$. Elementary analysis shows that $g(x) = xe^{-\frac{1}{2}x^2}$ has a global minimum at $x = -1$ of $-e^{-\frac{1}{2}} \rightarrow$ At $(x, y) = (-1, 0)$, f has a global minimum.

j. $Df(x, y) = [x^2 - x \;\; y^2 + y] \rightarrow$ Stationary points:
$x^2 - x = 0$ and $y^2 + y = 0 \rightarrow x(x - 1) = 0$ and $y(y + 1) = 0 \rightarrow$
($x = 0$ or $x = 1$) and ($y = 0$ or $y = -1$) $\rightarrow (x, y) = (0, 0)$, $(x, y) = (0, -1)$, $(x, y) = (1, 0)$ and $(x, y) = (1, -1)$.
The second-order derivative is: $Hf(x, y) = \begin{bmatrix} 2x - 1 & 0 \\ 0 & 2y + 1 \end{bmatrix} \rightarrow$

1. $Hf(0, 0) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \rightarrow Hf(1, 0)$ is indefinite $\rightarrow$
   At $(x, y) = (1, 0)$ there is a saddle point.
2. $Hf(0, -1) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \rightarrow Hf(0, -1)$ is negative definite $\rightarrow$
   At $(x, y) = (0, -1)$ there is a maximum.
   Note: $f(0, -1) = \frac{1}{6}$, whereas $f(0, 100) > f(0, -1) \rightarrow$ local maximum.
3. $Hf(1, 0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \rightarrow Hf(1, 0)$ is positive definite $\rightarrow$
   At $(x, y) = (1, 0)$ there is a minimum.
   Note: $f(1, 0) = -\frac{1}{6}$, whereas $f(-100, 0) < f(1, 0) \rightarrow$ local minimum.

4. $Hf(1,-1) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \rightarrow Hf(1,-1)$ is indefinite $\rightarrow$
   At $(x,y) = (1,-1)$ there is a saddle point.

EXERCISE 2.25 Determine for next functions f whether they are convex, respectively, concave.

a. $f(x,y) = (x-y)^2,\ x > 0,\ y > 0.$
b. $f(x,y) = x^3,\ x \in \mathbb{R},\ y \in \mathbb{R}.$
c. $f(x,y) = e^{-xy},\ x \in \mathbb{R},\ y \in \mathbb{R}.$
d. $f(x,y) = xy,\ x > 0,\ y > 0.$
e. $f(x,y,z) = x^2 + y^2 - z^2,\ x > 0,\ y > 0,\ z < 0.$
f. $f(x,y) = xy,\ x \in \mathbb{R},\ y = 1.$

ANSWER 2.25

a. $Df = [2(x-y) \quad -2(x-y)] \rightarrow Hf = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}.$
The principal matrices of Hf are the matrices

$$[2];\ [2];\ \text{ and } Hf.$$

Therefore, the principal minors are: $2,\ 2$, and $0 \rightarrow f$ is convex.

b. $Df = [3x^2 \ \ 0] \rightarrow Hf = \begin{bmatrix} 6x & 0 \\ 0 & 0 \end{bmatrix}.$
The principal matrices of Hf are the matrices

$$[6x];\ [0];\ \text{ and } Hf.$$

Therefore, the principal minors are: $6x,\ 0$, and 0. Note $6x < 0$ if $x < 0 \rightarrow f$ is not convex.
The principal matrices of $-Hf$ are the matrices

$$[-6x];\ [0];\ \text{ and } Hf.$$

Therefore, the principal minors are: $-6x,\ 0$, and 0. Note $-6x < 0$ if $x > 0 \rightarrow f$ is not concave.

c. $Df = [-ye^{-xy} \quad -xe^{-xy}] \rightarrow Hf = \begin{bmatrix} y^2e^{-xy} & (xy-1)e^{-xy} \\ (xy-1)e^{-xy} & x^2e^{-xy} \end{bmatrix}.$
The principal matrices of Hf are the matrices

$$[y^2e^{-xy}];\ [x^2e^{-xy}];\ \text{ and } Hf.$$

Therefore, the principal minors are: $y^2e^{-xy},\ x^2e^{-xy}$, and $((x^2y^2 - (xy-1)^2)e^{-2xy}$. Note $(2xy - 1)e^{-2xy}$ can attain negative values on $\mathbb{R}^2 \rightarrow f$ is not convex. It is easily verified that f is neither concave.

d. $Df = [y \;\; x] \to Hf = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.
The principal matrices of Hf are the matrices

$$[0];\; [0];\;\; \text{and } Hf.$$

Therefore, the principal minors are: $0,\; 0$, and $-1 \to f$ is not convex.
The principal matrices of $-Hf$ are the matrices

$$[0];\; [0];\;\; \text{and } -Hf.$$

Therefore, the principal minors are: $0,\; 0$, and $-1 \to f$ is not concave.

e. $Df = [2x \;\; 2y \;\; -2z] \to Hf = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$.
The principal matrices of Hf are the matrices

$$[2];\; [2];\; [-2];\; \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix};\; \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix};\; \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \text{ and } Hf.$$

Therefore, the principal minors are: $2,\; 2,\; -2,\; 4,\; -4,\; -4$, and $-8 \to f$ is not convex.
The principal matrices of $-Hf$ are the matrices

$$[-2];\; [-2];\; [2];\; \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix};\; \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix};\; \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix} \text{ and } -Hf.$$

Therefore, the principal minors are: $-2,\; -2,\; 2,\; 4,\; -4,\; -4$, and $-8 \to f$ is not concave.

f. $f(x) = x$, with $x \in \mathbb{R} \to f$ is both convex and concave.

EXERCISE 2.26 Determine for next functions f the tangent plane to the graph of f at the indicated point.

a. $f(x, y) = 2^{x+y}$ at $(x, y) = (1, 0)$.
b. $f(x, y) = \frac{x+y}{x-y}$ at $(x, y) = (1, 0)$.
c. $f(x, y, z) = x^2 + yz$ at $(x, y, z) = (1, 1, 1)$.
d. $f(x, y, z) = x + 2y + 3z$ at $(x, y, z) = (1, 1, 1)$.

ANSWER 2.26

a. $Df = [\ln(2)2^{x+y} \;\; \ln(2)2^{x+y}]$; note: both partial derivatives are continuous $\to Df(1, 0) = [2\ln(2) \;\; 2\ln(2)]$ and tangent plane to the graph at $(1, 0)$ is $z(x, y) = 2 + 2\ln(2)(x - 1) + 2\ln(2)y$.
b. $Df = [\frac{-2y}{(x-y)^2} \;\; \frac{2x}{(x-y)^2}]$; note: both partial derivatives are continuous in a neighborhood of $(1, 0) \to Df(1, 0) = [0 \;\; 2]$ and tangent plane to the graph at $(1, 0)$ is $z(x, y) = 1 + 2y$.
c. $Df = [2x \;\; z \;\; y]$; note: all partial derivatives are continuous $\to Df(1, 1, 1) = [2 \;\; 1 \;\; 1]$ and tangent plane to the graph at $(1, 1, 1)$ is $t(x, y, z) = 2 + 2(x - 1) + y - 1 + z - 1$.

d. $Df = [1 \;\; 2 \;\; 3]$; note: all partial derivatives are continuous $\rightarrow Df(1,1,1) = [1 \;\; 2 \;\; 3]$ and tangent plane to the graph at $(1,1,1)$ is $t(x,y,z) = 6 + (x-1) + 2(y-1) + 3(z-1) = x + 2y + 3z$.

EXERCISE 2.27 Let $\mathbb{R}^2_+ := \{(x,y) \in \mathbb{R}^2 \mid x > 0,\ y > 0\}$ and $\mathbb{R}^3_+ := \{(x,y,z) \in \mathbb{R}^3 \mid x > 0,\ y > 0,\ z > 0\}$. Motivate which functions f, below, are differentiable on the indicated set D.

a. $f(x,y) = \ln(1 - e^{-xy}),\ D = \mathbb{R}^2_+$.

b. $f(x,y) = e^{x+\ln(y)},\ D = \{(x,y) \mid x > -1,\ y > -1\}$.

c. $f(x,y) = \frac{2^{xy} - xy}{2^{-xy} + xy},\ D = \mathbb{R}^2_+$.

d. $f(x,y) = \frac{\sqrt{x+y-2}}{x^2+y^2+1},\ D = \mathbb{R}^2_+$.

e. $f(x,y) = |x + y + 1|,\ D = \mathbb{R}^2_+$.

f. $f(x,y,z) = \frac{\ln(1+e^{-xyz})}{e^{-xyz}+1},\ D = \mathbb{R}^3_+$.

ANSWER 2.27

a. $g(x,y) = 1 - e^{-xy}$ is differentiable on D. The range of $g(x,y)$ is $(0,1)$. $h(z) = \ln(z)$ is differentiable for all $z > 0 \rightarrow h \circ g(x,y)$ is differentiable on D.

b. $\ln(y)$ is not defined for $y \leq 0 \rightarrow f$ is not defined for all $(x,y) \in D$, so not differentiable for all $(x,y) \in D$.

c. $g(x,y) = 2^{-xy} + xy$ is differentiable on D. $g(x,y) > 0$ on D. $h(x,y) = 2^{xy} - xy$ is differentiable on $D \rightarrow \frac{h(x,y)}{g(x,y)}$ is differentiable on D.

d. $\sqrt{(x+y-2)}$ is not defined for $x + y < 2$ on $D \rightarrow f$ is not defined for all $(x,y) \in D$, so not differentiable for all $(x,y) \in D$.

e. $f(x,y) = x + y + 1$ on $D \rightarrow f$ is differentiable on D.

f. $g(x,y,z) = 1 + e^{-xyz}$ is differentiable on D. The range of $g(x,y,z)$ is $(1,2)$. $h(s) = \ln(s)$ is differentiable for all $s > 0 \rightarrow \frac{h \circ g(x,y,z)}{g(x,y,z)}$ is differentiable on D.

EXERCISE 2.28 Motivate which of next functions f has a global maximum on its domain D.

a. $f(x,y) = -x^2 - y^2,\ D = \{(x,y) \mid x \geq 0,\ y > 0\}$.

b. $f(x,y) = \frac{e^{xy} + 2^{xy}}{x+y+4},\ D = \{(x,y) \mid |x| \leq 1,\ |y| \leq 2\}$.

c. $f(x,y) = \frac{\ln(1+x^2+y^2)}{e^{1+x^2+y^2}},\ D = \{(x,y) \mid 1 \leq x^2 + y^2 \leq 4\}$.

d. $f(x,y) = x^2 + \sqrt{y^2 - 4},\ D = \{(x,y) \mid -4 \leq x \leq -2 \text{ and } -4 \leq y \leq -2;\ \text{or } 2 \leq x \leq 4 \text{ and } 2 \leq y \leq 4\}$.

e. $f(x,y) = \begin{cases} \frac{1}{x+y}, & x + y \neq 0; \\ 0, & x + y = 0; \end{cases} \quad D = \{(x,y) \mid |x| \leq 1,\ |y| \leq 1\}$.

f. $f(x,y) = \frac{1}{x+y+1},\ D = \{(x,y,z) \mid x \geq 0,\ y \geq 0\}$.

ANSWER 2.28

a. Note $f(x,y) \leq 0$, with $f(0,0) = 0$ on $\mathbb{R}^2$. Note that with $(0, \frac{1}{k}) \in D$, $f(0, \frac{1}{k}) = -\frac{1}{k^2}$ converges to 0 if k goes to infinity $\rightarrow$ we can approach the value of 0 arbitrarily close in D, but never attain it $\rightarrow$ f has not a global maximum on D.

b. $g(x,y) = x + y + 4 > 0$ for $(x,y) \in D \rightarrow f$ is continuous on D; D is compact $\rightarrow$ (Weierstrass' theorem) f attains a global maximum on D.

c. $g(x,y) = 1 + x^2 + y^2 > 0$ and $e^{g(x,y)} > 0 \rightarrow f$ is continuous on D; D is compact $\rightarrow$ (Weierstrass' theorem) f attains a global maximum on D.

d. $y^2 - 4 \geq 0$ on $D \rightarrow f$ is continuous on D; D is compact $\rightarrow$ (Weierstrass' theorem) f attains a global maximum on D.

e. With $(\frac{1}{k}, 0) \in D$, $f(\frac{1}{k}, 0) = k \rightarrow f$ attains arbitrarily large values if k goes to infinity $\rightarrow f$ has not a global maximum on D.

e. $f(0,0) = 1$, $f(x,y) \leq 1$ for $(x,y) \in D \rightarrow f$ has a global maximum on D.

EXERCISE 2.29 Determine the boundary and accumulation points of next sets S. Also motivate whether they are open/closed/bounded/compact, respectively.

a. $\{(x,y) \mid 1 < x \leq 2,\ y \in \mathbb{R};\ \text{or } 4 \leq x \leq 5,\ 0 \leq y \leq 1\}$.

b. $\{(x,y) \mid y \leq x \leq y + 1 \text{ and } 0 \leq y + x \leq 1\}$.

c. $\{(x,y) \mid x > 0 \text{ and } y > 0\}$.

d. $\{(x,y) \mid x \leq 0 \text{ and } y > 0\}$.

ANSWER 2.29

a. Boundary points: $\{(2,y),\ y \in R\}$, $\{(4,y),\ 0 \leq y \leq 1\}$, $\{(5,y),\ 0 \leq y \leq 1\}$, $\{(x,0),\ 4 \leq x \leq 5\}$ and $\{(x,1),\ 4 \leq x \leq 5\}$.
Accumulation points: all boundary points plus $\{(1,y),\ y \in R\}$.
Not open: Every neighborhood of $(2,0)$ contains points outside S.
Not closed: $(1,0)$ belongs to the complement set of S, whereas every neighborhood of $(1,0)$ contains points of S.
Not bounded: $(2,k) \in S$ for every $k \in \mathbb{R}$. $\|(2,k)\| = \sqrt{4 + k^2}$ goes to infinity if k goes to infinity.
Not compact: set is not closed/set is not bounded.

b. Boundary points: $\{(x,x),\ 0 \leq x \leq \frac{1}{2}\}$, $\{(x,1-x),\ \frac{1}{2} \leq x \leq 1\}$, $\{(x,-x),\ 0 \leq x \leq \frac{1}{2}\}$ and $\{(x,x-1),\ \frac{1}{2} \leq x \leq 1\}$.
Accumulation points: all boundary points.
Not open: Every neighborhood of $(0,0)$ contains points outside S.
Closed: $S_1 := \{(x,y) \mid y \leq x\}$ is closed; $S_2 := \{(x,y) \mid x \leq y+1\}$ is closed; $S_3 := \{(x,y) \mid 0 \leq y+x\}$ is closed; $S_4 := \{(x,y) \mid y + x \leq 1\}$ is closed $\rightarrow S = S_1 \cap S_2 \cap S_3 \cap S_4$ is closed.
Bounded: for every $(x,y) \in S$, $0 \leq x \leq 1$ and $-1 \leq y \leq 1 \rightarrow \|(x,y)\| \leq \sqrt{2}$.
Compact: set is closed and bounded.

c. Boundary points: none
Accumulation points: $\{(0,y),\ y \geq 0\}$ and $\{(x,0),\ x \geq 0\}$.
Open: For every point $(\bar{x}, \bar{y}) \in S$ there exists a circle with center $(\bar{x}, \bar{y})$ and radius r, where r is chosen smaller than the minimum of $\bar{x}$ and $\bar{y}$, such that this circle completely belongs to S.
Not closed: $(0,0)$ belongs to the complement set of S, whereas every neighborhood of $(0,0)$ contains points of S.
Not bounded: $(2,k) \in S$ for every $k > 0$. $\|(2,k)\| = \sqrt{4 + k^2}$ goes to infinity if k goes to infinity.
Not compact: set is not closed/set is not bounded.

d. Boundary points: $\{(0,y),\ y>0\}$.
Accumulation points: $\{(0,y),\ y\geq 0\}$ and $\{(x,0),\ x\leq 0\}$.
Not open: Every neighborhood of $(0,1)$ contains points outside S.
Not closed: $(-1,0)$ belongs to the complement set of S, whereas every neighborhood of $(-1,0)$ contains points of S.
Not bounded: $(-2,k)\in S$ for every $k>0$. $\|(-2,k)\|=\sqrt{4+k^2}$ goes to infinity if k goes to infinity.
Not compact: set is not closed/set is not bounded.

EXERCISE 2.30 Motivate which functions f, below, are continuous on the indicated set S.

a. $f(x,y)=\frac{\ln(x+y+1)}{x+y+1},\ S=\{(x,y)\mid x>0,\ y>0\}$.

b. $f(x,y)=|x|+|y|,\ S=\mathbb{R}^2$.

c. $f(x,y)=\begin{cases} x+y-1, \text{ if } x>0; \\ -x^2+xy+y-1, \text{ if } x\leq 0. \end{cases}$

d. $f(x,y)=\begin{cases} x, \text{ if } x\geq 0; \\ y, \text{ if } x<0. \end{cases}$

e. $f(x,y)=\begin{cases} \frac{xy+1}{xy}, \text{ if } xy\neq 0; \\ 0, \text{ if } xy=0. \end{cases}$

f. $f(x,y)=\sqrt{x^2+y^2+1},\ S=\mathbb{R}^2$.

ANSWER 2.30

a. $g(x,y)=x+y+1>0$ on $S\rightarrow \frac{\ln(g(x,y))}{g(x,y)}$ is continuous on S.

b. $f_1(x,y)=|x|$ is continuous on S; $f_2(x,y)=|y|$ is continuous on $S\rightarrow f=f_1+f_2$ is continuous on S.

c. $f_1(x,y)=x+y-1$ is continuous on $\mathbb{R}^2$; $f_2(x,y)=-x^2+xy+y-1$ is continuous on $\mathbb{R}^2$; $f_1(0,y)=y-1=f_2(0,y)\rightarrow f$ is continuous on S.

d. $f(-\frac{1}{k},1)=1$ for all $k>0$, and $f(0,1)=0\rightarrow$ with k going to infinity, $(-\frac{1}{k},1)$ converges to $(0,1)$, whereas $f(-\frac{1}{k},1)$ converges to 1, which differs from $f(0,1)=0\rightarrow f$ is not continuous.

e. $f(\frac{1}{k},\frac{1}{k})=k^2+1$ for all $k>0$, and $f(0,0)=0\rightarrow$ with k going to infinity, $(\frac{1}{k},\frac{1}{k})$ converges to $(0,0)$, whereas $f(\frac{1}{k},\frac{1}{k})$ goes to infinity, which differs from $f(0,0)=0\rightarrow f$ is not continuous.

f. $g(x,y)=x^2+y^2+1\geq 0$ on S, $h(s)=\sqrt{s}$ is continuous for $s\geq 0\rightarrow h\circ g(x,y)=\sqrt{g(x,y)}$ is continuous on S.

EXERCISE 2.31 Consider $f(x,y)=x^2+y^2+1$.

a. Draw the intersection curve at $(x,y)=(0,0)$ in the y-direction in a two-dimensional plot.

b. Draw the intersection curve at $(x,y)=(0,0)$ in the x-direction in a two-dimensional plot.

c. Draw the level curve with a level of 10 in a two-dimensional plot.

d. Show that $f(x,y)\geq 1$.

e. Draw the intersection curve at $(x,y)=(0,0)$ with the plane $y=x$ in a two-dimensional plot.

f. Draw the intersection curve at $(x,y)=(0,0)$ with the plane $y=\alpha x$ in a two-dimensional plot for $\alpha=2$ and $\alpha=-1$.

g. Based on the above results, make a plot of the graph of f in three-dimensional space.

ANSWER 2.31

a. $g(y) = y^2 + 1$.

b. $g(x) = x^2 + 1$.

c. $x^2 + y^2 + 1 = 10 \rightarrow x^2 + y^2 = 9 \rightarrow$ circle with center $(0, 0)$ and radius 3.

d. $x^2 + y^2 + 1 \geq 0 + 0 + 1 = 1$.

e. $g(x) = f(x, x) = 2x^2 + 1$.

f. Intersection curve with $y = 2x$: $g(x) = f(x, 2x) = x^2 + 4x^2 + 1 = 5x^2 + 1$.
Intersection curve with $y = -x$: $g(x) = f(x, -x) = x^2 + x^2 + 1 = 2x^2 + 1$.

g. Parabola with minimum value at $(0, 0)$ of 1, that is rotated around the z-axis.

EXERCISE 2.32 Determine which of the quadratic forms f, below, are either positive definite/positive semi-definite/negative definite/ negative semi-definite.

a. $f(x, y, z) = x^2 + 4xy + y^2$.

b. $f(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2xz + 2yz$.

c. $f(x, y, z) = -x^2 - yz$.

d. $f(x, y, z) = -x^2 + 2xy - 2y^2 + 4yz - 3z^2$.

ANSWER 2.32

a. $f(x, y) = [x\ y] \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \rightarrow$ Consider $A := \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. A is indefinite $\rightarrow f$ is indefinite.
Quicker: $f(x, y) = (x + 2y)^2 - 3y^2 \rightarrow f(1, 0) > 0$; $f(-2, 1) < 0 \rightarrow f$ is indefinite.

b. $f(x, y) = [x\ y\ z] \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \rightarrow$ Consider $A := \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}$.
The principal matrices of A are the matrices

$$[1];\ [1];\ [1];\ \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix};\ \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix};\ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \text{ and } A.$$

Therefore, the principal minors are: $1,\ 1,\ 1,\ 0,\ 0,\ 0$, and $0 \rightarrow A$ is positive semi-definite $\rightarrow f$ is positive semi-definite.
Quicker: $f(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2xz + 2yz = (x - y - z)^2 \geq 0 \rightarrow f$ is positive semi-definite.
Note, not positive definite, because if $x - y - z = 0$, $f(x, y, z) = 0$.

c. $f(x, y) = [x\ y\ z] \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \rightarrow$ Consider $A := \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 \end{bmatrix}$.
The principal matrices of A are the matrices

$$[-1];\ [0];\ [0];\ \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix};\ \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix};\ \begin{bmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix} \text{ and } A.$$

Therefore, the principal minors are: $-1,\ 0,\ 0,\ 0,\ 0,\ -\frac{1}{4}$, and $\frac{1}{4} \rightarrow A$ is not positive semi-definite.
The principal matrices of $-A$ are the matrices

$$[1];\ [0];\ [0];\ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix};\ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix};\ \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} \text{ and } -A.$$

Therefore, the principal minors are: $1,\ 0,\ 0,\ 0,\ 0,\ -\frac{1}{4}$, and $-\frac{1}{4} \rightarrow A$ is not negative semi-definite.
$\rightarrow$ Since A is neither positive or negative semi-definite f is indefinite.
Quicker: $f(1,0,0) < 0$; $f(0,-1,1) > 0 \rightarrow f$ is indefinite.

d. $f(x,y) = [x\ y\ z] \begin{bmatrix} -1 & 1 & 0 \\ 1 & -2 & 2 \\ 0 & 2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \rightarrow$ Consider $-A := \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -2 \\ 0 & -2 & 3 \end{bmatrix}$.
The principal matrices of $-A$ are the matrices

$$[1];\ [2];\ [3];\ \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix};\ \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix};\ \begin{bmatrix} 2 & -2 \\ -2 & 3 \end{bmatrix} \text{ and } -A.$$

Therefore, the principal minors are: $1,\ 2,\ 3,\ 1,\ 3,\ 2$, and $-1 \rightarrow A$ is not negative semi-definite.
It is easily verified that neither all principal minors of A are nonnegative $\rightarrow A$ is not positive semi-definite.
$\rightarrow$ Since A is neither positive or negative semi-definite f is indefinite.
Quicker: $f(x,y,z) = -(x^2 - 2xy + 2y^2 - 4yz + 3z^2) = -((x-y)^2 + y^2 - 4yz + 3z^2) = -((x - y)^2 + 3(z - \frac{2}{3}y)^2 - \frac{1}{3}y^2) \rightarrow f(1,0,0) < 0$; $f(3,3,2) > 0 \rightarrow f$ is indefinite.

Advanced, Deepening and Broadening Exercises

EXERCISE 2.33 Which of the next statements are true. Motivate your answer.

a. Every function $f(x,y)$ attains a global maximum on $S = \{(x,y) \mid x^2 + y^2 \leq 1\}$.
b. Every convex function f has the property that $f^{"}$ exists and is positive semi-definite on its domain.
c. Matrix A is negative definite if and only if all leading principal minors are smaller than zero.
d. Matrix A is positive definite if all principal minors are larger than zero.
e. Matrix A is negative semi-definite if and only if matrix $-A + \epsilon I$ is positive definite for every $\epsilon > 0$.
f. If both partial derivatives of $f(x,y)$ exist, $f(x,y)$ is differentiable.
g. If $f(x,y)$ is differentiable, both partial derivatives exist.
h. If $f(x,y)$ is differentiable, f is continuous.
i. If $f(x,y)$ is continuous, both partial derivatives exist.
j. If $f(x,y)$ is continuous, f is differentiable.
k. Assume set S is not compact. Then every boundary point of S is an element of S.

l. Assume S is a subset of $\{(x,y) \mid x^2+y^2 \leq 100\}$ and every sequence s_n that converges to some limit s, where $s_i,\ i=1,2,\cdots$ all belong to S, has the property that $s \in S$. Then $f(x,y)=x+y$ attains a global minimum on S.

m. Assume $f(x,y)$ is twice continuous differentiable; has a stationary point at $(0,0)$; and $Hf(0,0) \geq 0$. Then f has a local minimum at $(0,0)$.

n. If both partial derivatives of $f(x,y)$ exist, $f(x,y)$ is continuous.

ANSWER 2.33

a. False. Consider $f(x,y)=\frac{1}{xy}$ if $xy \neq 0$ and $f(x,y)=0$ if $xy=0$.

b. False. Consider $f(x)=0,\ x<0$ and $f(x)=x,\ x \geq 0 \rightarrow f$ is convex but not differentiable at $x=0$.

c. False. $A=\begin{bmatrix} -2 & 2 \\ 2 & -3 \end{bmatrix} \rightarrow \det(A)=2>0$, whereas all principal minors of $-A$ are positive (so A is negative definite).

d. True. See Proposition 2.49.

e. True. A is negative semi-definite iff. $-A$ is positive semi-definite. By Remark 2.50, $-A$ is positive semi-definite iff. $-A+\epsilon I$ is positive definite for every $\epsilon>0$.

f. False. See Example 2.20.

g. True. See Definition 2.23 of differentiability.

h. True. See Theorem 2.25.

i. False. $f(x,y)=|x|$ is continuous but its partial derivative w.r.t. x does not exist at, e.g., $(0,0)$.

j. False. See, e.g., item i.

k. True. By definition a point can only be a boundary point if it belongs to the set.

l. True. Set S is closed (see Definition 2.4). Moreover S is bounded. So S is compact. f is continuous. $\rightarrow$ Weierstrass' theorem implies that f attains a global minimum on S.

m. False. $f(x,y)=-x^4-y^4$ has a global maximum at $(0,0)$; $f'(0,0)=[0\ 0]$ whereas $Hf(0,0)$ equals the zero matrix, and therefore is positive semi-definite.

n. False. Consider $f(x,y)=0$ if $x=0$ or $y=0$ and $f(x,y)=1$, elsewhere.

EXERCISE 2.34 *(Example of quadratic convex function that has not a minimum)*

Consider $f(x,y)=\frac{1}{2}(x+y)^2+\alpha x+\beta y$, with $\alpha \neq \beta$.

a. Show that f is convex.

b. Show that f has no stationary points.

c. Show that f has not a minimum.

ANSWER 2.34

a. $Hf=\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. One easily verifies that Hf is positive semi-definite. So, f is convex on $\mathbb{R}^2$.

b. $f'=[x+y+\alpha,\ x+y+\beta]$. Clearly, $f'=[0\ 0]$ has no solutions it $\alpha \neq \beta$.

c. Assume w.l.o.g. that e.g. $\alpha < \beta$. Notice $f(x,y) = \frac{1}{2}(x+y)^2 + \alpha(x+y) + (\beta - \alpha)y$. Now, choose y an arbitrarily large negative value and $x = -y$. Then the value of f becomes arbitrarily large negative too.

EXERCISE 2.35 *(Cauchy-Schwarz and triangle inequality)*
Let $x,\ y \in \mathbb{R}^n$. Show (you may restrict the proof to the case $n = 2$) that

a. $|x^T y| \leq \|x\|\|y\|$ (Caucy-Schwartz inequality) Hint: Consider for a fixed x and y the quadratic polynomial $p(t) = \|x - ty\|^2$.

b. $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality) Hint: Expand $\|x + y\|^2$ and use item a.

c. Use item b. to show the so-called inverse triangle inequality: $|\|x\| - \|y\|| \leq \|x - y\|$. Hint: $x = x - y + y$.

ANSWER 2.35

a. Let $x =: [x_1\ x_2]^T$ and $y =: [y_1\ y_2]^T$, and define

$$p(t) = (x_1 - ty_1)^2 + (x_2 - ty_2)^2 \geq 0.$$

Note that

$$p(t) = (y_1^2 + y_2^2)t^2 - 2(x_1y_1 + x_2y_2)t + (x_1^2 + x_2^2),$$

which is a quadratic function in t with at most one zero. Therefore, the discriminant of $p(t) = 0$ must be non-positive, that is

$$4(x_1y_1 + x_2y_2)^2 - 4 * (x_1^2 + x_2^2) * (y_1^2 + y_2^2) \leq 0.$$

This implies

$$|x_1y_1 + x_2y_2| \leq \sqrt{x_1^2 + x_2^2} * \sqrt{y_1^2 + y_2^2}$$

or equivalently, $|x^T y| \leq \|x\|\|y\|$.

b. Using the notation from item a., we have:

$$\begin{aligned} \|x + y\|^2 &= (x_1 + y_1)^2 + (x_2 + y_2)^2 = x_1^2 + x_2^2 + 2(x_1y_1 + x_2y_2) + y_1^2 + y_2^2 \\ &= \|x\|^2 + 2x^T y + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

Hence, $\|x + y\| \leq \|x\| + \|y\|$.

c. Since $x = x - y + y$, we have

$$\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|, \text{ or equivalently } \|x\| - \|y\| \leq \|x - y\|.$$

Similarly,

$$\|y\| = \|y - x + x\| \leq \|y - x\| + \|x\| = \|x - y\| + \|x\|, \text{ or equivalently } \|y\| - \|x\| \leq \|x - y\|.$$

Combining these two inequalities, we conclude $|\|x\| - \|y\|| \leq \|x - y\|$.

EXERCISE 2.36 Assume f and g are continuous functions defined on open sets D_1 and D_2, respectively. Assume $f(x) = g(x)$ for all $x \in D_3 := D_1 \cap D_2$.
Consider the function $h(x) = \begin{cases} f(x) \text{ if } x \in D_1; \\ g(x) \text{ if } x \in D_2. \end{cases}$
Motivate why $h(x)$ is continuous on $D := D_1 \cup D_2$.

ANSWER 2.36
If the intersection set D_3 is empty, it is obvious that h is continuous on D.
Next, let $x_0 \in D_1$. Then, since $h(x) = f(x)$ and f is continuous at x_0, it follows that h is continuous at x_0.
Similarly, if $x_0 \in D_2$, then $h(x) = g(x)$. Because g is continuous at x_0, h is also continuous at x_0.
Now, let $x_0 \in D$ be arbitrary. Then x_0 belongs to either D_1 or D_2 (or possibly both). Hence h is continuous at x_0.

EXERCISE 2.37 *(Limits in $\mathbb{R}^n$, coordinate criterion)*
The intuition that a sequence $\{s_n\}_{n=1}^{\infty}$, with $s_n \in \mathbb{R}^2$, approaches the vector $s \in \mathbb{R}^2$ (or: the limit of the sequence $\{s_n\}_{n=1}^{\infty}$ is s) is formal defined as: for every $\epsilon > 0$ (however small that may be) we can find a number N_ϵ such that, for all $n > N_\epsilon$ the distance between s_n and s, that is $\|s_n - s\|$, is smaller than this ϵ.
Let $s_n := (x_n, y_n) \in \mathbb{R}^2$ and $s := (x, y)$. Use the above definition to show that

$$s_n \text{ approaches } s \text{ iff. } x_n \text{ approaches } x \text{ and } y_n \text{ approaches } y.$$

That is, a sequence in $\mathbb{R}^n$ has a limit iff. all its corresponding n coordinate sequences in $\mathbb{R}$ have a limit.

ANSWER 2.37
" $\Rightarrow$ " Let $\epsilon > 0$ be arbitrary. By definition of convergence, there exists an $N(\epsilon)$ such that

$$\|s_n - s\| < \epsilon, \text{ for all } n > N(\epsilon),$$

that is,

$$\sqrt{(x_n - x)^2 + (y_n - y)^2} < \epsilon, \text{ for all } n > N(\epsilon).$$

Since

$$\sqrt{(x_n - x)^2 + (y_n - y)^2} \geq \sqrt{(x_n - x)^2 + 0} = |x_n - x|,$$

it follows that $|x_n - x| < \epsilon$ for all $n > N(\epsilon)$, so x_n approaches x.
By a similar argument, we also conclude that y_n converges to y.
" $\Leftarrow$ " Let $\epsilon > 0$ be arbitrary. Since x_n converges to x, by definition there exists $N_1(\epsilon)$ such that

$$|x_n - x| < \epsilon, \text{ for all } n > N_1(\epsilon).$$

Similarly, since y_n converges to y, there exists $N_2(\epsilon)$ such that

$$|y_n - y| < \epsilon, \text{ for all } n > N_2(\epsilon).$$

Hence, for all $n > \max\{N_1(\epsilon), N_2(\epsilon)\}$, we have

$$\|s_n - s\| = \sqrt{(x_n - x)^2 + (y_n - y)^2} < \sqrt{\epsilon^2 + \epsilon^2} = \sqrt{2}\epsilon.$$

In other words, the distance between s_n and s can be made arbitrarily small as n goes to infinity. Therefore, s_n converges to s.

EXERCISE 2.38 *(Calculation rules with Limits)*
Let s_n and $t_n \in \mathbb{R}^2$ be sequences that approach s and $t \in \mathbb{R}^2$, respectively.
Define the sequence u_n by $u_n := s_n + t_n,\ n = 1, 2, \cdots$; and for a scalar λ, v_n by $v_n := \lambda s_n,\ n = 1, 2, \cdots$.
Show that

a. u_n approaches $s + t$ and v_n approaches λs.
b. Given an example of a sequence $u_n = s_n + t_n$, where u_n approaches zero, but both s_n and t_n don't have a limit.
c. Assume $s_n,\ t_n,\ s,\ t \in \mathbb{R}$. Show that $w_n := s_n t_n$ approaches st.
d. Assume $s_n,\ t_n,\ s,\ t \in \mathbb{R}$, where $t \neq 0$. Show that $x_n := \frac{1}{t_n}$ approaches $\frac{1}{t}$. Use this result and item c. to show that $y_n := \frac{s_n}{t_n}$ approaches $\frac{s}{t}$.

ANSWER 2.38

a. Let $\epsilon > 0$. Since s_n approaches s, there exists an N_1 such that, for all $n > N_1$, $\|s_n - s\| < \epsilon$. Similarly, Since t_n approaches t, there exists an N_2 such that, for all $n > N_2$, $\|t_n - t\| < \epsilon$. So, for all $n > \max\{N_1, N_2\}$ we have (using the triangle inequality (see Exercise 2.35)):

$$\|u_n - (s + t)\| = \|s_n - s + t_n - t\| \leq \|s_n - s\| + \|t_n - t\| \leq \epsilon + \epsilon = 2\epsilon.$$

That is, the distance between u_n and $s + t$ becomes arbitrarily small if n goes to infinity $\rightarrow u_n$ approaches $s + t$ if n goes to infinity.

Similarly we have that for all $n > N_1$:

$$\|v_n - \lambda s\| = \|\lambda s_n - \lambda s\| = \|\lambda(s_n - s)\| = |\lambda|\|s_n - s\| \leq |\lambda|\epsilon.$$

That is, the distance between v_n and λs becomes arbitrarily small if n goes to infinity $\rightarrow v_n$ approaches λs if n goes to infinity.

b. Consider, e.g., $s_n := [n \;\; n]$ and $t_n := [-n \;\; -n]$.

c. Using the notation from item a. we have for all $n > \max\{N_1, N_2\}$:

$$\begin{aligned} |w_n - st| &= |s_n t_n - st| = |s_n t_n - st_n + st_n - st| \leq |s_n t_n - st_n| + |st_n - st| \\ &= |t_n||s_n - s| + |s||t_n - t| \leq |t_n|\epsilon + |s|\epsilon = (|t_n| + |s|)\epsilon. \end{aligned}$$

Notice that for $n > N_2$, $||t_n| - |t|| \leq |t_n - t| \leq \epsilon \rightarrow |t_n| \leq \epsilon + |t|$ for $n > N_2 \rightarrow |w_n - st| \leq (|t| + \epsilon + |s|)\epsilon$ for all $n > \max\{N_1, N_2\}$. That is, the distance between w_n and st becomes arbitrarily small if n goes to infinity $\rightarrow w_n$ approaches st if n goes to infinity.

d. Using the notation from item a. we have for all $n > N_2$:

$$|x_n - \frac{1}{t}| = |\frac{1}{t_n} - \frac{1}{t}| = |\frac{t - t_n}{t_n t}| \leq \frac{1}{t_n t}\epsilon.$$

Along the lines of item c. it follows that $|t_n| \geq |t| - \epsilon \rightarrow |x_n - \frac{1}{t}| \leq \frac{1}{|t| - \epsilon}\epsilon$. Note that if $t \neq 0$, $\frac{1}{t}$ is a finite number too, and therefore, for small values of ϵ, $\frac{1}{|t| - \epsilon}$ too. $\rightarrow$ the distance between x_n and $\frac{1}{t}$ becomes arbitrarily small if n goes to infinity $\rightarrow x_n$ approaches $\frac{1}{t}$ if n goes to infinity.
Since $t \neq 0$, $v_n := \frac{1}{t_n}$ approaches $\frac{1}{t}$. Consequently, by item c., $y_n = s_n * v_n$ approaches $s * \frac{1}{t}$.

EXERCISE 2.39 *(combinations of continuous functions are continuous again)*
Let $f(x)$ and $g(x)$ be defined on an open set S, $x_0 \in S$, and λ and μ be any real numbers. Furthermore, let $k(x)$ be defined on $f(S)$ and continuous at $f(x_0)$. Show, using the definition, that if the functions $f(x)$ and $g(x)$ are continuous at x_0, also next functions are continuous at x_0.
a. $\lambda f(x) + \mu g(x)$. b. $f(x)g(x)$. c. $\frac{f(x)}{g(x)}$, if $g(x_0) \neq 0$. d. $k \circ f(x)$.

ANSWER 2.39

a. Let $\{s_n\}_{n=1}^{\infty}$ be a sequence that approaches x_0. Since both f and g are continuous at x_0, it follows that the sequences $f_n := f(s_n)$ and $g_n := g(s_n)$ approach $f(x_0)$ and $g(x_0)$, respectively, as s_n approaches x_0.
By Exercise 2.38, this implies that

$$h(s_n) := \lambda f(s_n) + \mu g(s_n) = \lambda f_n + \mu g_n \text{ approaches } \lambda f(x_0) + \mu g(x_0),$$

as s_n approaches x_0. In other words, $h(s_n)$ converges to $h(x_0)$. Therefore, $h(x)$ is continuous at x_0.

b. Using the notation from item a., we have by Exercise 2.38 again that

$$h_2(s_n) := f(s_n)g(s_n) = f_n g_n \text{ approaches } f(x_0)g(x_0),$$

as s_n approaches x_0. That is: $h_2(s_n)$ converges to $h_2(x_0)$. Therefore, $h_2(x)$ is continuous at x_0.

c. Using the notation from item a., we have by Exercise 2.38 again that

$$h_3(s_n) := \frac{f(s_n)}{g(s_n)} = \frac{f_n}{g_n} \text{ approaches } \frac{f(x_0)}{g(x_0)},$$

as s_n approaches x_0. In other words, $h_3(s_n)$ converges to $h_3(x_0)$. Therefore, $h_3(x)$ is continuous at x_0.

d. Using the notation from item a., let $h_3(s_n) := k(f(s_n))$. If n goes to infinity, f_n approaches $f(x_0)$. As k is continuous at $f(x_0)$, $k(f_n)$ approaches $k(f(x_0))$. That is, $k \circ f(x)$ is continuous at x_0.

EXERCISE 2.40 *(Proof Continuity property, Proposition 2.15)*
Let the continuous function f be defined on a connected set S. Let $I := f(S)$ and $s_i \in S,\ i = 1, 2$. Assume w.l.o.g. $f(s_1) \leq f(s_2)$ and $\lambda \in [f(s_1), f(s_2)]$. Use the Intermediate Value Theorem 1.18 to show that there exists a $s \in S$ such that $f(s) = \lambda$.

ANSWER 2.40 Since S is connected, there exists a path connecting s_1 and s_2. That is, there exists a continuous function $g : [0, 1] \to S$, with $g(0) = s_1$ and $g(1) = s_2$.
Because f is a continuous function, the composite function $fog(t) :\ [0, 1] \to S$ is also continuous. Therefore, by the Intermediate Value Function Theorem, there exists $t_1 \in [0, 1]$ such that $fog(t_1) = \lambda$. That is, with $s := g(t_1)$, $f(s) = \lambda$.

EXERCISE 2.41 Show, using the definition, that next functions are differentiable on $\mathbb{R}^2$.
a. $f(x, y) = 1$. b. $f(x, y) = x$. c. $f(x, y) = xy$. (Hint: use $2\Delta x\Delta y \leq (\Delta x)^2 + (\Delta y)^2$.)

ANSWER 2.41
a. $D_1 f(x_0, y_0) = D_2 f(x_0, y_0) = 0 \to$ The candidate tangent plane is:

$$z(x, y) = f(x_0, y_0) + D_1 f(x_0, y_0)(x - x_0) + D_2 f(x_0, y_0)(y - y_0) = 1.$$

$\to$ For every $(x, y) \neq (x_0, y_0)$:

$$\frac{f(x, y) - z(x, y)}{\|(x, y) - (x_0, y_0)\|} = 0.$$

Consequently, if (x, y) approaches (x_0, y_0), this fraction is zero too. So, by definition, f is differentiable at (x_0, y_0).

b. $D_1 f(x_0, y_0) = 1;\ D_2 f(x_0, y_0) = 0 \rightarrow$ The candidate tangent plane is:

$$z(x, y) = f(x_0, y_0) + D_1 f(x_0, y_0)(x - x_0) + D_2 f(x_0, y_0)(y - y_0) = x_0 + (x - x_0) = x.$$

$\rightarrow$ For every $(x, y) \neq (x_0, y_0)$:

$$\frac{f(x, y) - z(x, y)}{\|(x, y) - (x_0, y_0)\|} = 0.$$

Consequently, if (x, y) approaches (x_0, y_0), this fraction is zero too. So, by definition, f is differentiable at (x_0, y_0).

c. $D_1 f(x_0, y_0) = y;\ D_2 f(x_0, y_0) = x \rightarrow$ The candidate tangent plane is:

$$\begin{aligned} z(x, y) &= f(x_0, y_0) + D_1 f(x_0, y_0)(x - x_0) + D_2 f(x_0, y_0)(y - y_0) \\ &= x_0 y_0 + y_0(x - x_0) + x_0(y - y_0) = -x_0 y_0 + y_0 x + x_0 y \end{aligned}$$

$\rightarrow$ For every $(x, y) \neq (x_0, y_0)$:

$$\frac{f(x, y) - z(x, y)}{\|(x, y) - (x_0, y_0)\|} = \frac{xy + x_0 y_0 - y_0 x - x_0 y}{\|(x, y) - (x_0, y_0)\|} = \frac{(x - x_0)(y - y_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}}.$$

Or, with $\Delta x := x - x_0$ and $\Delta y := y - y_0$, we have:

$$\frac{|f(x, y) - z(x, y)|}{\|(x, y) - (x_0, y_0)\|} = \frac{|\Delta x \Delta y|}{\sqrt{(\Delta x)^2 + (\Delta y)^2}} \leq \frac{1}{2} \frac{(\Delta x)^2 + (\Delta y)^2}{\sqrt{(\Delta x)^2 + (\Delta y)^2}} = \frac{1}{2} \sqrt{(\Delta x)^2 + (\Delta y)^2}.$$

Consequently, $\frac{f(x,y) - z(x,y)}{\|(x,y) - (x_0, y_0)\|}$ goes to zero if both Δx and Δy go to zero. That is, f is differentiable at (x_0, y_0).

EXERCISE 2.42 *(Differentiability implies continuity)*
Using the definition of differentiability, show that if f is differentiable at (x_0, y_0), this implies that $f(x_0 + \Delta x, y_0 + \Delta y)$ approaches $f(x_0, y_0)$ if Δx and Δy approach zero. Or, stated differently, f is continuous at x_0.

ANSWER 2.42

By definition, f is differentiable if with $z(x, y) = f(x_0, y_0) + D_1 f(x_0, y_0)(x - x_0) + D_2 f(x_0, y_0)(y - y_0)$ the quotient

$$\frac{f(x, y) - z(x, y)}{\|(x, y) - (x_0, y_0)\|}$$

approaches zero if (x, y) approaches (x_0, y_0). Or, stated differently, with $x := x_0 + \Delta x$ and $y := y_0 + \Delta y$,

$$\frac{f(x_0 + \Delta x, y_0 + \Delta y) - z(x_0 + \Delta x, y_0 + \Delta y)}{\|(x_0 + \Delta x, y_0 + \Delta y) - (x_0, y_0)\|}$$

approaches zero if $(\Delta x, \Delta y)$ approaches $(0, 0)$. That is, if

$$\frac{f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) - D_1 f(x_0, y_0)\Delta x - D_2 f(x_0, y_0)\Delta y}{\sqrt{(\Delta x)^2 + \Delta y)^2}}$$

approaches zero if $(\Delta x, \Delta y)$ approaches $(0, 0)$.
Some elementary rewriting shows:

$$\begin{aligned}
&f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = \\
&\frac{f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) - D_1 f(x_0, y_0)\Delta x - D_2 f(x_0, y_0)\Delta y}{\sqrt{(\Delta x)^2 + \Delta y)^2}} \sqrt{(\Delta x)^2 + \Delta y)^2} \\
&+ D_1 f(x_0, y_0)\Delta x + D_2 f(x_0, y_0)\Delta y
\end{aligned}$$

By the definition of differentiability

$$\frac{f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) - D_1 f(x_0, y_0)\Delta x - D_2 f(x_0, y_0)\Delta y}{\sqrt{(\Delta x)^2 + \Delta y)^2}}$$

approaches zero as $(\Delta x, \Delta y)$ approaches $(0, 0)$.
Therefore, multiplying this expression by $\sqrt{(\Delta x)^2 + \Delta y)^2}$ shows that

$$\frac{f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) - D_1 f(x_0, y_0)\Delta x - D_2 f(x_0, y_0)\Delta y}{\sqrt{(\Delta x)^2 + \Delta y)^2}} \sqrt{(\Delta x)^2 + \Delta y)^2}$$

also approaches zero as $(\Delta x, \Delta y)$ approaches $(0, 0)$.
Furthermore, obviously also $D_1 f(x_0, y_0)\Delta x + D_2 f(x_0, y_0)\Delta y$ approaches zero as $(\Delta x, \Delta y)$ approaches $(0, 0)$.
Consequently, $f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$ approaches zero as $(\Delta x, \Delta y)$ approaches $(0, 0) \to f$ is continuous at (x_0, y_0).

EXERCISE 2.43 *(Mean Value Theorem)*
The *line segment* $[a, b]$ is defined as the set of points on the line through a and b, located between a and b: $\{x \mid x = (1 - \lambda)a + \lambda b,\ 0 \leq \lambda \leq 1\}$.
The Mean Value Theorem states the next property. Assume $f(x)$ is differentiable on an open convex set $S \subset \mathbb{R}^n$. Then, for every $a, b \in S$, there exists a point c on the line segment $[a, b]$, such that

$$f(b) - f(a) = Df(c)(b - a).$$

In item b., below, you are asked to prove this theorem.

a. Consider the intersection function of $f(x)$ on the line segment $[a,b]$: $F(t) = f(a + t(b-a)),\ 0 \leq t \leq 1$. Show that there exists a $\tau \in (0,1)$ such that: $F(1) - F(0) = F^{'}(\tau)(1-0)$.

b. Use the chain rule of differentiation to conclude that there exists a point $c \in [a,b]$ such that $f(b) - f(a) = Df(c)(b-a)$.

c. Consider $f(x,y) = e^{x-y}$. Show that for every two points (x_1,y_1) and (x_2,y_2) there exists a number $\lambda \in (0,1)$, such that $e^{x_1-y_1} - e^{x_2-y_2} = e^{\lambda(x_1-y_1)+(1-\lambda)(x_2-y_2)}(x_1 - y_1 + y_2 - x_2)$.

ANSWER 2.43

a. Note that the scalar function $F(t)$ is continuous on $[0,1]$ and differentiable on $(0,1)$. So, by the Mean Value Theorem (Exercise 1.36), there exists a $\tau \in (0,1)$ such that $F(1) - F(0) = F^{'}(\tau)(1-0)$.

b. By the chain rule of differentiation: $F^{'}(\tau) = Df(a + \tau(b-a))(b-a)$. Note, $F(1) = f(b)$ and $F(0) = f(a)$. The result follows next from item a.

c. Notice $f(x,y)$ is differentiable on $\mathbb{R}^2 \rightarrow$ the conditions for applying the Mean Value Theorem are met $\rightarrow$ For the points (x_1,y_1) and (x_2,y_2) there exists a point on the line segment between both these points, i.e. $(c_1,c_2) = (\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2)$ (where λ is some number in the open interval $(0,1)$), such that $e^{x_1-y_1} - e^{x_2-y_2} = [e^{c_1-c_2} \quad -e^{c_1-c_2}][x_1 - y_1 \ \ x_2 - y_2]^T$. Substitution of c_i and expanding the right-hand side of this equation yields then the advertised result.

EXERCISE 2.44 *(combination of convex functions that are convex too)*

Assume f, g are convex functions defined on some open convex set $S \subset \mathbb{R}^n$. Let $A \in \mathbb{R}^{n \times m}$ be an arbitrary matrix and $b \in \mathbb{R}^n$.

a. Show that $\lambda f + \mu g$ is convex, if $\lambda \geq 0$ and $\mu \geq 0$.

b. Show that $h(x) := f(Ax + b)$ is convex.

c. Show that $h(x) = \max\{f(x), g(x)\}$ is convex.

ANSWER 2.44

a. $(\lambda f + \mu g)(\alpha x + (1-\alpha)y) = \lambda f(\alpha x + (1-\alpha)y) + \mu g(\alpha x + (1-\alpha)y) \leq \lambda\alpha f(x) + (1-\alpha)f(y) + \mu\alpha g(x) + (1-\alpha)g(y) = \alpha(\lambda f(x) + \mu g(x)) + (1-\alpha)(\lambda f(y) + \mu g(y)) = \alpha(\lambda f + \mu g)(x) + (1-\alpha)(\lambda f + \mu g)(y)$.

b. $h(\alpha x + (1-\alpha)y) = f(A(\alpha x + (1-\alpha)y) + b) = f(\alpha(Ax + b) + (1-\alpha)(Ay + b)) \leq \alpha f(Ax + b) + (1-\alpha)f(Ay + b) = \alpha h(x) + (1-\alpha)h(y)$.

c. $h(\alpha x + (1-\alpha)y) = \max\{f(\alpha x + (1-\alpha)y), g(\alpha x + (1-\alpha)y)\}$. Assume the maximum is attained by $f(\alpha x + (1-\alpha)y)$. Then, since f is convex, $h(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) \leq \alpha h(x) + (1-\alpha)h(y)$. Where the last inequality follows from the fact that $h(z) = \max\{f(z), g(z)\} \geq f(z)$.

EXERCISE 2.45 *(Taylor's theorem for a twice differentiable function)*

If a multi-variable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is twice differentiable, Taylor's theorem states, we can find

for every x and a a point ζ that is located on the line segment between x and a such that,

$$f(x) = f(a) + f^{'}(a)(x-a) + \frac{1}{2}(x-a)^T Hf(\zeta)(x-a). \tag{2.1}$$

In this exercise you are asked to provide a proof of this result.

a. For a fixed a and x consider the line $y(t) = tx + (1-t)a,\ t \in \mathbb{R}$. Show that $F(t) := f(y(t))$ is a twice differentiable scalar function on $\mathbb{R}$.

b. Show that $F(1) = F(0) + F^{'}(0) + \frac{1}{2}F^{"}(t_0)$, where $t_0 \in (0,1)$.

c. Prove (2.1).

ANSWER 2.45

a. Notice $y(t)$ is a linear function. So, in particular, it is twice differentiable on $\mathbb{R}$. By assumption $f(x)$ is twice differentiable on $\mathbb{R}^n$. Therefore, by the chain rule of differentiation, $F(t)$ is twice differentiable on $\mathbb{R}$.

b. Since $F(t)$ is a twice differentiable function on $\mathbb{R}$ this result follows immediately from the scalar version of Taylor's theorem (see Exercise 1.39).

c. Notice $F(1) = f(x)$, $F(0) = f(a)$, $F^{'}(t) = f^{'}(t)(x-a)$ and $F^{"}(t) = (x-a)^T f^{"}(t)(x-a)$. Substitution of this into the result of item b yields immediately the advertized result.

EXERCISE 2.46 Make a Taylor series expansion (2.1) for the functions f below, with $a = (0,0)$ and $a = (1,2)$, respectively.

a. $f(x) = e^{x+y}$. b. $f(x) = \frac{y}{1+x}$. c. $f(x) = x^2 + xy + 2y^2 + 1$.

ANSWER 2.46

a. $Df(x,y) = [e^{x+y} \ \ e^{x+y}] \rightarrow Hf(x,y) = \begin{bmatrix} e^{x+y} & e^{x+y} \\ e^{x+y} & e^{x+y} \end{bmatrix}$

$\rightarrow Df(0,0) = [1 \ \ 1] \rightarrow$ Taylor expansion:

$$\begin{aligned} f(x,y) &= 1 + x + y + \frac{1}{2}[x\ y] \begin{bmatrix} e^{\zeta_1+\zeta_2} & e^{\zeta_1+\zeta_2} \\ e^{\zeta_1+\zeta_2} & e^{\zeta_1+\zeta_2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= 1 + x + y + \frac{1}{2}e^{\zeta_1+\zeta_2}(x^2 + 2xy + y^2), \end{aligned}$$

where $\begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix} + (1-\lambda) \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ for some $\lambda \in (0,1)$

$\rightarrow \zeta_1 = \lambda x$ and $\zeta_2 = \lambda y$ for some $\lambda \in (0,1)$.

At $(1,2) \rightarrow Df(1,2) = [e^3 \;\; e^3] \rightarrow$ Taylor expansion:

$$\begin{aligned} f(x,y) &= e^3 + e^3(x-1) + e^3(y-2) + \frac{1}{2}[x-1 \; y-2] \begin{bmatrix} e^{\zeta_1+\zeta_2} & e^{\zeta_1+\zeta_2} \\ e^{\zeta_1+\zeta_2} & e^{\zeta_1+\zeta_2} \end{bmatrix} \begin{bmatrix} x-1 \\ y-2 \end{bmatrix} \\ &= e^3 + e^3(x-1) + e^3(y-2) + \frac{1}{2}e^{\zeta_1+\zeta_2}((x-1)^2 + 2(x-1)(y-2) + (y-2)^2), \end{aligned}$$

where $\begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix} + (1-\lambda) \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ for some $\lambda \in (0,1)$
$\rightarrow \zeta_1 = \lambda x + 1 - \lambda$ and $\zeta_2 = \lambda y + 2 - 2\lambda$ for some $\lambda \in (0,1)$.

b. $Df(x,y) = [-\frac{y}{(1+x)^2} \;\; \frac{1}{1+x}] \rightarrow Hf(x,y) = \begin{bmatrix} \frac{2y}{(1+x)^3} & -\frac{1}{(1+x)^2} \\ -\frac{1}{(1+x)^2} & 0 \end{bmatrix}$
$\rightarrow Df(0,0) = [0 \;\; 1] \rightarrow$ Taylor expansion:

$$f(x,y) = y + \frac{1}{2}[x \; y] \begin{bmatrix} \frac{2\zeta_2}{(1+\zeta_1)^3} & -\frac{1}{(1+\zeta_1)^2} \\ -\frac{1}{(1+\zeta_1)^2} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = y + \frac{\zeta_2}{(1+\zeta_1)^3}x^2 - \frac{1}{(1+\zeta_1)^2}xy,$$

where $\begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix} + (1-\lambda) \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ for some $\lambda \in (0,1)$
$\rightarrow \zeta_1 = \lambda x$ and $\zeta_2 = \lambda y$ for some $\lambda \in (0,1)$.

At $(1,2) \rightarrow Df(1,2) = [-\frac{1}{2} \;\; \frac{1}{2}] \rightarrow$ Taylor expansion:

$$\begin{aligned} f(x,y) &= 1 - \frac{1}{2}(x-1) + \frac{1}{2}(y-2) + \frac{1}{2}[x-1 \; y-2] \begin{bmatrix} \frac{2\zeta_2}{(1+\zeta_1)^3} & -\frac{1}{(1+\zeta_1)^2} \\ -\frac{1}{(1+\zeta_1)^2} & 0 \end{bmatrix} \begin{bmatrix} x-1 \\ y-2 \end{bmatrix} \\ &= 1 - \frac{1}{2}(x-1) + \frac{1}{2}(y-2) + \frac{\zeta_2}{(1+\zeta_1)^3}(x-1)^2 - \frac{1}{(1+\zeta_1)^2}(x-1)(y-2), \end{aligned}$$

where $\begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix} + (1-\lambda) \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ for some $\lambda \in (0,1)$
$\rightarrow \zeta_1 = \lambda x + 1 - \lambda$ and $\zeta_2 = \lambda y + 2 - 2\lambda$ for some $\lambda \in (0,1)$.

c. $Df(x,y) = [2x+y \;\; x+4y] \rightarrow Hf(x,y) = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$
$\rightarrow Df(0,0) = [0 \;\; 0] \rightarrow$ Taylor expansion:

$$f(x,y) = 1 + \frac{1}{2}[x \; y] \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 1 + x^2 + xy + 2y^2.$$

At $(1,2) \rightarrow Df(1,2) = [4 \;\; 9] \rightarrow$ Taylor expansion:

$$\begin{aligned} f(x,y) &= 12 + 4(x-1) + 9(y-2) + \frac{1}{2}[x-1 \; y-2] \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x-1 \\ y-2 \end{bmatrix} \\ &= 12 + 4(x-1) + 9(y-2) + (x-1)^2 + (x-1)(y-2) + 2(y-2)^2 = 1 + x^2 + xy + 2y^2. \end{aligned}$$

EXERCISE 2.47 *(Taylor's theorem for a vector-valued twice differentiable function)*
Let $f_j : \mathbb{R}^n \to \mathbb{R},\ j = 1, \cdots, m,$ be twice continuously differentiable on a neighborhood S of a. Assume, moreover, the second-order derivatives of f_j are bounded on S, i.e. there exists a scalar d such that $-dI \leq Hf_j(x) \leq dI$, for all $x \in S$ and $j = 1, \cdots, m$.
Let $f := [f_1 \cdots f_m]^T : \mathbb{R}^n \to \mathbb{R}^m$. Consider the notation $f(x) = O(x)$ as $x \to 0$, which is defined as that there exist scalars c and ϵ such that $\|f(x)\| \leq c\|x\|$ when $0 < \|x\| \leq \epsilon$.
Then Taylor's theorem states that, with $f^{'}(a) := \begin{bmatrix} Df_1(a) \\ \vdots \\ Df_m(a) \end{bmatrix}$:

$$f(x) = f(a) + f^{'}(a)(x-a) + O(\|x-a\|^2). \tag{2.2}$$

In this exercise you are asked to provide a proof of this result.
a. Show that for every f_j we can find a point ζ_j that is located on the line segment between x and a such that,

$$f_j(x) = f_j(a) + f_j^{'}(a)(x-a) + \frac{1}{2}(x-a)^T Hf_j(\zeta_j)(x-a).$$

b. Show that for every f_j, $|f_j(x) - (f_j(a) + f_j^{'}(a)(x-a))| \leq \frac{1}{2}d\|x-a\|^2$, for all $x \in S$.
c. Show that for all $x \in S$, $\|f(x) - (f(a) + f^{'}(a)(x-a))\| \leq \frac{1}{2}\sqrt{m}d\|x-a\|^2$. Conclude (2.2).

ANSWER 2.47
a. This follows immediately from Taylor's theorem, Exercise 2.45.
b. Notice, since f_j is twice continuously differentiable Hf_j is symmetric. Consequently, by item a, $|f_j(x) - (f_j(a) + f_j^{'}(a)(x-a))| = |\frac{1}{2}(x-a)^T Hf_j(\zeta_j)(x-a)| \leq \frac{1}{2}d(x-a)^T(x-a) = \frac{1}{2}d\|x-a\|^2$, for all j and $x \in S$.
c. Using item b we have: $\|f(x) - (f(a) + f^{'}(a)(x-a))\|^2 = \sum_{i=1}^m (f_j(x) - (f_j(a) + f_j^{'}(a)(x-a)))^2 \leq \sum_{i=1}^m \frac{1}{4}d^2\|x-a\|^4$. So, $\|f(x) - (f(a) + f^{'}(a)(x-a))\| \leq \frac{1}{2}\sqrt{m}d\|x-a\|^2$.

EXERCISE 2.48 *(Newton-Raphson's algorithm to find a stationary point of a multi-variable function)*
Assume the partial derivatives of a function $f : \mathbb{R}^n \to \mathbb{R}$ are twice continuously differentiable on a neighborhood S of a and f has a stationary point at $x = a$. Assume, moreover, that $H(D_j f),\ j = 1, \cdots, n,$ are bounded on S (i.e., there exists a d such that $-dI \leq H(D_j f(x)) \leq dI$ on S for all j). In this exercise we consider the multi-variable extension of Newton-Raphson's algorithm

$$x_{N+1} = x_N - Hf^{-1}(x_N)Df^T(x_N),\ x_0 = \bar{x} \in S, \tag{2.3}$$

to produce successively better approximations of this stationary point a.
a. Show that if $x_N \in S$ is an initial guess for the stationary point a and $x_{N+1} \in S$ an update of this guess, $Df^T(x_{N+1}) = Df^T(x_N) + Hf(x_N)(x_{N+1} - x_N) + O(\|x_{N+1} - x_N\|^2$.

b. Assume we are looking for an update of x_N that has the property that $Df^T(x_{N+1}) = 0$ and $Hf(x)$ is invertible on S. Show that the update (2.3) makes sense.

c. Consider the Newton-Raphson approximations (2.3) of the stationary point a.
Let c be the absolute smallest eigenvalue of $Hf(x)$ on S (which implies that $Hf^{-1}(x)Hf^{-1}(x) \leq c^2 I$ for all $x \in S$). Assume $Hf(x)$ satisfies a so-called Lipschitz condition $\|Hf(x) - Hf(y)\| \leq L\|x - y\|$ on S (with $L > 0$)[1].
Show that with $\epsilon_N := \|x_N - a\|$, this update error satisfies: $\epsilon_{N+1} \leq \frac{1}{c}(L + \frac{1}{2}\sqrt{n}d)\epsilon_N^2$, with $\epsilon_0 = \|\bar{x} - a\|$.

d. Let $\tilde{\epsilon}_N := \frac{1}{c}(L + \frac{1}{2}\sqrt{n}d)\epsilon_N$. Show that $\tilde{\epsilon}_{N+1} \leq \tilde{\epsilon}_N^2$ with $\tilde{\epsilon}_0 = \frac{1}{c}(L + \frac{1}{2}\sqrt{n}d)\|\bar{x} - a\|$.

e. Use item d to conclude that $\tilde{\epsilon}_{N+1} \leq \tilde{\epsilon}_0^{2^N}$.
Assume $\frac{1}{c}(L+\frac{1}{2}\sqrt{n}d)\epsilon_0 < 1$. Show that the Newton-Raphson algorithm converges to the stationary point a.

f. Formulate a recursive Newton-Raphson procedure that will provide the location (x, y) at which the function $f(x, y) = x^4 - 4x + y^4 - 32y$ attains its minimum.

ANSWER 2.48

a. Consider $g(x) := [D_1 f(x) \cdots D_n f(x)]^T$. Then all conditions mentioned in Exercise 2.47 are met and we deduce from Taylor's Theorem (2.2) that: $g(x_{N+1}) = g(x_N) + g'(x_N)(x_{N+1} - x_N) + O(\|x_{N+1} - x_N\|^2)$. From which the advertized result follows immediately.

b. From item a we have: $Df^T(x_{N+1}) = Df^T(x_N) + Hf(x_N)(x_{N+1} - x_N) + O(\|x_{N+1} - x_N\|^2)$. So, if x_N is already close to the stationary point a of f, and we are looking for an update of x_N which is close to a, it makes sense to choose x_{N+1} such that $0 = Df^T(x_N) + Hf(x_N)(x_{N+1} - x_N)$. This, since one may expect that the distance between x_{N+1} and x_N will be small.

c. With $x_{N+1} = x_N - Hf^{-1}(x_N)Df^T(x_N)$ we have that $x_{N+1} - a = x_N - a - Hf^{-1}(x_N)Df^T(x_N)$. Or, with $\Delta_N := x_N - a$,

$$\begin{aligned}
\Delta_{N+1} &= \Delta_N - Hf^{-1}(x_N)Df^T(x_N) \\
&= \Delta_N - Hf^{-1}(x_N)Df^T(a + \Delta_N) \\
&= Hf^{-1}(x_N)(Hf(x_N)\Delta_N - Df^T(a + \Delta_N)) \\
&= Hf^{-1}(x_N)(Hf(x_N)\Delta_N - Hf(a)\Delta_N + Hf(a)\Delta_N - Df^T(a + \Delta_N)) \quad (i).
\end{aligned}$$

By Taylor's Theorem (2.2) we have $Df^T(a + \Delta_N) = Df^T(a) + Hf(a)\Delta_N + O(\|\Delta_N\|^2) = Hf(a)\Delta_N + O(\|\Delta_N\|^2)$. Or, more precisely (see item c Exercise 2.47):

$$\|Df^T(a + \Delta_N) - Hf(a)\Delta_N\| \leq \frac{1}{2}\sqrt{n}d\|\Delta_N\|^2, \text{ for all } x_N \in S. \quad (ii)$$

1 Here we use the 2-norm as matrix norm. That is $\|A\|$ is the largest singular value of matrix A, i.e., $\|A\|$ is the square root of the largest eigenvalue of matrix $A^T A$. Using this matrix norm one also easily verifies the inequality $\|Ax\| \leq \|A\|\|x\|$ applies.

With $z := Hf(x_N)\Delta_N - Hf(a)\Delta_N + Hf(a)\Delta_N - Df^T(a+\Delta_N)$ we have from (i): $\|\Delta_{N+1}\|^2 = z^T Hf^{-1}(x_N)Hf^{-1}(x_N)z \leq \frac{1}{c^2} z^T z = \frac{1}{c^2}\|z\|^2$
$\leq \frac{1}{c^2}(\|(Hf(x_N) - Hf(a))\Delta_N\| + \|Hf(a)\Delta_N - Df^T(a+\Delta_N)\|)^2$. Where we used the triangle inequality to arrive at the last inequality.
Now, $\|(Hf(x_N) - Hf(a))\Delta_N\| \leq \|(Hf(x_N) - Hf(a))\|\|\Delta_N\| \leq L\|x_N - a\|\|\Delta_N\| = L\|\Delta_N\|^2$. Furthermore, by (ii), $\|Df^T(a+\Delta_N) - Hf(a)\Delta_N\| \leq \frac{1}{2}\sqrt{n}d\|\Delta_N\|^2$.
So we conclude that $\|\Delta_{N+1}\|^2 \leq \frac{1}{c^2}(L\|\Delta_N\|^2 + \frac{1}{2}\sqrt{n}d\|\Delta_N\|^2)^2$. Or: $\|\Delta_{N+1}\| \leq \frac{1}{c}(L + \frac{1}{2}\sqrt{n}d)\|\Delta_N\|^2$. Which yields the advertized inequality.

d. By definition: $\tilde{\epsilon}_{N+1} = \frac{1}{c}(L + \frac{1}{2}\sqrt{n}d)\epsilon_{N+1} \leq \frac{1}{c^2}(L + \frac{1}{2}\sqrt{n}d)^2\epsilon_N = \tilde{\epsilon}_N^2$.

e. Using the inequality from item d recursively we get: $|\tilde{\epsilon}_{N+1}| \leq \tilde{\epsilon}_0^{2^N}$. Or: $|\epsilon_{N+1}| \leq \frac{(\frac{1}{c}(L+\frac{1}{2}\sqrt{n}d)\epsilon_0)^{2^N}}{\frac{1}{c}(L+\frac{1}{2}\sqrt{n}d)}$. So, if $\frac{1}{c}(L + \frac{1}{2}\sqrt{n}d)\epsilon_0 < 1$, ϵ_N converges exponentially to zero. Or, stated differently, the algorithm yields estimates that converge exponentially to a.

f. With $f(x,y) = x^4 - 4x + y^4 - 32y$, we have $f'(x,y) = [4x^3 - 4 \;\; 4y^3 - 32]$. From this one easily concludes that f has a unique stationary point that is located within the set $[\frac{1}{2}, 2] \times [1, 3]$.
Furthermore, $Hf(x,y) = \begin{bmatrix} 12x^2 & 0 \\ 0 & 12y^2 \end{bmatrix}$. Consequently,
$\|Hf(x_1,y_1) - Hf(x_2,y_2)\| = \| \begin{bmatrix} 12(x_1^2 - x_2^2) & 0 \\ 0 & 12(y_1^2 - y_2^2) \end{bmatrix} \|$. So, the square root of the largest eigenvalue of $(H(x_1,y_1) - H(x_2,y_2))^T(H(x_1,y_1) - H(x_2,y_2))$ is $12\max\{|x_1^2 - x_2^2|, |y_1^2 - y_2^2|\}$. Therefore, if $S \subset [\frac{1}{2}, 2] \times [1, 3]$, $\|H(x_1,y_1) - H(x_2,y_2)\| \leq L\|(x_1,y_1) - (x_2,y_2)\|$ if e.g. $L = 12 * 9 = 108$. Moreover, $Hf(x,y) \geq \frac{1}{4}I$ on this set.
Next, let $g_1(x,y) := D_1 f(x,y) = 4x^3 - 4$ and $g_2(x,y) := D_2 f(x,y) = 4y^3 - 32$. Notice $g_1' = [12x^2 \; 0]$, $Hg_1(x,y) = \begin{bmatrix} 24x & 0 \\ 0 & 0 \end{bmatrix}$, $g_2' = [0 \; 12y^2]$ and $Hg_2(x,y) = \begin{bmatrix} 0 & 0 \\ 0 & 24y \end{bmatrix}$. So, if $S \subset [\frac{1}{2}, 2] \times [1, 3]$, $Hg_i(x,y) \leq 72I$ on S.
Consequently, $L + \frac{1}{2}\sqrt{m}d = 108 + 36\sqrt{2}$. Notice the stationary point of f is $(1, 2)$. Therefore, if we start the Newton-Raphson algorithm
$(x_{N+1}, y_{N+1})^T = (x_N, y_N)^T - \begin{bmatrix} 12x_N^2 & 0 \\ 0 & 12y_N^2 \end{bmatrix}^{-1} \begin{bmatrix} 4x_N^3 - 4 \\ 4y_N^3 - 32 \end{bmatrix} = (\frac{2}{3}x_N + \frac{1}{3x_N^2}, \frac{2}{3}y_N + \frac{8}{3y_N^2})^T$,
with (x_0, y_0) within a ball around $(1, 2)$ that has a radius smaller than $\frac{1}{4*(108+36\sqrt{2})}$ the iterates converge exponentially to this stationary point $(1, 2)$.
That f has a minimum at the stationary point follows from the fact that $Hf(x,y)$ is positive definite on S.

EXERCISE 2.49 *(Stationary point of a convex function is a global minimum)*
In this exercise we use Taylor's theorem (2.1) to show that if a function f is convex, twice continuous differentiable and has a stationary point at $x = a$, $f(a)$ is a global minimum.

a. Show that if f has a stationary point at $x = a$ there exists a point ζ, that is located somewhere between x and a, such that $f(x) = f(a) + \frac{1}{2}(x-a)^T Hf(\zeta)(x-a)$.

b. Use item a to show that if f is convex, $f(x) \geq f(a)$.
c. Show that $f(x,y) = (x-1)^4 + y^2 - 2y - 4$ has a global minimum at $(1,1)$.
d. Show that $f(x,y) = (x-y)^2$ has a global minimum at all points $(z,z),\ z \in \mathbb{R}$.

ANSWER 2.49

a. Substitution of $f'(a) = 0$ into (2.1) gives the result.
b. f is convex $\rightarrow$ HF is positive semi-definite on its domain $\rightarrow$ $HF(\zeta)$ is positive semi-definite $\rightarrow$ $(x-a)^T Hf(\zeta)(x-a) \geq 0 \rightarrow f(x) = f(a) + \frac{1}{2}(x-a)^T Hf(\zeta)(x-a) \geq f(a)$.
c. $Df(x,y) = [4(x-1)^3 \ \ 2y-2] \rightarrow Df(1,1) = [0\ 0]$. $Hf(x,y) = \begin{bmatrix} 12(x-1)^2 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow$
The principal matrices of Hf are the matrices

$$[12(x-1)^2];\ [2];\ \text{ and } Hf.$$

Therefore, the principal minors are: $12(x-1)^2,\ 2$, and $24(x-1)^2 \rightarrow f$ is convex $\rightarrow$ by item b., f has a global minimum at $(1,1)$.
d. Easy way: $f(x,y) \geq 0$, $f(z,z) = 0 \rightarrow$ conclusion follows.
Using item b.:
$Df(x,y) = [2(x-y) \ \ -2(x-y)] \rightarrow Df(z,z) = [0\ 0]$, for all $z \in \mathbb{R}$. $Hf(x,y) = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \rightarrow$
The principal matrices of Hf are the matrices

$$[2];\ [2];\ \text{ and } Hf.$$

Therefore, the principal minors are: $2,\ 2$, and $0 \rightarrow f$ is convex $\rightarrow$ by item b., for all $z \in \mathbb{R}$, f has a global minimum at (z,z).

EXERCISE 2.50 Assume f and g are continuous functions on $\mathbb{R}^2$. Consider the function $m(x,y) =$ minimum$\{f(x,y), g(x,y)\}$ for all $(x,y) \in \mathbb{R}^2$. Motivate why $m(x,y)$ attains both a global maximum and a global minimum on $S = \{(x,y) \mid |x| \leq 1 \text{ and } |y| \leq 1\}$.

ANSWER 2.50

$$m(x,y) = \begin{cases} f(x,y) \text{ if } (x,y) \in S_1 := \{(x,y) \mid f(x,y) < g(x,y)\}; \\ g(x,y) \text{ if } (x,y) \in S_2 := \{(x,y) \mid f(x,y) > g(x,y)\}; \\ g(x,y) \text{ if } (x,y) \in S_3 := \{(x,y) \mid f(x,y) = g(x,y)\}. \end{cases}$$

Now let (x_0, y_0) be an arbitrary element of $\mathbb{R}^2$. If $(x_0,y_0) \in S_1$, then by the continuity of f, there exists a neighborhood of (x_0,y_0) where $f(x,y) < g(x_0,y_0)$. So, on this neighborhood of (x_0,y_0), $m(x,y) = f(x,y)$, which implies that m is continuous at (x_0,y_0). A similar argument shows that m is continuous at (x_0,y_0) if $(x_0,y_0) \in S_2$.
Next consider the case $(x_0,y_0) \in S_3$. Let $\{s_n\}_{n=1}^{\infty}$ be an arbitrary sequence that converges to (x_0,y_0).

Since f is continuous on $\mathbb{R}^2$, $f(s_n)$ converges to $f(x_0, y_0)$. Similarly, since g is continuous on $\mathbb{R}^2$, $g(s_n)$ converges to $g(x_0, y_0) = f(x_0, y_0)$. Therefore,

$$m(s_n) = \min\{f(s_n), g(s_n)\} \text{ converges to } f(x_0, y_0) = \min\{f(x_0, y_0), g(x_0, y_0)\} = m(x_0, y_0).$$

Thus, m is continuous at every point $(x_0, y_0) \in S_3$.
Hence, $m(x, y)$ is continuous on S. Since, moreover, S is compact, the Weierstrass Theorem implies that $m(x, y)$ attains both a global maximum and a global minimum on S.

EXERCISE 2.51 *(Proof Theorem 2.28)*
Prove Theorem 2.28, item i) and ii), using the definition of differentiability. Hint: for item ii), add and subtract into the definition the term $f(x_0 + \Delta x, y_0 + \Delta y)g(x_0, y_0)$ and next recollect terms.

ANSWER 2.51

i) Let $z_f(x, y) = f(x_0, y_0) + D_1 f(x_0, y_0)(x - x_0) + D_2 f(x_0, y_0)(y - y_0)$ and $z_g(x, y) = g(x_0, y_0) + D_1 g(x_0, y_0)(x - x_0) + D_2 g(x_0, y_0)(y - y_0)$. Since both f and g are differentiable at (x_0, y_0), with $x := x_0 + \Delta x$ and $y := y_0 + \Delta y$, next quotients

$$\frac{f(x_0 + \Delta x, y_0 + \Delta y) - z_f(x_0 + \Delta x, y_0 + \Delta y)}{\|(x_0 + \Delta x, y_0 + \Delta y) - (x_0, y_0)\|} \text{ and}$$
$$\frac{g(x_0 + \Delta x, y_0 + \Delta y) - z_g(x_0 + \Delta x, y_0 + \Delta y)}{\|(x_0 + \Delta x, y_0 + \Delta y) - (x_0, y_0)\|}$$

approach zero if $(\Delta x, \Delta y)$ approaches $(0, 0)$.
Now consider $z_{f+g}(x, y) :=$
$f(x_0, y_0) + g(x_0, y_0) + (D_1 f(x_0, y_0) + D_1 g(x_0, y_0))(x - x_0) + (D_2 f(x_0, y_0) + D_2 g(x_0, y_0))(y - y_0)$
Notice: $z_{f+g}(x, y) = z_f(x, y) + z_g(x, y)$. Therefore,

$$\frac{(f+g)(x_0 + \Delta x, y_0 + \Delta y) - z_{f+g}(x_0 + \Delta x, y_0 + \Delta y)}{\|(x_0 + \Delta x, y_0 + \Delta y) - (x_0, y_0)\|} =$$
$$\frac{f(x_0 + \Delta x, y_0 + \Delta y) - z_f(x_0 + \Delta x, y_0 + \Delta y) + g(x_0 + \Delta x, y_0 + \Delta y) - z_g(x_0 + \Delta x, y_0 + \Delta y)}{\|(x_0 + \Delta x, y_0 + \Delta y) - (x_0, y_0)\|} =$$
$$\frac{f(x_0 + \Delta x, y_0 + \Delta y) - z_f(x_0 + \Delta x, y_0 + \Delta y)}{\|(x_0 + \Delta x, y_0 + \Delta y) - (x_0, y_0)\|} + \frac{g(x_0 + \Delta x, y_0 + \Delta y) - z_g(x_0 + \Delta x, y_0 + \Delta y)}{\|(x_0 + \Delta x, y_0 + \Delta y) - (x_0, y_0)\|}.$$

As noted above, both last terms approach zero if $(\Delta x, \Delta y)$ approaches $(0, 0) \to (f + g)$ is differentiable at (x_0, y_0) with derivative $[D_1 f(x_0, y_0) + D_1 g(x_0, y_0) \quad D_2 f(x_0, y_0) + D_2 g(x_0, y_0)] = Df(x_0, y_0) + Dg(x_0, y_0)$.

ii) Using the notation from item i., consider

$$\begin{aligned} z_{fg}(x, y) \quad = \quad & f(x_0, y_0)g(x_0, y_0) + (D_1 f(x_0, y_0)g(x_0, y_0) + D_1 g(x_0, y_0)f(x_0, y_0))(x - x_0) + \\ & (D_2 f(x_0, y_0)g(x_0, y_0) + D_2 g(x_0, y_0)f(x_0, y_0))(y - y_0). \end{aligned}$$

Then,

$$\frac{(fg)(x_0+\Delta x,y_0+\Delta y)-z_{fg}(x_0+\Delta x,y_0+\Delta y)}{\|(x_0+\Delta x,y_0+\Delta y)-(x_0,y_0)\|}=$$
$$\frac{1}{\|(\Delta x,\Delta y)\|}\{(fg)(x_0+\Delta x,y_0+\Delta y)$$
$$-f(x_0+\Delta x,y_0+\Delta y)g(x_0,y_0)+f(x_0+\Delta x,y_0+\Delta y)g(x_0,y_0)$$
$$-f(x_0,y_0)g(x_0,y_0)-(D_1f(x_0,y_0)g(x_0,y_0)+D_1g(x_0,y_0)f(x_0,y_0))\Delta x-$$
$$(D_2f(x_0,y_0)g(x_0,y_0)+D_2g(x_0,y_0)f(x_0,y_0))\Delta y\}=$$
$$f(x_0+\Delta x,y_0+\Delta y)\frac{g(x_0+\Delta x,y_0+\Delta y)-(g(x_0,y_0)+D_1g(x_0,y_0)\Delta x+D_2g(x_0,y_0)\Delta y)}{\|(\Delta x,\Delta y)\|}$$
$$+g(x_0,y_0)\frac{f(x_0+\Delta x,y_0+\Delta y)-(f(x_0,y_0)+D_1f(x_0,y_0)\Delta x+D_2f(x_0,y_0)\Delta y)}{\|(\Delta x,\Delta y)\|}$$
$$+\frac{(f(x_0+\Delta x,y_0+\Delta y)-f(x_0,y_0))(D_1g(x_0,y_0)\Delta x+D_2g(x_0,y_0)\Delta y}{\|(\Delta x,\Delta y))\|}.$$

Next note that $D_1g(x_0,y_0)\Delta x+D_2g(x_0,y_0)\Delta y=[D_1g(x_0,y_0)\ D_2g(x_0,y_0)][\Delta x\ \Delta y]^T$. Therefore, using the Cauchy-Schwartz inequality, $D_1g(x_0,y_0)\Delta x+D_2g(x_0,y_0)\Delta y\leq\|Dg(x_0,y_0)\|\|(\Delta x,\Delta y)\|$. As, moreover, f is continuous at (x_0,y_0) it follows that

$$\frac{(f(x_0+\Delta x,y_0+\Delta y)-f(x_0,y_0))(D_1g(x_0,y_0)\Delta x+D_2g(x_0,y_0)\Delta y)}{\|(\Delta x,\Delta y))\|}$$

goes to zero if $(\Delta x,\Delta y)$ approaches $(0,0)$.
Furthermore, since f and g are differentiable at (x_0,y_0) it follows also that the first two terms in the above formula approach zero if $(\Delta x,\Delta y)$ approaches $(0,0)\to(fg)$ is differentiable at (x_0,y_0) with derivative

$$[D_1f(x_0,y_0)g(x_0,y_0)+D_1g(x_0,y_0)f(x_0,y_0)\ \ D_2f(x_0,y_0)g(x_0,y_0)+D_2g(x_0,y_0)f(x_0,y_0)]=$$
$$Df(x_0,y_0)g(x_0,y_0)+Dg(x_0,y_0)f(x_0,y_0).$$

EXERCISE 2.52 *(Example of function which partial derivatives exist, but is not differentiable)*
Consider $f(x,y)=\begin{cases}x,\ \text{if } y=0;\\ 0,\ \text{if } y\neq 0.\end{cases}$.
a. Consider the section function $f(x,0)$. Show that $\frac{\partial f}{\partial x}(0,0)=1$.
b. Consider the section function $f(0,y)$. Show that $\frac{\partial f}{\partial x}(0,0)=0$.
c. Motivate why $f(x,y)$ is not differentiable at $(0,0)$.
d. Let $\epsilon\neq 0$. Motivate why $\frac{\partial f}{\partial y}(\epsilon,0)$ does not exist.
e. Motive why Theorem 2.26 does not apply here.

ANSWER 2.52

a. $f(x,0) = x \rightarrow \frac{\partial f}{\partial x}(0,0) = 1$.

b. $f(0,y) = 0$ if $y \neq 0$, $f(0,0) = 0 \rightarrow f(0,y) = 0$ for all $y \in \mathbb{R} \rightarrow \frac{\partial f}{\partial x}(0,0) = 0$.

c. Candidate tangent plane: $z(x,y) = x \rightarrow$ To see whether f is differentiable at $(0,0)$ consider the quotient:

$$\frac{f(0+\Delta x, 0+\Delta y) - z(0+\Delta x, 0+\Delta y)}{\sqrt{(\Delta x)^2 + \Delta y)^2}} = \frac{f(\Delta x, \Delta y) - \Delta x}{\sqrt{(\Delta x)^2 + \Delta y)^2}}.$$

Now, consider the case that we approach $(0,0)$ along the line $y = x$. That is, $\Delta x = \Delta y$. Then, since $\Delta y \neq 0$, $f(\Delta x, \Delta y) = 0$. Therefore the above quotient reduces to

$$\frac{-\Delta x}{\sqrt{2(\Delta x)^2}} = \frac{-\Delta x}{\sqrt{2}|\Delta x|} = -\frac{1}{\sqrt{2}}\text{sign}(\Delta x).$$

Clearly, the above number does not converge to zero if Δx $(\neq 0)$ approaches zero. So, f is not differentiable at $(0,0)$.

Geometrically, this is also clear. Note that the intersection curve with f along the line $y = x$ is $f(x,x) = 0 \rightarrow$ the tangent line to this intersection curve is also a horizontal line. Clearly this horizontal line in the $y = x$ direction at $(0,0)$ is not part of the plane $z(x,y) = x$.

d. Let $\epsilon \neq 0$. Then $f(\epsilon, y) = 0$, if $y \neq 0$ and $f(\epsilon, 0) = \epsilon \rightarrow$ The intersection curve in the y direction at $(\epsilon, 0)$ is not continuous $\rightarrow$ The partial derivative w.r.t. y at $(\epsilon, 0)$ does not exist.

e. For every neighborhood of $(0,0)$ there exist points where the partial derivative of f w.r.t. y does not exist. Therefore there does not exist a neighborhood of $(0,0)$ such that, at every point (x_0, y_0) in this neighborhood, the partial derivative of f w.r.t. y is continuous.

EXERCISE 2.53 *(Example of function which intersection function on any line segment through its stationary point has a minimum, but has not a minimum)*

Consider $f(x,y) = (y - x^2)(y - 3x^2)$.

a. Show that f has a stationary point at $(0,0)$.

b. Show that $f"(0,0)$ is positive semi-definite.

c. Consider the section function $f(x, kx)$, where k is an arbitrary real number. Show that for every k, this section function has a minimum at $x = 0$.

d. Consider the section function $f(0,y)$. Show that this section function has a minimum at $y = 0$.

e. Show that $f(x,y)$ has a saddle point at $(0,0)$.

ANSWER 2.53

a. $Df(x,y) = [-8xy + 12x^3 \;\; 2y - 4x^2] \rightarrow Df(0,0) = [0 \;\; 0]$.

b. $Hf(x,y) = \begin{bmatrix} -8y + 36x^2 & -8x \\ -8x & 2 \end{bmatrix} \rightarrow Hf(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow Hf(0,0)$ is positive semi-definite.

c. $f(x,kx) = k^2x^2 - 4kx^3 + 3x^4 \rightarrow f^{'}(x,kx) = 2k^2x - 12kx^2 + 12x^3 \rightarrow f^{'}(x,kx) = 0$ at $x = 0$. Furthermore, $f^{"}(x,kx) = 2k^2 - 24kx + 36x^2 \rightarrow f^{"}(x,kx) > 0$ at $x = 0$, if $k \neq 0 \rightarrow f(x,kx)$ has a minimum at $x = 0$, if $k \neq 0$. Moreover, if $k = 0$, $f(x,kx) = 3x^4 \rightarrow f(x,kx)$ has a minimum at $x = 0$ in that case too.

d. $f(0,y) = y^2 \rightarrow f(0,y)$ has a minimum at $y = 0$.

e. Consider the intersection curve $f(x,2x^2) \rightarrow f(x,2x^2) = -x^4 \rightarrow f(x,2x^2) \leq 0$ and has a maximum at $x = 0$ on this intersection curve.
$\rightarrow$ In every neighborhood of $(0,0)$ there exists a point $(x,2x^2)$, where $f < 0$, and a point (x,x) where $f > 0 \rightarrow f$ has a saddle point at $(0,0)$.

EXERCISE 2.54 *(Example of function which is continuous but which partial derivatives don't exist)*

Consider $f(x,y) = \begin{cases} |x|, \text{ if } y = 0; \\ |y|, \text{ if } x = 0; \\ 0, \text{ elsewhere.} \end{cases}$.

a. Consider the section function $f(x,0)$. Show that $\frac{\partial f}{\partial x}(0,0)$ does not exist.

b. Consider the section function $f(0,y)$. Show that $\frac{\partial f}{\partial x}(0,0)$ does not exist.

c. Motivate why $f(x,y)$ is continuous at $(0,0)$.

ANSWER 2.54

a. $f(x,0) = |x|$ is not differentiable at $x = 0 \rightarrow \frac{\partial f}{\partial x}(0,0)$ does not exist.

b. $f(0,y) = |y|$ is not differentiable at $y = 0 \rightarrow \frac{\partial f}{\partial y}(0,0)$ does not exist.

c. Let $\{s_n\}_{n=1}^{\infty}$ with $s_n = (x_n, y_n)$ be an arbitrary sequence that converges to $(0,0)$. Then, either $f(s_n) = |x_n|$ or $f(s_n) = |y_n|$. However, in both cases $f(s_n)$ converges to $0 = f(0,0)$ if s_n converges to $(0,0) \rightarrow f$ is continuous at $(0,0)$.

EXERCISE 2.55 *(Example of function that is differentiable, but which partial derivatives are not continuous)*

Consider $f(x,y) = (x^2y^2)^{\frac{1}{3}}$..

a. Show that both partial derivatives of $f(x,y)$ exist at $(0,0)$.

b. Calculate $D_1f(x_0,y_0)$ at all (x_0,y_0) for which $x_0y_0 \neq 0$.

c. Show that $D_1f(x,y)$ is not continuous at $(0,0)$.

d. Show that $f(x,y)$ is differentiable at $(0,0)$.

ANSWER 2.55

a. $f(x,0) = 0$ is differentiable at $x = 0 \rightarrow \frac{\partial f}{\partial x}(0,0) = 0$.
$f(0,y) = 0$ is differentiable at $y = 0 \rightarrow \frac{\partial f}{\partial y}(0,0) = 0$.

b. $f(x,y_0) = y_0^{\frac{2}{3}}x^{\frac{2}{3}}$ is differentiable at $x = x_0 \neq 0 \rightarrow \frac{\partial f}{\partial x}(x_0,y_0)$ exists and equals $\frac{2}{3}y_0^{\frac{2}{3}}x_0^{\frac{-1}{3}}$.

c. With $D_1(0,0) = 0$ and $D_1(x,y) = \frac{2}{3}y^{\frac{2}{3}}x^{\frac{-1}{3}}$ it follows that if we approach $(0,0)$ along the curve $s := (x, x^{\frac{1}{6}})$, $D_1(s) = \frac{2}{3}$. So $D_1(s)$ does not approach 0 if s approaches zero. So, $D_1(x,y)$ is not continuous at $(0,0)$.

d. The tangent plane to the graph of f at $(0,0)$ is $z(x,y) = 0$. So, $\lim_{(x,y)\to(0,0)} \frac{f(x,y)-z(x,y)}{\|(x,y)-(0,0)\|} = \lim_{(x,y)\to(0,0)} \frac{x^{\frac{2}{3}}y^{\frac{2}{3}}}{\sqrt{x^2+y^2}} = 0$. Where the last equality can be obtained from the fact that $|xy| \leq \frac{1}{2}(x^2 + y^2)$.

EXERCISE 2.56 *(Extension of notion of differentiability)*

Let $f(x)$ be defined on a subset C of $\mathbb{R}^n$ and $x_0 \in C$. Then f is said to be *differentiable at* x_0 if there exists a vector $a \in \mathbb{R}^n$ (where a may depend on x_0) such that for all vectors $h \neq 0$ such that $x_0 + h \in C$

$$\lim_{h\to 0} \frac{f(x_0+h) - f(x_0) - a^T h}{\|h\|} = 0.$$

If this limit exists, a is called the derivative of f at x_0.

Notice there are no restrictions here on the set C and x_0.

a. Consider $f(x,y) = |y|$ on the constraint set $C = \{(x,y) \mid y = x^2\}$. Show that f is differentiable at $(0,0)$ on C.

b. Consider $f(x,y) = |y|$ on the constraint set $C = \{(x,y) \mid x = y^2\}$. Show that f is not differentiable at $(0,0)$ on C.

c. Consider $f(x,y) = |y|$ on the constraint set $C = \{(x,y) \mid y = |x|\}$. Show that f is not differentiable at $(0,0)$ on C.

d. Consider $f(x,y) = y$ on the constraint set $C = \{(x,y) \mid x^2 = y^2\}$. Show that f is not differentiable at $(0,0)$ on C.

ANSWER 2.56

a. Notice $f(x,y)$ coincides on C with the function $\tilde{f}(x) = x^2,\ x \in \mathbb{R}$. Obviously, $\tilde{f}(x)$ is differentiable at $x = 0$. So, f is differentiable at $(0,0)$.

b. Notice that on one branch of C, $f(x, \sqrt{x}) = \sqrt{x}$, whereas on the other branch of C, $f(x, -\sqrt{x}) = \sqrt{x}$. Clearly $\tilde{f}(x) = \sqrt{x}$ is not differentiable at $x = 0$. So, f is not differentiable at $(0,0)$. Obviously, $\tilde{f}(x)$ is differentiable at $x = 0$. So, f is differentiable at $(0,0)$.

c. Notice that on the branch $(x, -x),\ x < 0$, $f(x,y)$ coincides with the function $\tilde{f}(x) = -x,\ x < 0$, whereas on the other branch of C, $(x,x),\ x > 0$, $f(x,y)$ coincides with $\tilde{f}(x) = x,\ x > 0$. Clearly $\tilde{f}$ is not differentiable at $x = 0$. So, f is not differentiable at $(0,0)$.

d. Notice, e.g., that on the branch $(x, -x),\ x < 0$, $f(x,y)$ coincides with the function $\tilde{f}(x) = -x,\ x < 0$, whereas on the branch of C, $(x,x),\ x > 0$, $f(x,y)$ coincides with $\tilde{f}(x) = x,\ x > 0$. Clearly $\tilde{f}$ is not differentiable at $x = 0$. So, f is not differentiable at $(0,0)$.

EXERCISE 2.57 *(Brouwer's Fixed Point Theorem)*

Let $S \subset \mathbb{R}^2$ be a compact, convex set (nonempty). Assume $f_i : S \to \mathbb{R},\ i = 1, 2$, are continuous

functions. Consider the range of f defined as: $\{(f_1(x,y), f_2(x,y)) \in \mathbb{R}^2 \mid (x,y) \in S\}$.
Then Brouwer's Fixed Point Theorem states that if the range of f is a subset of S, there exists $(\bar{x}, \bar{y}) \in S$ such that $f_1(\bar{x}, \bar{y}) = \bar{x}$ and $f_2(\bar{x}, \bar{y}) = \bar{y}$. $(\bar{x}, \bar{y})$ is called a fixed point of f in that case. (This result generalizes for n functions which domain S is a compact, convex set in $\mathbb{R}^n$).
Unfortunately, currently, a proof using elementary arguments we developed so far of this result is lacking in literature. Below, you are asked to show one of the many applications Brouwer's theorem has. Here we consider the issue to find a stationary point.

a. Let S be a compact, convex set in $\mathbb{R}^2$. Show that a continuous differentiable function $f : S \to \mathbb{R}$ has a stationary point $(\bar{x}, \bar{y})$ if and only if the function $g(x,y) := (D_1 f(x,y) + x, D_2 f(x,y) + y)$ has a fixed point on S.
b. Consider the function $f(x,y) = \sqrt{1 + \frac{1}{2}x^2y^2} + e^{-\frac{1}{4}(x^2+y^2)} - \frac{1}{2}(x^2+y^2)$. Show that f has a stationary point within the set $S = \{(x,y) \mid |x| \leq 1,\ |y| \leq 1\}$.

ANSWER 2.57

a. Notice that, since by assumption f is continuous differentiable, $D_i f$ are continuous, and therefore both $g_1(x,y) := D_1 f(x,y) + x$ and $g_2(x,y) := D_2 f(x,y) + y$ are continuous.
" $\Rightarrow$ " assume f has a stationary point $(\bar{x}, \bar{y}) \in S$, i.e., $D_i f(\bar{x}, \bar{y}) = 0 \to D_1 f(\bar{x}, \bar{y}) + \bar{x} = \bar{x}$ and $D_2 f(\bar{x}, \bar{y}) + \bar{y} = \bar{y}$, for $(\bar{x}, \bar{y}) \in S \to g(x,y)$ has a fixed point on S.
" $\Leftarrow$ " If the function $g(x,y) := (D_1 f(x,y) + x, D_2 f(x,y) + y)$ has a fixed point on S, by definition there is $(\bar{x}, \bar{y}) \in S$ such that $D_1 f(\bar{x}, \bar{y}) + \bar{x} = \bar{x}$ and $D_2 f(\bar{x}, \bar{y}) + \bar{y} = \bar{y} \to D_1 f(\bar{x}, \bar{y}) = 0$ and $D_2 f(\bar{x}, \bar{y}) = 0 \to f$ has a stationary point $(\bar{x}, \bar{y}) \in S$.
b. $Df = [\frac{\frac{1}{2}xy^2}{\sqrt{1+x^2y^2}} - \frac{1}{2}xe^{-\frac{1}{4}(x^2+y^2)} - x \quad \frac{\frac{1}{2}x^2y}{\sqrt{1+x^2y^2}} - \frac{1}{2}ye^{-\frac{1}{4}(x^2+y^2)} - y]$. Note that the range of $[D_1 f + x \quad D_2 f + y](S) \subset S$. So, by Brouwer's theorem f has a stationary point on this set S.

EXERCISE 2.58 *(Fixed Point Contraction Mapping Theorem)*
Let $S \subset \mathbb{R}^n$ be a (nonempty) closed convex set and $f_i : S \to \mathbb{R}$, $i = 1, \cdots, n$. Consider the function $f := (f_1, \cdots, f_n)$ on S. Assume $f(x) \in S$ for all $x \in S$ and, moreover, there exists a constant $0 \leq c < 1$ such that for all $x, y \in S$, $\|f(x) - f(y)\| \leq c\|x - y\|$.
Then the Contraction Mapping Theorem states that under these conditions f has a *unique* fixed point $\bar{x} \in S$.
Furthermore, $\bar{x}$ can be found as follows: start with an arbitrary element $x_0 \in S$ and define a sequence $\{s_n\}_{n=1}^{\infty}$ by $s_n = f(s_{n-1})$, then s_n converges to $\bar{x}$.
In items a. and b., below, you are asked to provide a proof of this result under the additional assumption that S is bounded and f is continuous.

a. Assume f is continuous and S compact. Show that f has a fixed point on S.
b. Show that f has a unique fixed point on S. Hint: assume there is another one.
c. Let $x \in S$. Consider the sequence $\{s_n\}_{n=1}^{\infty}$ defined by: $s_1 = x$ and $s_{n+1} = f(s_n),\ n = 2, 3, ...$ Show that $\|f(s_n) - s_n\| \leq c^n \|s_2 - s_1\|$. (Using this, one can show that this implies that f must have a

fixed point, i.e., that the Contraction Mapping Theorem applies also without assuming continuity of f and compactness of S).

d. Consider the function $f(x,y) = (\frac{1}{2}y + 3 \;\; \frac{1}{2}x)$. Show, using the Contraction Mapping Theorem, that f has a unique fixed point on $\mathbb{R}^2$. Consider $s_1 := (0,0)$. Calculate $s_{n+1} = f(s_n)$, for $n = 2,3,4,5$. Calculate the fixed point.

e. Show, using the Contraction Mapping Theorem, that $f(x,y) = \frac{1}{2}\ln(2 + x^2 + y^2) - \frac{1}{2}(x^2 + y^2)$ has a unique stationary point. Hint: consider the approach used in Exercise 2.57, item a. Use the Mean Value Theorem to estimate the involved partial derivatives.

ANSWER 2.58

a. Note: i. $f(x)$ is continuous S; ii. for all $x \in S$, $f(x) \in S \rightarrow$ Brouwer's theorem: f has a fixed point on S.

b. Assume f has two fixed points $\bar{x}_1$ and $\bar{x_2}$. Then

$$\|\bar{x}_1 - \bar{x}_2\| = \|f(\bar{x}_1) - f(\bar{x}_2)\| \leq c\|\bar{x}_1 - \bar{x}_2\| < \|\bar{x}_1 - \bar{x}_2\|.$$

This is a contradiction. Therefore, the assumption is false, and there must be a unique fixed point.

c. Note that $s_{n+1} = f(s_n)$. Then

$$\|f(s_n) - s_n\| = \|s_{n+1} - s_n\| = \|f(s_n) - f(s_{n-1})\| \leq c\|s_n - s_{n-1}\|$$

Applying this estimate recursively yields

$$\|s_{n+1} - s_n\| \leq c\|s_n - s_{n-1}\| \leq c^2\|s_{n-1} - s_{n-2}\| \leq ... \leq c^n\|s_2 - s_1\|.$$

d. With $f(x,y) = (\frac{1}{2}y + 3 \;\; \frac{1}{2}x))$ we have

$$\begin{aligned} \|f(x_1,y_1) - f(x_2,y_2)\| &= \|(\frac{1}{2}(y_1 - y_2) \;\; \frac{1}{2}(x_1 - x_2))\| = \frac{1}{2}\sqrt{(y_1 - y_2)^2 + (x_1 - x_2)^2} \\ &= \frac{1}{2}\|(x_1 - x_2 \;\; y_1 - y_2)\| = \frac{1}{2}\|(x_1 \; y_1) - (x_2 \; y_2)\|. \end{aligned}$$

Thus, with $c = \frac{1}{2}$, the Contraction Mapping Theorem applies, and there is a unique fixed point. With $s_{n+1} := f(s_n)$ and $s_1 = (0 \; 0)$, the first five elements of the sequence are:

$$s_2 = (3 \; 0);\; s_3 = (3 \; \frac{3}{2});\; s_4 = (\frac{15}{4} \; \frac{3}{2});\; s_5 = (\frac{15}{4} \; \frac{15}{8}).$$

Solving the system $\frac{1}{2}y + 3 = x$ and $\frac{1}{2}x = y$ gives the fixed point $(4,2)$.

e. First, note that $Df = [\frac{x}{2+x^2+y^2} - x \;\; \frac{y}{2+x^2+y^2} - y]$.
Following the approach of Exercise 2.57, item a., we observe that f has a stationary point $(\bar{x},\bar{y})$ if and only if the function

$$g(x,y) = (D_1 f(x,y) + x, D_2 f(x,y) + y)$$

has a fixed point $(\bar{x}, \bar{y})$. That is,

$$g(x,y) = (\frac{x}{2+x^2+y^2}, \frac{y}{2+x^2+y^2})$$

must have a fixed point.
Note that $g \in C^1$. By the Mean Value Theorem, there exists a point (ξ_1, ξ_2) on the line segment between (x_1, y_1) and (x_2, y_2) such that

$$g_1(x_1,y_1) - g_1(x_2,y_2) = \frac{x_1}{2+x_1^2+y_1^2} - \frac{x_2}{2+x_2^2+y_2^2} = [D_1 g_1(\xi_1,\xi_2) \;\; D_2 g_1(\xi_1,\xi_2)][x_1 - x_2 \;\; y_1 - y_2]^T.$$

By the Cauchy-Schwartz inequality,

$$[D_1 g_1(\xi_1,\xi_2) \;\; D_2 g_1(\xi_1,\xi_2)][x_1 - x_2 \;\; y_1 - y_2]^T \leq \|[D_1 g_1(\xi_1,\xi_2) \;\; D_2 g_1(\xi_1,\xi_2)]\| \|[x_1 - x_2 \;\; y_1 - y_2]\|.$$

Now, $D_1 g_1(\xi_1,\xi_2) = \frac{2-\xi_1^2+\xi_2^2}{(2+\xi_1^2+\xi_2^2)^2}$, so $|D_1 g_1(\xi_1,\xi_2)| \leq \frac{2+\xi_1^2+\xi_2^2}{(2+\xi_1^2+\xi_2^2)^2} \leq \frac{1}{2}$.
Similarly, $|D_2 g_1(\xi_1,\xi_2)| = |\frac{-2\xi_1\xi_2}{(2+\xi_1^2+\xi_2^2)^2}| \leq \frac{1}{4}$.
Hence,

$$\|[D_1 g_1(\xi_1,\xi_2) \;\; D_2 g_1(\xi_1,\xi_2)]\| \leq \sqrt{(\frac{1}{2})^2 + (\frac{1}{4})^2} = \frac{\sqrt{5}}{4}.$$

Therefore,

$$|g_1(x_1,y_1) - g_1(x_2,y_2)| \leq \frac{\sqrt{5}}{4} \|[x_1 - x_2 \;\; y_1 - y_2]\|.$$

By a similar argument,

$$|g_2(x_1,y_1) - g_2(x_2,y_2)| \leq \frac{\sqrt{5}}{4} \|[x_1 - x_2 \;\; y_1 - y_2]\|.$$

It follows that

$$\begin{aligned} \|g(x_1,y_1) - g(x_2,y_2)\| &= \sqrt{(g_1(x_1,y_1) - g_1(x_2,y_2))^2 + (g_2(x_1,y_1) - g_2(x_2,y_2))^2} \\ &\leq \sqrt{(\frac{\sqrt{5}}{4}\|[x_1 - x_2 \;\; y_1 - y_2]\|)^2 + (\frac{\sqrt{5}}{4}\|[x_1 - x_2 \;\; y_1 - y_2]\|)^2} \\ &= \frac{\sqrt{10}}{4}\|[x_1 - x_2 \;\; y_1 - y_2]\|. \end{aligned}$$

Thus, g satisfies the conditions of the Contraction Mapping Theorem. Therefore, g has a unique fixed point, and consequently f has a unique stationary point.

Exercises with an Economic Content

EXERCISE 2.59 *(generalized Cobb-Douglas production function)*
Show that the production function $F(x_1, x_2, x_3) = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}$, with $0 < \alpha_i < 1$ is concave on its domain $D = \{(x_1, x_2, x_3) \mid x_1 > 0,\ x_2 > 0,\ x_3 > 0\}$.

ANSWER 2.59
Straightforward calculations show that

$$D_{ii}F(x_1, x_2, x_3) = \alpha_i(\alpha_i - 1)\frac{F(x_1, x_2, x_3)}{x_i^2} \text{ and } D_{ij}F(x_1, x_2, x_3) = \alpha_i\alpha_j\frac{F(x_1, x_2, x_3)}{x_i x_j}.$$

Using these expressions, further (though somewhat lengthy) computations show that all principal minors of $-HF(x_1, x_2, x_3)$ are nonnegative. Hence, by Theorem 2.45, F is concave on its domain.

EXERCISE 2.60 A country i considers entering an economic union with N members (including the country itself). Welfare of country i equals $U_i = y_i - g_i + \alpha_i h(g_i + \beta \sum_{j \neq i} g_j)$, where $y_i - g_i$ is the amount of private consumption and g_i is the local public good level in country i. Here, $h(x)$ is an increasing function. The *first-best union* is the union where the sum of welfare of the individual countries is maximized.

a. Determine the first-order conditions for the public good for the first-best union if $N = 3$.

b. Let $h(x) = \ln(x)$. Solve the first-order conditions you found in item a. Hint: introduce $x_i = \frac{\alpha_i}{g_i + \beta \sum_{j \neq i} g_j}$ and rewrite the system of equations in terms of x_i.

c. Solve the first-order conditions in item b for an arbitrary number of countries.

ANSWER 2.60
a. The problem is to maximize

$$U(g_1, g_2, g_3) := \sum_{i=1}^{3} U_i(g_1, g_2, g_3),$$

which can be rewritten as

$$\max y_1 + y_2 + y_3 - (g_1 + g_2 + g_3) + \alpha_1 h(g_1 + \beta(g_2 + g_3)) + \alpha_2 h(g_2 + \beta(g_1 + g_3)) + \alpha_3 h(g_3 + \beta(g_1 + g_2)).$$

The stationary points are characterized by $U' = 0$, that is

$$U'^T = \begin{bmatrix} -1 + \alpha_1 h'(g_1 + \beta(g_2 + g_3)) + \alpha_2\beta h'(g_2 + \beta(g_1 + g_3)) + \alpha_3\beta h'(g_3 + \beta(g_1 + g_2)) \\ -1 + \alpha_1\beta h'(g_1 + \beta(g_2 + g_3)) + \alpha_2 h'(g_2 + \beta(g_1 + g_3)) + \alpha_3\beta h'(g_3 + \beta(g_1 + g_2)) \\ -1 + \alpha_1\beta h'(g_1 + \beta(g_2 + g_3)) + \alpha_2\beta h'(g_2 + \beta(g_1 + g_3)) + \alpha_3 h'(g_3 + \beta(g_1 + g_2)) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

b. Elementary calculations yield

$$\begin{aligned} U' &= [-1+\frac{\alpha_1}{g_1+\beta(g_2+g_3)}+\frac{\alpha_2\beta}{g_2+\beta(g_1+g_3)}+\frac{\alpha_3\beta}{g_3+\beta(g_1+g_2)}, \\ &\quad -1+\frac{\alpha_1\beta}{g_1+\beta(g_2+g_3)}+\frac{\alpha_2}{g_2+\beta(g_1+g_3)}+\frac{\alpha_3\beta}{g_3+\beta(g_1+g_2)}, \\ &\quad -1+\frac{\alpha_1\beta}{g_1+\beta(g_2+g_3)}+\frac{\alpha_2\beta}{g_2+\beta(g_1+g_3)}+\frac{\alpha_3}{g_3+\beta(g_1+g_2)}]. \end{aligned}$$

Let $x_i := \frac{\alpha_i}{g_i+\beta\sum\limits_{j\neq i} g_j}$. Then

$$U' = [-1+x_1+\beta(x_2+x_3),\ -1+x_2+\beta(x_1+x_3),\ -1+x_3+\beta(x_1+x_2)].$$

Thus, the stationary points are obtained by solving $U' = [0\ 0\ 0]$, i.e.:

$$\begin{bmatrix} 1 & \beta & \beta \\ \beta & 1 & \beta \\ \beta & \beta & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

This yields $x_i = \frac{1}{2\beta+1}$, for all i.
Using the definition of x_i, we obtain the system for the original variables:

$$\begin{bmatrix} 1 & \beta & \beta \\ \beta & 1 & \beta \\ \beta & \beta & 1 \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} = (2\beta+1) \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}.$$

Solving this system gives $\begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} = \frac{1}{\beta-1} \begin{bmatrix} ((\alpha_2-\alpha_1+\alpha_3)*\beta-\alpha_1) \\ ((\alpha_1-\alpha_2+\alpha_3)*\beta-\alpha_2) \\ ((\alpha_1+\alpha_2-\alpha_3)*\beta-\alpha_3) \end{bmatrix}$.

c. Following item b, we obtain $g_j = \frac{1}{\beta-1}(\beta\sum_{i\neq j}\alpha_i - (1+\beta)\alpha_j)$.

EXERCISE 2.61 *(optimal carbondioxide emissions)*
Consider N countries who want to optimize CO_2 emissions. Welfare of country i equals: $W_i = R_i(e_i) - D_i(\sum_{j=1}^{N} e_j)$, where R_i are welfare revenues associated with the production of emissions e_i and D_i is the associated pollution damage. All functions are assumed to be twice continuous differentiable, $R_i(x)$ concave, and $D_i(x)$ convex. Assume countries agree to maximize total welfare in the countries.

a. Consider $N = 3$. Show that, at the optimum (e_1^*, e_2^*, e_3^*), necessarily $R_i'(e_i^*) = R_j'(e_j^*),\ i,j = 1,2,3$.

b. Show that the total welfare function is a concave function.

c. Assume the welfare function has a stationary point (e_1^*, e_2^*, e_3^*). Show that there exist a unique set of emissions, given by (e_1^*, e_2^*, e_3^*), where total welfare is maximized.

d. Consider $N = 3$, with $R_i(x) = 2\alpha_i\sqrt{x}$ and $D_i(x) = \beta_i x$, where $\alpha_i > 0$ and $\beta_i > 0$. Determine the optimal emissions. Show that this optimum indeed yields a maximum.

ANSWER 2.61

a. The problem is to minimize

$$W := \sum_{i=1}^{3} W_i = \sum_{i=1}^{3} R_i(e_i) - D_i(\sum_{j=1}^{3} e_j).$$

The derivative is

$$DW(e_1, e_2, e_3) = [R_1^{'} - \sum_{i=1}^{3} D_i^{'} \;\; R_2^{'} - \sum_{i=1}^{3} D_i^{'} \;\; R_3^{'} - \sum_{i=1}^{3} D_i^{'}].$$

Hence, the stationary points satisfy

$$R_j^{'} - \sum_{i=1}^{3} D_i^{'} = 0, \; j = 1, 2, 3,$$

which implies

$$R_1^{'}(e_1^*) = R_2^{'}(e_2^*) = R_3^{'}(e_3^*) = \sum_{i=1}^{3} D_i^{'}(e_1^* + e_2^* + e_3^*).$$

b. The Hessian of W is

$$\begin{aligned} HW &= \begin{bmatrix} R_1^{"} - \sum_{i=1}^{3} D_i^{"} & -\sum_{i=1}^{3} D_i^{'} & -\sum_{i=1}^{3} D_i^{'} \\ -\sum_{i=1}^{3} D_i^{'} & R_2^{"} - \sum_{i=1}^{3} D_i^{"} & -\sum_{i=1}^{3} D_i^{'} \\ -\sum_{i=1}^{3} D_i^{'} & -\sum_{i=1}^{3} D_i^{'} & R_3^{"} - \sum_{i=1}^{3} D_i^{"} \end{bmatrix} \\ &= \begin{bmatrix} R_1^{"} & 0 & 0 \\ 0 & R_2^{"} & 0 \\ 0 & 0 & R_3^{"} \end{bmatrix} - \sum_{i=1}^{3} D_i^{"} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} R_1^{"} & 0 & 0 \\ 0 & R_2^{"} & 0 \\ 0 & 0 & R_3^{"} \end{bmatrix} - \sum_{i=1}^{3} D_i^{"} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}. \end{aligned}$$

By assumption, $D_i(x)$ is convex, which implies $D_i^{"}(x) \geq 0$ for all x, and therefore $\sum_{i=1}^{3} D_i^{"}(x) \geq 0$ for all x.

Furthermore, for any $z \in \mathbb{R}^3$, $z^T \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} z = y^T y \geq 0$. Hence,

$$\sum_{i=1}^{3} D_i^{"}(x) z^T \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} z \geq 0 \text{ for arbitrary } z \in \mathbb{R}^3,$$

which implies that $\sum_{i=1}^{3} D_i^{"}(e_1+e_2+e_3)\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ is positive semi-definite.
On the other hand, R_i is concave, so $R_i^{"}(e_i) \leq 0$ for all e_i. Therefore, for any $z \in \mathbb{R}^3$ and any (e_1, e_2, e_3),

$$z^T \begin{bmatrix} R_1^{"} & 0 & 0 \\ 0 & R_2^{"} & 0 \\ 0 & 0 & R_3^{"} \end{bmatrix} z = z_1^2 R_1^{"}(e_1) + z_2^2 R_2^{"}(e_2) + z_3^2 R_3^{"}(e_3) \leq 0,$$

which implies that $\begin{bmatrix} R_1^{"} & 0 & 0 \\ 0 & R_2^{"} & 0 \\ 0 & 0 & R_3^{"} \end{bmatrix}$ is negative semi-definite.
Combining these results, HW is negative semi-definite, which implies W is concave.

c. Since the welfare function is concave, by Exercise 2.49, the function attains its global maximum at the stationary point.

d. With the advertized functions,

$$W(e_1, e_2, e_3) = \sum_{i=1}^{3} \{2\alpha_i \sqrt{e_i} - \beta_i (\sum_{j=1}^{3} e_j)\}.$$

The first-order condition is

$$W_{e_i}^{'} = \frac{\alpha_i}{\sqrt{e_i}} - \sum_{j=1}^{3} \beta_j = 0, \text{ which implies } e_i^* = \frac{\alpha_i^2}{(\sum_{j=1}^{3} \beta_j)^2}.$$

The Hessian is $HW = \begin{bmatrix} -\frac{\alpha_1}{2e_1^{\frac{3}{2}}} & 0 & 0 \\ 0 & -\frac{\alpha_2}{2e_2^{\frac{3}{2}}} & 0 \\ 0 & 0 & -\frac{\alpha_3}{2e_3^{\frac{3}{2}}} \end{bmatrix}$. Clearly, $HW(e_1^*, e_2^*, e_3^*)$ is negative definite.
Hence W attains a maximum at (e_1^*, e_2^*, e_3^*).

EXERCISE 2.62 A producer's production function of shavers is given by $Y(K, L) = 2K^{\frac{1}{2}} L^{\frac{1}{4}}$. The cost of one unit of capital, K, is 8 and of one unit of labour, L, is 2. The market price of shavers is 32.

a. Show that the profit function of the producer is $\pi(K, L) = 64K^{\frac{1}{2}} L^{\frac{1}{4}} - 8K - 2L$.

b. Show that $\pi(K, L)$ is a concave function on the domain $K > 0$ and $L > 0$.

c. Show that $\pi(K, L)$ has a stationary point at $(K, L) = (512, 1024)$.

d. Determine the maximum profits the producer can make.

ANSWER 2.62

a. $\pi(K, L) = 32 * Y - 8K - 2L = 64K^{\frac{1}{2}} L^{\frac{1}{4}} - 8K - 2L$

b. Straightforward calculations show

$$D\pi(K,L) = [32K^{-\frac{1}{2}}L^{\frac{1}{4}} - 8 \;\; 16K^{\frac{1}{2}}L^{-\frac{3}{4}} - 2], \text{ and } \pi^{"}(K,L) = \begin{bmatrix} -16K^{-\frac{3}{2}}L^{\frac{1}{4}} & 8K^{-\frac{1}{2}}L^{-\frac{3}{4}} \\ 8K^{-\frac{1}{2}}L^{-\frac{3}{4}} & -12K^{\frac{1}{2}}L^{-\frac{7}{4}} \end{bmatrix}.$$

Hence, $-\pi^{"}(K,L) = \begin{bmatrix} 16K^{-\frac{3}{2}}L^{\frac{1}{4}} & -8K^{-\frac{1}{2}}L^{-\frac{3}{4}} \\ -8K^{-\frac{1}{2}}L^{-\frac{3}{4}} & 12K^{\frac{1}{2}}L^{-\frac{7}{4}} \end{bmatrix}$.

Note that

$$16K^{-\frac{3}{2}}L^{\frac{1}{4}} > 0, \; 12K^{\frac{1}{2}}L^{-\frac{7}{4}} > 0, \text{ and } \det -\pi^{"} = 128K^{-1}L^{-\frac{3}{2}} > 0.$$

Therefore, $-\pi^{"}(K,L)$ is positive definite, which implies that $\pi^{"}(K,L)$ is negative definite, and hence $\pi(K,L)$ is concave.

c. $D\pi(512, 1024) = [0 \;\; 0]$.

d. Since π is a concave function, its maximum is attained at the stationary point. Evaluating at $(K,L) = (512, 1024)$ gives $\pi(512, 1024) = 2048$.

EXERCISE 2.63 A baker can produce both bread, b, and cookies, c. His cost function is $C(b,c) = \frac{2}{3}b^2 + \frac{1}{3}bc + \frac{1}{6}c^2$. The market price of bread is 3 and of cookies is 1.

a. Determine the profit function of the producer.

b. Determine the amount of bread and cookies that maximize his profit.

c. Determine the maximum profits the producer can make.

ANSWER 2.63

a. $\pi(b,c) = 3 * b + c - C(b,c) = 3 * b + c - \frac{2}{3}b^2 - \frac{1}{3}bc - \frac{1}{6}c^2$.

b. The derivative of π is

$$D\pi(b,c) = [3 - \frac{4}{3}b - \frac{1}{3}c \;\; 1 - \frac{1}{3}b - \frac{1}{3}c],$$

which equals $[0 \;\; 0]$ at $(b,c) = (2,1)$.

The Hessian is

$$\pi^{"} = \begin{bmatrix} -\frac{4}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}, \text{ hence, } -\pi^{"}(b,c) = \begin{bmatrix} \frac{4}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}.$$

Note that $\frac{4}{3} > 0$, $\frac{1}{3} > 0$ and $\det -\pi^{"} = \frac{1}{3} > 0$. Therefore, $-\pi^{"}(b,c)$ is positive definite, which implies $\pi^{"}(b,c)$ is negative definite. Hence $\pi(b,c)$ is concave, and the maximum profit is attained at $(b,c) = (2,1)$.

c. $\pi(2,1) = 3\frac{5}{6}$.

EXERCISE 2.64 A travel agency has a fixed budget of 14 (in millions of euro) to advertize a cruise. It has three gremia for this purpose: newspaper, television and internet. Let x_1 denote the amount of money spent on advertizing in the newspaper, x_2 the amount of money spent on advertizing on the television and x_3 the amount of money spent on advertizing on the internet.

The number of customers $N_j,\ j = n, t, i$, when (x_1, x_2, x_3) is spent on advertizing is:

$$N_n(x_1, x_2, x_3) = 2x_1 + x_2 + x_3;\ N_t(x_1, x_2, x_3) = x_1 + 3x_2 + x_3;\ N_i(x_1, x_2, x_3) = \frac{1}{2}x_1 + \frac{1}{2}x_2 + x_3.$$

Assume the price of the cruise is p.

a. Show that the number of customers is determined by maximizing
$f(x_1, x_2) = 2x_1^2 + 3x_2^2 + x_3^2 + 2x_1x_2 + \frac{3}{2}(x_1 + x_2)x_3$, where $x_3 = 14 - x_1 - x_2$.

b. Determine how much of the price p consists of advertising cost.

ANSWER 2.64

a. $f(x_1, x_2, x_3) = x_1N_n + x_2N_t + x_3N_i$.

b. Profit is given by

$$\pi(x_1, x_2) = pf(x_1, x_2, x_3) - 100, \text{ with } x_3 = 14 - x_1 - x_2.$$

Its derivative is

$$D\pi(x_1, x_2) = [4x_1 - 2x_3 + 2x_2 + \frac{3}{2}(x_3 - x_1 - x_2)\ \ 6x_2 - 2x_3 + 2x_1 + \frac{3}{2}(x_3 - x_2 - x_1)].$$

Setting $D\pi(x_1, x_2) = [0\ \ 0]$ gives the system of equations $x_1 + 5x_2 - 7 = 0$ and $3x_1 + x_2 - 7 = 0$, which has the solution $(x_1, x_2) = (2, 1)$. This implies optimal advertising cost is $x_3 = 11$.

EXERCISE 2.65 A farmer has 100 (acres) of land. He has to decide how much land he wants to use to grow corn, potatoes and wheat, respectively. On x (acres) of land he can produce $3x$, $6x$ and $2x$ kilos (in hundreds) of corn, potatoes and wheat, respectively. The cost involved of growing x (acres) of corn, potatoes and wheat are $\ln(1 + 3x)$, $2\ln(1 + 6x)$ and $2\ln(1 + 2x)$, respectively. The price of selling 100 kilos of corn, potatoes and wheat is 2, 1 and 2 euro s (in hundreds), respectively.

a. Show that the profit function of the farmer is given by $\Pi(x_1, x_2, x_3) = 6x_1 + 6x_2 + 4x_3 - \ln(1 + 3x_1) - 2\ln(1 + 6x_2) - 2\ln(1 + 2x_3)$ where $x_i,\ i = 1, 2, 3$, is the land used for growing corn, potatoes and wheat, respectively, and $x_3 = 100 - x_1 - x_2$.

b. Determine how much acres the farmer has to use for growing corn, potatoes and wheat, respectively, in order to maximize his profits.

ANSWER 2.65

a. $\Pi(x_1, x_2, x_3) = 2 * 3x_1 - \ln(1 + 3x_1) + 1 * 6x_2 - 2\ln(1 + 6x_2) + 2 * 2x_3 - 2\ln(1 + 2x_3)$.

b. With $x_3 = 100 - x\_x_2$ the derivative of the profit function is

$$D\Pi(x_1, x_2) = [2 - \frac{3}{1 + 3x_1} + \frac{4}{1 + 2x_3}\ \ 2 - \frac{12}{1 + 6x_2} + \frac{4}{1 + 2x_3}].$$

The stationary points satisfy $D\Pi(x_1, x_2) = [0 \;\; 0]$, which leads to

$$\frac{3}{1+3x_1} = \frac{12}{1+6x_2}, \text{ and } 2 - \frac{3}{1+3x_1} + \frac{4}{1+2x_3} = 0.$$

From the first equation we get $x_2 = \frac{1}{2} + 2x_1$. Substituting $x_3 = 100 - x_1 - x_2 = 99 - \frac{3}{2}x_1$ into the second equation gives

$$2 - \frac{3}{1+3x_1} + \frac{4}{200-6x_1} = 0, \text{ or, } 2(1+3x_1)(200-6x_1) - 3(200-6x_1) + 4(1+3x_1) = 0.$$

Solving this yields two solutions:

$$x_1 = 0.16 \text{ (with } x_2 = 0.82, x_3 = 99) \text{ or } x_1 = 33.67 \text{ (with } x_2 = 67.84, x_3 = -1.51).$$

The second solution is not feasible because $x_3 < 0$.

One can easily verify that the Hessian $H\Pi$ is positive definite, which implies that profits are *minimized* at the feasible stationary point.

Since the domain is compact and Π is continuous, Weierstrass' theorem guarantees that a maximum exists, which must therefore be attained somewhere on the boundary of the domain.

So, let's consider $\Pi(x_1, x_2)$ on the boundaries of its domain.

Case 1: $x_2 = 0, \; x_1 \in [0, 100]$: $\Pi(x_1) = 6x_1 + 4x_3 - \ln(1 + 3x_1) - 2\ln(1 + 2x_3) \rightarrow \Pi'(x_1) = 2 - \frac{3}{1+3x_1} - \frac{4}{1+2x_3} \rightarrow \Pi"(x_1) = \frac{9}{(1+3x_1)^2} + \frac{8}{(1+2x_3)^2} \rightarrow$ So, if Π has a stationary point it will be a minimum $\rightarrow$ maximum attained at either $x_1 = 0$ or $x_1 = 100$.

Case 2: $x_1 = 0, \; x_2 \in [0, 100]$: In a similar way it follows that also $\Pi(x_2)$ attains its maximum at either $x_2 = 0$ or $x_2 = 100$.

Evaluation of $\Pi(100, 0, 0)$, $\Pi(0, 100, 0)$ and $\Pi(0, 0, 100)$ shows that the maximum value is attained at $(x_1, x_2, x_3) = (100, 0, 0)$.

EXERCISE 2.66 An investor wants to decide how much of his capital he wants to put on the bank, earning a risk-free interest rate, $r = 0.03$; investing it in stock 1, getting an expected return of $\mu_1 = .04$, at a risk $\sigma_1 = .0316$; or investing it in stock 2, getting an expected return of $\mu_2 = .08$, at a risk $\sigma_2 = .05$. Let the fractions he invests in stock 1 and 2 and he puts on the bank be $\omega_i, \; i = 1, 2, 3$, respectively. Assume the expected return of such a portfolio is $\mu = .04\omega_1 + .08\omega_2 + .03\omega_3$ and the risk $\sqrt{0.001\omega_1^2 + 0.0025\omega_2^2 - 0.002\omega_1\omega_2}$. Assume the investor likes to maximize his utility function $U := \mu - 100\sigma^2$.

a. Show that $U(\omega_1, \omega_2, \omega_3) = -0.1\omega_1^2 - 0.25\omega_2^2 + 0.2\omega_1\omega_2 + 0.04\omega_1 + 0.08\omega_2 + 0.03\omega_3$. where $\omega_3 = 1 - \omega_1 - \omega_2$.

b. Determine the fractions he should invest in the stocks in order to maximize his utility $U(\omega_1, \omega_2)$. What is the fraction he should invest on the bank?

ANSWER 2.66

a. Calculate $U := \mu - 100\sigma^2$, with $\mu = .04\omega_1 + .08\omega_2 + .03\omega_3$ and $\sigma = \sqrt{0.001\omega_1^2 + 0.0025\omega_2^2 - 0.002\omega_1\omega_2}$.

b. The deriative of U is

$$DU(\omega_1, \omega_2) = [-0.2\omega_1 + 0.2\omega_2 + 0.01 \;\; 0.2\omega_1 - 0.5\omega_2 + 0.05].$$

The stationary points are found by setting $DU(\omega_1, \omega_2) = [0 \;\; 0]$, which gives $(\omega_1, \omega_2) = (0.25, 0.2)$ and consequently, $\omega_3 = 0.55$.

Since the Hessian HU is negative definite, it follows that $(\omega_1, \omega_2, \omega_3) = (0.25, 0.2, 0.55)$ corresponds to a maximum of U.

EXERCISE 2.67 *(optimal pollution)*

Let $U(y, p)$ measure the total welfare of society, where y is the number of produced goods and p the associated level of pollution. Assume that welfare increases if y increases, and decreases if p increases.

a. Show there exists not an optimal level of welfare under the above model assumptions.

b. Substantiate why the above model might not be a correct model of reality.

c. Consider the variable $z = \frac{p}{y}$. Motivate why for large values of z utility, $U(z)$, might be negative.

d. Consider $U(z) = -\frac{1}{2}z^2 + z$. Determine the optimal ratio between pollution and goods production that maximizes welfare.

ANSWER 2.67

a. By increasing y and decreasing p welfare increases forever.

b. Likely, y decreases if p increases. This relationship lacks.

c. z large implies that p is much larger than y. Very likely this is disliked by society.

d. The derivative of U is $U^{'}(z) = -z + 1$. Setting $U^{'}(z) = 0$ gives $z = 1$. Since $U^{"} < 0$, $z = 1$ corresponds to a global maximum. Therefore, U is maximized when $p = y$.

EXERCISE 2.68 *(homogeneous functions)*

The production function $Y(K, L)$ is called homogeneous of degree k if a multiplication of both inputs K and L by a factor t leads to a multiplication of the output Y of t^k. That is, if $Y(tK, tL) = t^k Y(K, L)$.

a. Consider the Cobb-Douglas production function $Y(K, L) = cK^{\alpha}L^{\beta}$, where c, α and β are positive constants. Show that $Y(K, L)$ is homogeneous of degree 1 if and only if $\alpha + \beta = 1$.

b. Explain in item a. what the effect on production is of doubling the amount of K and L if the degree of homogeneity is 1, smaller than 1 and larger than 1, respectively.

c. Consider $f(t) = Y(tK, tL)$. Show that $f^{'}(t) = D_1Y(tK, tL)K + D_2Y(tK, tL)L$.

d. Consider $f(t) = t^kY(K, L)$. Show that $f^{'}(t) = kt^{k-1}Y(K, L)$.

e. Assume $Y(K, L)$ is homogeneous of degree k. Show that $DY(K, L)\begin{bmatrix} K \\ L \end{bmatrix} = kY(K, L)$. Hint: use item c. and d. with $t = 1$.

f. Use item e. to show that if a homogeneous function has degree k:
$D_1Y(K,L)\frac{K}{Y(K,L)} + D_2Y(K,L)\frac{L}{Y(K,L)} = k$. (In economic terminology: the sum of the partial elasticities equals the degree of homogeneity).

g. Consider the Cobb-Douglas production function from item a., with $\alpha = \frac{1}{4}$ and $\beta = \frac{3}{4}$. Verify the correctness of item f.

ANSWER 2.68

a. Note that

$$Y(tK,tL) = c(tK)^{\alpha}(tL)^{\beta} = t^{\alpha+\beta}cK^{\alpha}L^{\beta} = t^{\alpha+\beta}Y(K,L).$$

Hence, Y is homogeneous of degree 1 if and only if $\alpha + \beta = 1$.

b. $k = 1$: output exactly doubles; $k < 1$: output less than doubles; $k > 1$: output more than doubles.

c. $f(t) = Y(g(t))$, where $g(t) = [tK\ tL]^T$. By the chain rule of differentiation, we have:

$$f^{'}(t) = DY(tK,tL)Dg = [D_1Y\ D_2Y][K\ L]^T = D_1Y(tK,tL)K + D_2Y(tK,tL)L.$$

d+e.If Y is homogeneous of degree k, $Y(tK,tL) = t^kY(K,L)$. Therefore, $f(t) = t^kY(K,L)$, and consequently $f^{'}(t) = kt^{k-1}Y(K,L)$. Using the result from item c. it follows that

$$f^{'}(t) = kt^{k-1}Y(K,L) = D_1Y(tK,tL)K + D_2Y(tK,tL)L.$$

This holds for any t. In particular, for $t = 1$ we obtain:

$$kY(K,L) = D_1Y(K,L)K + D_2Y(K,L)L.$$

f. By item e, $kY(K,L) = D_1Y(K,L)K + D_2Y(K,L)L$. Dividing both sides by Y gives the stated result.

g. The partial derivatives are

$$D_1Y = c\frac{1}{4}K^{-\frac{3}{4}}L^{\frac{3}{4}}, \text{ and } D_2Y = c\frac{3}{4}K^{\frac{1}{4}}L^{-\frac{1}{4}}.$$

It follows that

$$\begin{aligned} D_1Y(K,L)K + D_2Y(K,L)L &= c\frac{1}{4}K^{-\frac{3}{4}}L^{\frac{3}{4}}K + c\frac{3}{4}K^{\frac{1}{4}}L^{-\frac{1}{4}}L \\ &= c\frac{1}{4}K^{\frac{1}{4}}L^{\frac{3}{4}} + c\frac{3}{4}K^{\frac{1}{4}}L^{\frac{3}{4}} \\ &= \frac{1}{4}Y(K,L) + \frac{3}{4}Y(K,L) = Y(K,L). \end{aligned}$$

Dividing both sides by $Y(K,L)$ gives $D_1Y(K,L)\frac{K}{Y(K,L)} + D_2Y(K,L)\frac{L}{Y(K,L)} = 1$.

EXERCISE 2.69 Consider a producer of a product who can sell this product on two separate markets. The price he gets on both markets depends on the amount he supplies to that market. For market 1 this relationship is given by $p(x) = 20 - \frac{1}{2}x$, whereas for market 2 it is given by $p(x) = 25 - 2x$. Assume the cost of producing x units of the product is $C(x) = 5x + 60$.

a. Determine the profit function $\Pi(x_1, x_2)$ for this producer, where x_i denotes the amount of products sold on market $i,\ i = 1, 2$.

b. Calculate his maximal profit and the amounts of the product he will sell on both markets.

c. Assume the producer is not allowed to ask different prices on both markets. Calculate his profit under this additional constraint. What are his losses compared to the case that price discrimination is allowed.

ANSWER 2.69

a. Let p_i and x_i denote the price and quantity sold in market i. Then the profit function is

$$\Pi(x_1, x_2) = p_1x_1 + p_2x_2 - C(x_1 + x_2) = (20 - \frac{1}{2}x_1)x_1 + (25 - 2x_2)x_2 - 5(x_1 + x_2) - 60.$$

b. The derivative of Π is $D\Pi(x_1, x_2) = [15 - x_1 \ \ 20 - 4x_2]$, which gives the stationary point $(x_1, x_2) = (15, 5)$.
Note that the Hessian $H\Pi(x_1, x_2)$ is negative definite, so $(x_1, x_2) = (15, 5)$ corresponds to a global maximum with $\Pi(15, 5) = 102.5$.

c. Prices in both markets must be equal, so $p_1 = p_2$, which gives

$$20 - \frac{1}{2}x_1 = 25 - 2x_2, \text{ which gives } x_2 = \frac{1}{4}x_1 + 2.5.$$

Substituting this into the profit function, we get

$$\begin{aligned} \Pi(x_1, x_2) &= \Pi(x_1) \\ &= (20 - \frac{1}{2}x_1)x_1 + (25 - 2x_2)x_2 - 5(x_1 + x_2) - 60 \\ &= (20 - \frac{1}{2}x_1)x_1 + (20 - \frac{1}{2}x_1)(\frac{1}{4}x_1 + 2.5) - 5(x_1 + \frac{1}{4}x_1 + 2.5) - 60. \end{aligned}$$

Differentiating with respect to x_1 gives $\Pi'(x_1) = -\frac{10}{8}x_1 + \frac{70}{4}$. Setting $\Pi'(x_1) = 0$ yields $x_1 = 14$. One can easily verify that this gives a global maximum, with $\Pi(14) = 100$, which corresponds to a profit loss of 2.5 compared to the unconstrained optimum.

EXERCISE 2.70 Consider the Cob-Douglas production function $Y(K, L) = K^{\frac{2}{3}}L^{\frac{2}{3}}$, where K and L denote the amount of capital and labour used, respectively. Assume the price of capital is $\frac{1}{3}$ and of labour is $\frac{1}{3}$ too. Assume the price the producer gets for his product depends as follows on his production $p = 3Y^{-\frac{1}{2}}$.

a. Determine the profit function, $\Pi(K, L)$ for this producer.

b. Calculate his maximal profit and the corresponding production.

ANSWER 2.70

a. $\Pi(K,L) = pY - \frac{1}{3}K - \frac{1}{3}L = 3Y^{-\frac{1}{2}}Y - \frac{1}{3}K - \frac{1}{3}L = 3Y^{\frac{1}{2}} - \frac{1}{3}K - \frac{1}{3}L = 3K^{\frac{1}{3}}L^{\frac{1}{3}} - \frac{1}{3}K - \frac{1}{3}L$.

b. The derivative of Π is

$$D\Pi(K,L) = [K^{-\frac{2}{3}}L^{\frac{1}{3}} - \frac{1}{3} \quad K^{\frac{1}{3}}L^{-\frac{1}{3}} - \frac{1}{3}].$$

Setting $D\Pi(K,L) = [0 \quad 0]$ requires $K = L$ and $K^{-\frac{2}{3}}L^{\frac{1}{3}} = \frac{1}{3}$. That is, $K = L$ and $K^{\frac{1}{3}} = 3$. This gives $K = L = 27$. So maximum profit and corresponding production is $\Pi(27,27) = 25$ and $Y = 81$.

EXERCISE 2.71 *(profit maximization by a firm)*

Assume the production function of a firm is $Y(K,L)$, where K and L denote the amount of capital and labour used, respectively. Assume the price of capital is r and of labour is w. Assume the price the firm gets for his product is p.

a. Assume p is fixed. Show that maximum profits are obtained only where next two equalities hold $Y_K^{'} = \frac{r}{p}$ and $Y_L^{'} = \frac{w}{p}$.

b. Assume p depends on the production Y, i.e. $p = P(Y)$. Show that maximum profits are obtained only where next two equalities hold $Y_K^{'}(P^{'}Y + P) = r$ and $Y_L^{'}(P^{'}Y + P) = w$. Show that if profits are maximized then: $\frac{r}{Y_K^{'}} = \frac{w}{Y_L^{'}}$. Verify these conditions are satisfied in Exercise 2.70.

c. Assume p depends directly on the inputs K and L used for production, i.e. $p = P(K,L)$. Show that maximum profits are obtained only where next two equalities hold $\frac{P_K^{'}}{P} + \frac{Y_K^{'}}{Y} = \frac{r}{PY}$ and $\frac{P_L^{'}}{P} + \frac{Y_L^{'}}{Y} = \frac{w}{PY}$. Verify these conditions are satisfied in Exercise 2.70, using $P(K,L) = 3K^{-\frac{1}{3}}L^{-\frac{1}{3}}$.

ANSWER 2.71

a. Profits are given by $\Pi(K,L) = pY - rK - wL$, so the derivative is

$$D\Pi(K,L) = [pY_K^{'} - r \quad pY_L^{'} - w].$$

Setting $D\Pi(K,L) = [0 \quad 0]$ requires $Y_K^{'} = \frac{r}{p}$ and $Y_L^{'} = \frac{w}{p}$.

b. Profits are given by $\Pi(K,L) = P(Y)Y - rK - wL$, so the derivative is

$$D\Pi(K,L) = [(P^{'}(Y)Y + P(Y))Y_K^{'} - r \quad (P^{'}(Y)Y + P(Y))Y_L^{'} - w].$$

Setting $D\Pi(K,L) = [0 \quad 0]$ requires $Y_K^{'}(P^{'}Y + P) = r$ and $Y_L^{'}(P^{'}Y + P) = w$. Therefore, at the optimum,

$$P^{'}Y + P = \frac{r}{Y_K^{'}} = \frac{w}{Y_L^{'}}$$

For Exercise 2.70: with $r = w = \frac{1}{3}$, $Y_K^{'} = \frac{2}{3}K^{-\frac{1}{3}}L^{\frac{2}{3}} = \frac{2}{3}\frac{1}{3}9 = 2$ and, similarly, $Y_L^{'} = 2$, we have $\frac{r}{Y_K^{'}} = \frac{w}{Y_L^{'}}$.

c. Profits are given by $\Pi(K,L) = P(K,L)Y - rK - wL$, so the derivative

$$D\Pi(K,L) = [P_K^{'}Y + PY_K^{'} - r \;\; P_L^{'}Y + PY_L^{'} - w].$$

Setting $D\Pi(K,L) = [0 \;\; 0]$ requires $P_K^{'}Y + PY_K^{'} = r$ and $P_L^{'}Y + PY_L^{'} = w$, or equivalently,

$$\frac{P_K^{'}}{P} + \frac{Y_K^{'}}{Y} = \frac{r}{PY}, \text{ and } \frac{P_L^{'}}{P} + \frac{Y_L^{'}}{Y} = \frac{w}{PY}.$$

For Exercise 2.70: with $r = w = \frac{1}{3}$, $K = L = 27$, $Y_K^{'} = Y_L^{'} = 2$, $P = 3 * \frac{1}{3} * \frac{1}{3} = \frac{1}{3}$ and $Y = 81$.
Furthermore, $P_K^{'} = -K^{-\frac{4}{3}}L^{-\frac{1}{3}} = -\frac{1}{81} * \frac{1}{3}$ and $P_L^{'} = -\frac{1}{81} * \frac{1}{3}$.
Hence, $\frac{P_K^{'}}{P} + \frac{Y_K^{'}}{Y} = \frac{r}{PY}$ and $\frac{P_L^{'}}{P} + \frac{Y_L^{'}}{Y} = \frac{w}{PY}$.

EXERCISE 2.72 *(Price Discrimination)*
Consider a producer of a product who can sell this product on two separate markets. The price he gets on both markets depends on the amount, $x_i, \; i = 1,2$, he supplies to that market. Assume the cost of producing x units of the product is $C(x)$.

a. Show that the profit function $\Pi(x_1,x_2)$ for this producer is $\Pi(x_1,x_2) = P_1(x_1)x_1 + P_2(x_2)x_2 - C(x_1+x_2)$, where $P_i(x_i)$ is the inverse demand function for the product in market $i, \; i = 1,2$.
b. Show that maximum profits are obtained only where next two equalities hold $P_i^{'}(x_i)x_i + P_i(x_i) = C^{'}(x_1+x_2), \; i = 1,2$. Interpret this equality in terms of profits obtained in both markets.
c. Verify the conditions obtained in item b for Exercise 2.69.

ANSWER 2.72

a. Clear from the context that $\Pi(x_1,x_2) = P_1(x_1)x_1 + P_2(x_2)x_2 - C(x_1+x_2)$.
b. The derivative of Π is

$$D\Pi(x_1,x_2) = [P_1^{'}(x_1)x_1 + P_1(x_1) - C^{'}(x_1+x_2) \;\; P_2^{'}(x_2)x_2 + P_2(x_2) - C^{'}(x_1+x_2)].$$

Setting $D\Pi(x_1,x_2) = [0 \;\; 0]$ requires

$$P_i^{'}(x_i)x_i + P_i(x_i) = C^{'}(x_1+x_2), \; i = 1,2.$$

Thus, the marginal revenue in each market must equal the marginal cost of total output.
c. For Exercise 2.69: $(x_1,x_2) = (15,5)$, $p_1 = 12.5$, $p_2 = 15$, $P_1^{'} = -\frac{1}{2}$, $P_2^{'} = -2$ and $C^{'} = 5$. It follows that $P_i^{'}(x_i)x_i + P_i(x_i) = C^{'}(x_1+x_2), \; i = 1,2$.

CHAPTER 3

Constrained Optimization: Equality Constraints

3.1 The Substitution Method

EXERCISE 3.1 Determine all optima (location, value, kind (local/global)) of the next constrained optimization problems.

a. Optimize $f(x,y) = xy^2$, subject to the constraint $g(x,y) := x + y = 1$.

b. Optimize $f(x,y) = e^x y^2$, subject to the constraint $g(x,y) := x + y^2 = 1$.

c. Optimize $f(x,y,z) = x^2 + y^2 + z^2$, subject to the constraints $g_1(x,y,z) := x + y + z = 0$ and $g_2(x,y,z) := x - y + z = 2$.

d. Optimize $f(x,y,z) = \frac{1}{2}x^2 y + z$ subject to the constraints $g_1(x,y,z) := \frac{1}{2}x^2 + 1 - y = 0$ and $g_2(x,y,z) := \frac{4}{3}x - \frac{4}{3}xy - z = 0$.

e. Optimize $f(x,y,z) = x^2 + y^2 + z^2$, subject to the constraint $g(x,y,z) := x - y + 2z = 6$.

f. Optimize $f(x,y,z) = e^x + e^{3-y} + e^z$, subject to the constraint $g(x,y) := x - y + z = 0$.

ANSWER 3.1

a. $g(x,y) \to y = 1 - x \to F(x) = x(1-x)^2$

$F'(x) = (1-x)^2 - 2x(1-x) = (1-x)(1-x-2x) \to$ stationary points: $x = 1$ and $x = \frac{1}{3}$.

```
   +         -         +     F'(x)
-------0--------0-------
       1/3       1           x
```

$\to F$ has a local maximum at $x = \frac{1}{3}$ and a local minimum at $x = 1$. Note: $F(100) > f(\frac{1}{3})$ and $F(-100) < f(\frac{1}{3}) \to F$ has no global extremal points.

$\to$ Problem has local maximum at $(\frac{1}{3}, \frac{2}{3})$ and a local minimum at $(1,0)$.

b. $g(x,y) \to x = 1 - y^2 \to F(y) = y^2 e^{1-y^2}$

$F'(y) = (2y - 2y^3)e^{1-y^2} \to$ stationary points: $y = -1$, $y = 0$ and $y = 1$.

```
  +        -        +        -     F'(y)
-------0-------0-------0------
      -1        0        1         y
```

$\rightarrow F$ has a local maximum at $y = -1$ and $y = 1$, and a local minimum at $y = 0$. Note: $F(y) \geq 0$ and goes to zero if either y goes to infinity or minus infinity. As $F(0) = 0$, it follows that F has a global minimum at $y = 0$. As $F(1) = F(-1)$ it follows from the sign scheme of the derivative that F has a global maximum at both $y = -1$ and $y = 1$.

$\rightarrow$ Problem has global maximum at $(0, -1)$ and $(0, 1)$; and a global minimum at $(1, 0)$.

c. From g_1 and g_2 we have: $y = -1$ and $z = 1 - x \rightarrow F(x) = x^2 + 1 + (1 - x)^2$
$F^{'}(x) = 2x - 2(1 - x) \rightarrow$ stationary point: $x = \frac{1}{2}$.

```
  -        +    F'(x)
-------0-----
      1/2       x
```

$\rightarrow F$ has a global minimum at $x = \frac{1}{2}$.

$\rightarrow$ Problem has global minimum at $(\frac{1}{2}, -1, \frac{1}{2})$.

d. From $g_1 : y = 1 + \frac{1}{2}x^2$; from $g_2 : z = \frac{4}{3}x - \frac{4}{3}x(1 + \frac{1}{2}x^2)$
$\rightarrow F(x) = \frac{1}{2}x^2(1 + \frac{1}{2}x^2) + \frac{4}{3}x - \frac{4}{3}x(1 + \frac{1}{2}x^2) = \frac{1}{4}x^4 - \frac{2}{3}x^3 + \frac{1}{2}x^2$
$F^{'}(x) = x^3 - 2x^2 + x = x(x^2 - 2x + 1) \rightarrow$ stationary points: $x = 0$ and $x = 1$.

```
  -        +        +    F'(x)
-------0-------0-------
       0        1          x
```

$\rightarrow F$ has a global minimum at $x = 0$.

$\rightarrow$ Problem has global minimum at $(0, 1, 0)$.

e. From $g : y = x + 2z - 6. \rightarrow F(x, z) = x^2 + (x + 2z - 6)^2 + z^2$.
$F_x^{'} = 2x + 2(x + 2z - 6)$ and $F_z^{'} = 2z + 4(x + 2z - 6) \rightarrow$ stationary point: $(x, z) = (1, 2)$.
$HF = \begin{bmatrix} 4 & 4 \\ 4 & 10 \end{bmatrix} \rightarrow F$ is convex has a global minimum at $(x, z) = (1, 2)$.

$\rightarrow$ Problem has global minimum at $(1, -1, 2)$.

f. From $g : y = x + z. \rightarrow F(x, z) = e^x + e^{3-x-z} + e^z$.
$F_x^{'} = e^x - e^{3-x-z}$ and $F_z^{'} = e^z - e^{3-x-z} \rightarrow$ stationary point: From $F_x^{'} = 0$ and $F_z^{'} = 0$ it follows, by subtracting both equations, that $x = z \rightarrow$ stationary points are obtained from $F_x^{'}(x) = e^x - e^{3-2x} = 0 \rightarrow x = 3 - 2x \rightarrow x = 1 \rightarrow z = 1 \rightarrow$ Stationary point: $(x, z) = (1, 1)$.
$HF = \begin{bmatrix} e^x + e^{3-x-z} & e^{3-x-z} \\ e^{3-x-z} & e^z + e^{3-x-z} \end{bmatrix} \rightarrow HF(1, 1) = \begin{bmatrix} 2e & e \\ e & 2e \end{bmatrix} \rightarrow F$ has a local minimum at $(x, z) = (1, 1)$.
Note: $HF(x, z)$ is convex $\rightarrow F$ has a global minimum at $(x, z) = (1, 1)$.

$\rightarrow$ Problem has global minimum at $(1, 2, 1)$.

EXERCISE 3.2 Determine all optima (location, value, kind (local/global)) of the next constrained optimization problems.

a. Optimize $f(x,y) = xy^2$, subject to the constraint $g(x,y) := x^2 + 2y^2 = 3$.

b. Optimize $f(x,y) = \frac{x}{y} + \frac{y}{x}$, subject to the constraint $g(x,y) := xy = 1$.

c. Optimize $f(x,y) = \frac{1}{3}y^2e^x - y^2$, subject to the constraint $g(x,y) := e^x - y = -1$.

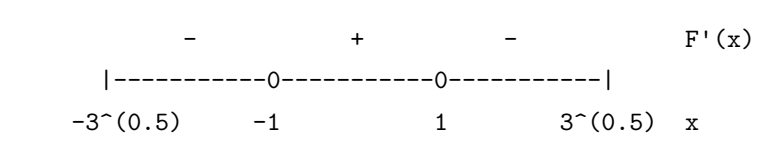
d. Optimize $f(x,y,z) = xz^2 - 7x - 10y + 6xy$ subject to the constraints $g_1(x,y,z) := x^2 + y^2 + z^2 = 9$ and $g_2(x,y,z) := x - y = 0$.

e. Optimize $f(x,y,z) = x^2 + y^2 + z^2$, subject to the constraints $g_1(x,y,z) := x^2 + y^2 = 4$ and $g_2(x,y,z) = x^2 + z^2 = 9$.

f. Optimize $f(x,y,z) = (z-3)^2 - xy$, subject to the constraint $g_1(x,y,z) := (x-1)^2 + (y-2)^2 + (z-3)^2 = 9$ and $g_2(x,y,z) := x + y = 0$.

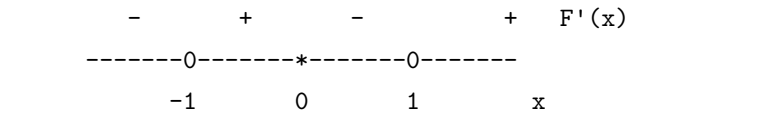
ANSWER 3.2

a. $g(x,y) \to y^2 = \frac{3}{2} - \frac{1}{2}x^2 \to F(x) = x(\frac{3}{2} - \frac{1}{2}x^2)$, with $\frac{3}{2} - \frac{1}{2}x^2 \geq 0$
$\to F(x) = x(\frac{3}{2} - \frac{1}{2}x^2)$, with $x \in [-\sqrt{3}, \sqrt{3}]$.
$F'(x) = \frac{3}{2} - \frac{3}{2}x^2 \to$ stationary points: $x = -1$ and $x = 1$.



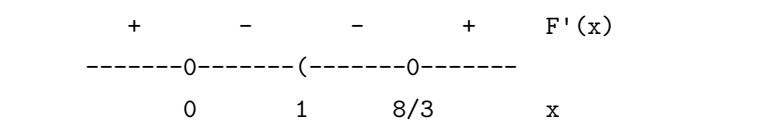
$\to F$ has a local maximum at $x = -\sqrt{3}$ and $x = 1$; and a local minimum at $x = -1$ and $x = \sqrt{3}$.
Note: $F(-\sqrt{3}) = F(\sqrt{3}) = 0$; $F(-1) = -1$; and $F(1) = 1 \to F$ has local maximum at $x = -\sqrt{3}$ a global maximum at $x = 1$; a global minimum at $x = -1$ and a local minimum at $x = \sqrt{3}$.
$\to$ Problem has a local maximum at $(-\sqrt{3}, 0)$; a global maximum at $(1,1)$ and $(1,-1)$; a global minimum at $(-1,1)$ and $(-1,-1)$; and a local minimum at $(\sqrt{3}, 0)$.

b. $g(x,y) \to y = \frac{1}{x},\ x \neq 0 \to F(x) = x^2 + \frac{1}{x^2}$, with $x^2 \neq 0$
$F'(x) = 2x - 2x^{-3} \to$ stationary points: $x = -1$ and $x = 1$.



$\to F$ has a local minimum at $x = -1$ and $x = 1$. Note: F has a vertical asymptote at $x = 0$, where F goes to infinity; and $F(1) = F(-1) = 2 \to F$ has a global minimum at $x = -1$ and $x = 1$.
$\to$ Problem has a global minimum at $(-1,-1)$ and $(1,1)$.

c. $g(x,y) \to e^x = y - 1 \to F(y) = \frac{1}{3}y^2(y-1) - y^2$, with $y > 1$
$F'(y) = y^2 - \frac{8}{3}y \to$ stationary points: $y = 0$ and $y = \frac{8}{3}$.



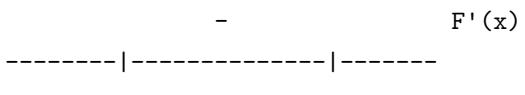
$\rightarrow F$ has a global minimum at $y = \frac{8}{3}$.

$\rightarrow$ Problem has a global minimum at $(\ln(\frac{5}{3}), \frac{8}{3})$.

d. $g_2(x, y) \rightarrow y = x$ and $g_1(x, y) \rightarrow z^2 = 9 - 2x^2$

$\rightarrow F(x) = x(9 - 2x^2) - 17x + 6x^2 = -2x^3 - 8x + 6x^2$, with $x \in [-a, a]$, where $a = \frac{3}{2}\sqrt{2}$.

$F^{'}(x) = -6x^2 + 12x - 8 < 0 \rightarrow$ no stationary points.

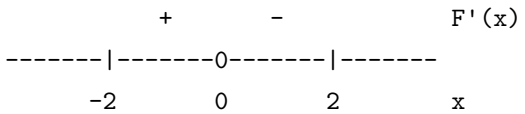


$\rightarrow F$ has a global minimum at $x = a$ and a global maximum at $x = -a$.

$\rightarrow$ Problem has a global minimum at $(a, a, 0)$ and a global maximum at $(-a, -a, 0)$.

e. $g_1(x, y) \rightarrow y^2 = 4 - x^2$ and $g_2(x, y) \rightarrow z^2 = 9 - x^2 \rightarrow F(x) = 13 - x^2$, with $x \in [-2, 2]$.

$F^{'}(x) = -2x \rightarrow$ stationary point: $x = 0$.



$\rightarrow F$ has a global maximum at $x = 0$. Note: $F(2) = F(-2) = 9 \rightarrow F$ has a global minimum at $x = 2$ and $x = -2$

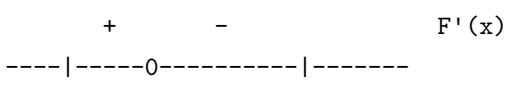
$\rightarrow$ Problem has a global maximum at $(0, 2, 3)$, $(0, 2, -3)$, $(0, -2, 3)$ and $(0, -2, -3)$; and a global minimum at $(2, 0, \sqrt{5})$, $(2, 0, -\sqrt{5})$, $(-2, 0, \sqrt{5})$ and $(-2, 0, -\sqrt{5})$.

f. $g_1(x, y) \rightarrow (z - 3)^2 = 9 - (x - 1)^2 - (y - 2)^2$ and $g_2(x, y) \rightarrow y = -x$

$\rightarrow F(x) = 9 - (x - 1)^2 - (x + 2)^2 + x^2$, with $9 - (x - 1)^2 - (x + 2)^2 \geq 0$

$\rightarrow F(x) = -x^2 - 2x + 4$, with $x \in [-2, 1]$.

$F^{'}(x) = -2x - 2 \rightarrow$ stationary point: $x = -1$.



$\rightarrow F$ has a global maximum at $x = -1$. Note: $F(-2) = 4$ and $F(1) = 1 \rightarrow F$ has a global minimum at $x = 1$ and a local minimum at $x = -2$

$\rightarrow$ Problem has a global maximum at $(-1, 1, 5)$ and $(-1, 1, 1)$; a global minimum at $(1, -1, 0)$ and a local minimum at $(-2, 2, 0)$.

3.2 Two variables and one constraint: a Geometric Approach

EXERCISE 3.3 Solve the next constrained optimization problems.

a. Maximize $f(K, L) = KL$ subject to $K + L = 36$ (you may assume a maximum exists).

b. Optimize $f(x,y) = 2x + 3y$ subject to $x^2 + xy + y^2 = 21$.

c. Maximize $f(x,y) = x^{\frac{2}{3}}y^{\frac{1}{3}}$ $(x > 0,\ y > 0)$ subject to $x + 2y = 10$ (you may assume a maximum exists).

d. Optimize $f(x,y) = x^2 + y^2$ subject to $x^2 + xy + y^2 = 3$.

ANSWER 3.3

a. $Df = [L\ K]$, $Dg = [1\ 1] \rightarrow$ FOC (3.4): $\frac{1}{1} = \frac{L}{K} \rightarrow L = K$; Note: $K + L = 36 \rightarrow K = L = 18$.

b. $Df = [2\ 3]$, $Dg = [2x + y\ x + 2y] \rightarrow$ FOC (3.4): $\frac{2x+y}{x+2y} = \frac{2}{3} \rightarrow y = 4x$; Note: $x^2 + xy + y^2 = 21 \rightarrow x^2 = 1 \rightarrow x = 1$ or $x = -1 \rightarrow (1,4)$ (with $f = 14$) or $(-1,-4)$ (with $f = -14$). Note: constraint set is compact, f is continuous $\rightarrow$ global max at $(1,4)$ and global min at $(-1,-4)$.

c. $Df = [\frac{2}{3}x^{-\frac{1}{3}}y^{\frac{1}{3}}\ \ \frac{1}{3}x^{\frac{2}{3}}y^{-\frac{2}{3}}]$, $Dg = [1\ 2] \rightarrow$ FOC (3.4): $\frac{1}{2} = \frac{2y}{x} \rightarrow x = 4y$; Note: $x + 2y = 10 \rightarrow y = \frac{5}{3},\ x = \frac{20}{3}$.

d. $Df = [2x\ 2y]$, $Dg = [2x + y\ x + 2y] \rightarrow$ FOC (3.4): $\frac{2x+y}{x+2y} = \frac{x}{y} \rightarrow y^2 = x^2 \rightarrow y = x$ or $y = -x$; Note: $x^2 + xy + y^2 = 3 \rightarrow x^2 = 1$, with $y = x$, or $x^2 = 3$, with $y = -x \rightarrow (1,1)$ and $(-1,-1)$, where $f = 2$, or $(\sqrt{3}, -\sqrt{3})$ and $(-\sqrt{3}, \sqrt{3})$, where $f = 6$. Note: constraint set is compact, f is continuous $\rightarrow$ global max at $(\sqrt{3}, -\sqrt{3})$ and $(-\sqrt{3}, \sqrt{3})$ and global min at $(1,1)$ and $(-1,-1)$.

EXERCISE 3.4 Determine for next constrained sets $S := \{(x,y) \mid f(x,y) = b\}$ all points (x_0, y_0) where in an interval around x_0 there exists a continuously differentiable function $y = h(x)$ such that, according the Implicit Function Theorem, $g(x, h(x)) = b$. Calculate $h'(x_0)$ at those points.

a. $S = \{(x,y) \mid (x-y)^2 = 4\}$.

b. $S = \{(x,y) \mid (x-1)^2 + (y-2)^2 = 4\}$.

ANSWER 3.4

a. Consider $g(x,y) = (x-y)^2 \rightarrow Dg = [2(x-y)\ \ -2(x-y)] \rightarrow D_2g(x,y) = 0$ if $-2(x-y) = 0 \rightarrow D_2g(x,y) = 0$ if $x = y \rightarrow$ We cannot conclude, using the Implicit Funtion Theorem, that $y = h(x)$ exists at those points where $x = y$ and $g(x,y) = 4 \rightarrow$ Implicit Funtion Theorem applies everywhere. $h'(x_0) = -\frac{D_1g(x_0,y_0)}{D_2g(x_0,y_0)} = -\frac{2(x_0-y_0)}{-2(x_0-y_0)} = 1$ at all points (x_0, y_0) satisfying $(x_0 - y_0)^2 = 4$.

b. Consider $g(x,y) = (x-1)^2 + (y-2)^2 \rightarrow Dg = [2(x-1)\ 2(y-2)] \rightarrow D_2g(x,y) = 0$ if $2(y-2) = 0 \rightarrow D_2g(x,y) = 0$ if $y = 2 \rightarrow$ We cannot conclude, using the Implicit Funtion Theorem, that $y = h(x)$ exists at those points where $y = 2$ and $g(x,y) = 4 \rightarrow$ We cannot conclude, using the Implicit Funtion Theorem that $y = h(x)$ exists at $(-1,2)$ and $(3,2)$.
Otherwise, $h'(x_0) = -\frac{D_1g(x_0,y_0)}{D_2g(x_0,y_0)} = -\frac{2(x_0-1)}{2(y_0-2)}$ at all points (x_0, y_0) satisfying $(x_0-1)^2 + (y_0-2)^2 = 4$.

EXERCISE 3.5 Determine, using Theorem 3.9, all points (x,y) where an extremum may occur for next constrained optimization problems.

a. Optimize $x^2 + y^2$ subject to $xy = 1$.

b. Optimize $e^{x^2+y^2}$ subject to $x^4 + y^4 = 2$.

ANSWER 3.5

a. $Df = [2x\ 2y]$, $Dg = [y\ x] \rightarrow$ FOC (3.4): $\frac{2x}{2y} = \frac{y}{x} \rightarrow y^2 = x^2 \rightarrow y = \pm x$; Note: $xy = 1 \rightarrow x^2 = 1$ (with $y = x$) or $x^2 = -1$ (with $y = -x$) $\rightarrow x = 1$ (with $y = 1$) or $x = -1$ (with $y = -1$).

b. $Df = [2xe^{x^2+y^2}\ 2ye^{x^2+y^2}]$, $Dg = [4x^3\ 4y^3] \rightarrow$ FOC (3.4): $\frac{4x^3}{4y^3} = \frac{2xe^{x^2+y^2}}{2ye^{x^2+y^2}} \rightarrow \frac{x^3}{y^3} = \frac{x}{y} \rightarrow y^2 = x^2 \rightarrow y = \pm x$; Note: $x^4 + y^4 = 2 \rightarrow x^4 = 1$ (with $y = \pm x$) $\rightarrow x = 1$ (with $y = \pm 1$) or $x = -1$ (with $y = \pm 1$).

EXERCISE 3.6 Use Lagrange's method to solve the next constrained optimization problems.

a. Optimize $f(x, y) = xy$ subject to $4x^2 + y^2 = 32$.

b. Optimize $f(x, y) = x + y$ subject to $x^2 + xy + y^2 = 3$.

c. Optimize $Q(K, L) = K + 20L$ subject to $\sqrt{K} + L = 30$, $K \geq 0$, $L \geq 0$.

d. Maximize $f(x, y) = x$ subject to $x^3 + y^2 = 0$.

ANSWER 3.6

a. Following Algorithm 3.13:

Step 1. $f' = [y\ x]$ and $g' = [8x\ 2y]$. Clearly, all partial derivatives are continuous. So, $f, g \in C^1$.

Step 2. $g'(x, y) = [0\ 0]$ only in case $(x, y) = (0, 0)$. However, $g(0, 0) = 0 \neq 32$. So, $(0, 0)$ is not in the feasible region. We, therefore, don't find any candidate solutions in this step.

Step 3. The Lagrange function is $L(x, y, \lambda) = xy - \lambda(4x^2 + y^2 - 32)$.

Step 4. Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L' = [y - 8\lambda x\ \ x - 2\lambda y\ \ 4x^2 + y^2 - 32] = [0\ 0\ 0].$$

$L_1' = 0 \rightarrow y = 8\lambda x \rightarrow$ Substitution into $L_2' = 0$ yields: $x - 16\lambda^2 x = 0 \rightarrow x = 0$ (with $y = 0$) or $\lambda = \frac{1}{4}$ (with $y = 2x$) or $\lambda = -\frac{1}{4}$ (with $y = -2x$) $\rightarrow (0, 0)$, but this does not satisfy $g(x, y) = 32$, or $\lambda = \frac{1}{4}$ (with $y = 2x$ and $8x^2 = 32$) or $\lambda = -\frac{1}{4}$ (with $y = -2x$ and $8x^2 = 32$) $\rightarrow (2, 4)$ or $(-2, -4)$, with $\lambda = \frac{1}{4}$, or $(2, -4)$ or $(-2, 4)$, with $\lambda = -\frac{1}{4}$.

Note: the constraint set is compact. So, by Weierstrass' theorem both a global maximum and minimum exist.

$f(2, 4) = f(-2, -4) = 8$; $f(-2, 4) = f(2, -4) = -8 \rightarrow$ Maximum occurs at $(2, 4)$ and $(-2, -4)$; minimum occurs at $(-2, 4)$ and $(2, -4)$.

b. Following Algorithm 3.13:

Step 1. $f' = [1\ 1]$ and $g' = [2x + y\ x + 2y]$. Clearly, all partial derivatives are continuous. So, $f, g \in C^1$.

Step 2. $g'(x, y) = [0\ 0]$ only in case $(x, y) = (0, 0)$. However, $g(0, 0) = 0 \neq 3$. So, $(0, 0)$ is not in the feasible region. We, therefore, don't find any candidate solutions in this step.

Step 3. The Lagrange function is $L(x, y, \lambda) = x + y - \lambda(x^2 + xy + y^2 - 3)$.

Step 4. Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [1 - 2\lambda x - \lambda y \;\; 1 - \lambda x - 2\lambda y \;\; x^2 + xy + y^2 - 3] = [0\ 0\ 0].$$

Notice $\lambda = 0$ does not yield a solution $\rightarrow [L_1^{'}\ L_2^{'}] = [0\ 0]$ yields $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
$\rightarrow x = y = \frac{1}{3\lambda} \rightarrow$ Substitution into $L_3^{'} = 0$ yields: $\lambda^2 = \frac{1}{9} \rightarrow \lambda = \frac{1}{3}$ (with $x = y = 1$) or $\lambda = -\frac{1}{3}$ (with $x = y = -1$).

Note: the constraint set is compact. So, by Weierstrass' theorem both a global maximum and minimum exist.

$f(1,1) = 2$; $f(-1,-1) = -2$. $\rightarrow$ Maximum occurs at $(1,1)$; minimum occurs at $(-1,1)$.

c. Following Algorithm 3.13:

Step 1. $f^{'} = [1\ 20]$ and $g^{'} = [\frac{1}{2}K^{-\frac{1}{2}}\ 1]$. Clearly, all partial derivatives are continuous. So, $f, g \in C^1$.

Step 2. $g^{'}(x,y) = [0\ 0]$ does not occur. We, therefore, don't find any candidate solutions in this step.

Step 3. The Lagrange function is $L(K,L,\lambda) = K + 20L - \lambda(\sqrt{K} + L - 30)$.

Step 4. Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [1 - \frac{1}{2}\frac{\lambda}{\sqrt{K}} \;\; 20 - \lambda \;\; \sqrt{K} + L - 30] = [0\ 0\ 0].$$

$L_2^{'} = 0 \rightarrow \lambda = 20$. Substitution into $L_1^{'} = 0$ yields: $K = 100$. $\rightarrow$ Substitution into $L_3^{'} = 0$ yields:$L = 20$.

Note: the constraint set is compact. So, by Weierstrass' theorem both a global maximum and minimum exist.

$Q(100,20) = 500$; $Q(0,30) = 600$, $Q(900,0) = 900 \rightarrow$ Maximum occurs at $(900,0)$; minimum occurs at $(100,20)$.

d. Following Algorithm 3.13:

Step 1. $f^{'} = [1\ 0]$ and $g^{'} = [3x^2\ 2y]$. Clearly, all partial derivatives are continuous. So, $f, g \in C^1$.

Step 2. $g^{'}(x,y) = [0\ 0]$ only in case $(x,y) = (0,0)$. So, $(0,0)$ is a candidate solution in this step.

Step 3. The Lagrange function is $L(x,y,\lambda) = x - \lambda(x^3 + y^2)$.

Step 4. Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [1 - 3\lambda x^2 \;\; -2\lambda y \;\; x^3 + y^2] = [0\ 0\ 0].$$

Notice $\lambda = 0$ does not yield a solution. $\rightarrow L_2^{'} = 0 \rightarrow y = 0 \rightarrow$ Substitution into $L_3^{'} = 0$ yields: $x = 0$, but $(0,0)$ does not satisfy $L_1^{'} = 0$. $\rightarrow$ only candidate solution is $(0,0)$.

Note $f(0,0) = 0$. Furthermore, it follows from the restriction $x^3 = -y^2$ that $x \leq 0 \rightarrow f(x,y) \leq 0$ $\rightarrow f$ has maximum at $(0,0)$.

EXERCISE 3.7 Use Lagrange's method to solve the constrained optimization problems from Exercise 3.3.

ANSWER 3.7

a. Following Algorithm 3.13:

Step 1. $f^{'} = [L\ K]$ and $g^{'} = [1\ 1]$. Clearly, all partial derivatives are continuous. So, $f, g \in C^1$.

Step 2. $g^{'}(x, y \neq [0\ 0]$. We, therefore, don't find any candidate solutions in this step.

Step 3. The Lagrange function is $L(K, L, \lambda) = KL - \lambda(K + L - 36)$.

Step 4. Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [L - \lambda\ \ K - \lambda\ \ K + L - 36] = [0\ 0\ 0].$$

$[L_1^{'}\ L_2^{'}] = [0\ 0] \rightarrow K = L = \lambda \rightarrow$ Substitution into $L_3^{'} = 0$ yields: $\lambda = 18 \rightarrow$ only candidate solution: $(18, 18)$ with $\lambda = 18$.

b. Following Algorithm 3.13:

Step 1. $f^{'} = [2\ 3]$ and $g^{'} = [2x + y\ x + 2y]$. Clearly, all partial derivatives are continuous. So, $f, g \in C^1$.

Step 2. $g^{'}(x, y) = [0\ 0]$ only holds at $(0, 0)$, but $g(0, 0) = 0 \neq 21$. We, therefore, don't find any candidate solutions in this step.

Step 3. The Lagrange function is $L(x, y, \lambda) = 2x + 3y - \lambda(x^2 + xy + y^2 - 21)$.

Step 4. Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [2 - 2\lambda x - \lambda y\ \ 3 - \lambda x - 2\lambda y\ \ \ x^2 + xy + y^2 - 21] = [0\ 0\ 0].$$

Notice $\lambda = 0$ does not yield a solution $\rightarrow [L_1^{'}\ L_2^{'}] = [0\ 0]$ yields $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

$\rightarrow x = \frac{1}{3\lambda};\ y = \frac{4}{3\lambda}$. $\rightarrow$ Substitution into $L_3^{'} = 0$ yields:$\lambda = \pm\frac{1}{3} \rightarrow (1, 4)$, with $\lambda = \frac{1}{3}$, or $(-1, -4)$, with $\lambda = -\frac{1}{3}$.

Note: the constraint set is compact. So, by Weierstrass' theorem both a global maximum and minimum exist.

$f(1, 4) = 14$; $f(-1, -4) = -14 \rightarrow$ Maximum occurs at $(1, 4)$; minimum occurs at $(-1, -4)$.

c. Following Algorithm 3.13:

Step 1. $f^{'} = [\frac{2}{3}x^{-\frac{1}{3}}y^{\frac{1}{3}}\ \ \frac{1}{3}x^{\frac{2}{3}}y^{-\frac{2}{3}}]$ and $g^{'} = [1\ 2]$. Clearly, all partial derivatives are continuous. So, $f, g \in C^1$.

Step 2. $g^{'}(x, y) \neq [0\ 0]$. We, therefore, don't find any candidate solutions in this step.

Step 3. The Lagrange function is $L(x, y, \lambda) = x^{\frac{2}{3}}y^{\frac{1}{3}} - \lambda(x + 2y - 10)$.

Step 4. Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [\frac{2}{3}x^{-\frac{1}{3}}y^{\frac{1}{3}} - \lambda\ \ \frac{1}{3}x^{\frac{2}{3}}y^{-\frac{2}{3}} - 2\lambda\ \ x + 2y - 10] = [0\ 0\ 0].$$

Notice $\lambda = 0$ does not yield a solution $\rightarrow [L_1^{'}\ L_2^{'}] = [0\ 0]$ yields $\lambda = \frac{2}{3}x^{-\frac{1}{3}}y^{\frac{1}{3}} = \frac{1}{6}x^{\frac{2}{3}}y^{-\frac{2}{3}}$ $\rightarrow x = 4y$. $\rightarrow$ Substitution into $L_3^{'} = 0$ yields:$y = \frac{5}{3} \rightarrow (\frac{20}{3}, \frac{5}{3})$, with $\lambda = \frac{1}{3}$, or $(-1, -4)$, with $\lambda = \frac{2}{3}4^{-\frac{1}{3}}$.

d. Following Algorithm 3.13:

Step 1. $f^{'} = [2x\ 2y]$ and $g^{'} = [2x + y\ x + 2y]$. Clearly, all partial derivatives are continuous. So, $f, g \in C^1$.

Step 2. $g^{'}(x, y) = [0\ 0]$ only holds at $(0, 0)$, but $g(0, 0) = 0 \neq 3$. We, therefore, don't find any candidate solutions in this step.

Step 3. The Lagrange function is $L(x, y, \lambda) = x^2 + y^2 - \lambda(x^2 + xy + y^2 - 21)$.

Step 4. Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [2x - 2\lambda x - \lambda y\ \ 2y - \lambda x - 2\lambda y\ \ \ x^2 + xy + y^2 - 3] = [0\ 0\ 0].$$

$[L_1^{'}\ L_2^{'}] = [0\ 0]$ yields $\begin{bmatrix} 2(1-\lambda) & -\lambda \\ -\lambda & 2(1-\lambda) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Note: $(x, y) = (0, 0)$ does not provide a solution $\rightarrow$ matrix $\begin{bmatrix} 2(1-\lambda) & -\lambda \\ -\lambda & 2(1-\lambda) \end{bmatrix}$ must be singular $\rightarrow \det \begin{bmatrix} 2(1-\lambda) & -\lambda \\ -\lambda & 2(1-\lambda) \end{bmatrix} = 0 \rightarrow 4(1-\lambda)^2 - \lambda^2 = 0 \rightarrow \lambda = \frac{2}{3}$ or $\lambda = 2$. $\rightarrow \lambda = \frac{2}{3}$ implies $x = y$. So, $x^2 = 1$. $\lambda = 2$ implies $y = -x$. So, $y^2 = 3$.

Note: the constraint set is compact. So, by Weierstrass' theorem both a global maximum and minimum exist.

The results from Exercise 4.3.d follow now immediately again.

EXERCISE 3.8 A consumer has a utility function $U(x, y) = bx^{\alpha}y^{1-\alpha}$, where x and y are the amount of product 1 and 2 the consumer can choose from, respectively. Assume the price of product 1 is p_1, the price of product 2 is p_2 and his income is I.

a. Assume the consumer wants to maximize his utility. Formulate the corresponding constrained optimization problem.

b. Determine the solution of this constrained optimization problem.

c. Calculate the approximate change in the demand for good x and utility U if the income, I, is changed by one unit.

ANSWER 3.8

a. Maximize $U(x, y)$ subject to $I = p_1x + p_2y$.

b. $\frac{U_x^{'}}{U_y^{'}} = \frac{p_1}{p_2} \rightarrow \frac{\alpha b x^{\alpha-1}y^{1-\alpha}}{(1-\alpha)bx^{\alpha}y^{-\alpha}} = \frac{p_1}{p_2} \rightarrow \frac{\alpha y}{(1-\alpha)x} = \frac{p_1}{p_2} \rightarrow y = \frac{(1-\alpha)p_1}{\alpha p_2}x$.
Substitution of this into the budget constraint yields: $I = (p_1 + \frac{(1-\alpha)p_1}{\alpha})x \rightarrow x = \frac{\alpha}{p_1}I$ and $y = \frac{1-\alpha}{p_2}I$.

c. $\frac{\partial x}{\partial I} = \frac{\alpha}{p_1} \rightarrow$ If I changes by one unit, x will approximately change by $\frac{\alpha}{p_1}$.
$U(I) = bx^{\alpha}y^{1-\alpha} = b(\frac{\alpha}{p_1}I)^{\alpha}(\frac{1-\alpha}{p_2}I)^{1-\alpha} = b(\frac{\alpha}{p_1})^{\alpha}(\frac{1-\alpha}{p_2})^{1-\alpha}I \rightarrow$ If I changes by one unit, U will

approximately change by $b(\frac{\alpha}{p_1})^{\alpha}(\frac{1-\alpha}{p_2})^{1-\alpha}$.

Note: This result can also be obtained by calculating the value of the costate variable if one uses the Lagrange method to solve the problem.

3.3 The Envelope Theorem and Lagrange Multiplier

EXERCISE 3.9

a. What is the effect if the parameter $a = a_0$ $(a > 0)$ is increased by one unit on the maximum value of $f(x, a) = -x^4 + ax + a^3$.

b. What is the effect if the parameter $a = a_0$ $(a > 0)$ is increased by one unit on the maximum value of $f(x, y, a) = -x^2 - y^4 + ax + a^2y^2 + 4$.

ANSWER 3.9

a. $f' = [-4x^3 + a\ x + 3a^2] \rightarrow f'' = \begin{bmatrix} -12x^2 & 1 \\ 1 & 6a \end{bmatrix} \rightarrow$ All second-order partial derivatives exist and are continuous $\rightarrow f \in C^2$.

$D^2f = -12x^2$, so at maximum location (with $x_0 = (\frac{1}{4}a)^{\frac{1}{3}}$), D^2f is invertible. $\rightarrow$ by Envelope Theorem, $\frac{df}{da} = x_0 + 3a^2 = (\frac{1}{4}a)^{\frac{1}{3}} + 3a^2$.

b. $f' = [-2x + a \quad -4y^3 + 2a^2y \quad x + 2ay^2] \rightarrow f'' = \begin{bmatrix} -2 & 0 & 1 \\ 0 & -12y^2 + 2a^2 & 4ay \\ 1 & 4ay & 2y^2 \end{bmatrix} \rightarrow$ All second-order partial derivatives exist and are continuous $\rightarrow f \in C^2$.

$D^2f = \begin{bmatrix} -2 & 0 \\ 0 & -12y^2 + 2a^2 \end{bmatrix}$, so at the maximum locations (with $x_0 = \frac{1}{2}a$ and $y_0 = \pm\sqrt{\frac{1}{2}a}$), D^2f is invertible. $\rightarrow$ by Envelope Theorem, $\frac{df}{da} = x_0 + 2ay_0^2 = \frac{1}{2}a + a^3$.

EXERCISE 3.10 *(profit maximization by a firm)*

Consider a firm producing a good. The cost of the firm for producing x units of the good is $C(x)$. Assume that this cost function has the next monotonicity properties: $C'(x) > 0$ and $C''(x) > 0$. Assume that a fraction α of the amount of produced product is sold on the market at a price p per unit. Show there is a unique amount of goods where profit is maximized. Calculate by how much maximum profit increases if α increases.

ANSWER 3.10 Profit: $\pi(x, \alpha) = \alpha px - C(x) \rightarrow \pi' = [\alpha p - C' \quad px] \rightarrow \pi'' = \begin{bmatrix} -C'' & p \\ p & 0 \end{bmatrix} \rightarrow$ All second-order partial derivatives exist and are continuous $\rightarrow \pi \in C^2$.

$D^2\pi = -C''$ is invertible. $\rightarrow$ by Envelope Theorem, $\frac{d\pi}{d\alpha} = px_0$, where x_0 solves $\alpha p - C'(x_0) = 0$. (Note that due to the monotonicity properties of the cost function, there is at most one solution to this equation, which yields a maximum).

EXERCISE 3.11 Reconsider items a), b) and c) from Exercise 3.6. Use Lagrange's method to solve the next constrained optimization problems. Compare the optimum value with the value predicted by Corollary 3.19, using the Lagrange parameter you calculated in Exercise 3.6 for the corresponding item.

a. Optimize $f(x, y) = xy$ subject to $4x^2 + y^2 = 32.01$.

b. Optimize $f(x, y) = x + y$ subject to $x^2 + xy + y^2 = 4$.

c. Optimize $Q(K, L) = K + 20L$ subject to $\sqrt{K} + L = 29.9$, $K \geq 0$, $L \geq 0$.

ANSWER 3.11

a. We use Algorithm 3.13 to solve the problem. For Step 1-Step 3 see Exercise 3.6.

Step 4.: Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [y - 8\lambda x \;\; x - 2\lambda y \;\; 4x^2 + y^2 - 32.01] = [0\ 0\ 0].$$

$L_1^{'} = 0 \rightarrow y = 8\lambda x \rightarrow$ Substitution into $L_2^{'} = 0$ yields: $x - 16\lambda^2 x = 0 \rightarrow x = 0$ (with $y = 0$) or $\lambda = \frac{1}{4}$ (with $y = 2x$) or $\lambda = -\frac{1}{4}$ (with $y = -2x$) $\rightarrow (0, 0)$, but this does not satisfy $g(x, y) = 32$, or $\lambda = \frac{1}{4}$ (with $y = 2x$ and $8x^2 = 32.01$) or $\lambda = -\frac{1}{4}$ (with $y = -2x$ and $8x^2 = 32.01$) $\rightarrow$ $(2.0003, 4.0006)$ or $(-2.0003, -4.0006)$, with $\lambda = \frac{1}{4}$, or $(2.0003, -4.0006)$ or $(-2.0003, 4.0006)$, with $\lambda = -\frac{1}{4}$.

Note: the constraint set is compact. So, by Weierstrass' theorem both a global maximum and minimum exist.

$f(2.0003, 4.0006) = f(-2.0003, -4.0006) = 8.0025$; $f(-2.0003, 4.0006) = f(2.0003, -4.0006) = -8.0025 \rightarrow$ Maximum occurs at $(2.0003, 4.0006)$ and $(-2.0003, -4.0006)$; minimum occurs at $(-2.0003, 4.0006)$ and $(2.0003, -4.0006)$.

The solutions predicted by Exercise 3.6 are: for the maximum value $8 + 0.01 * \frac{1}{4}$; for the minimum value $-8 - 0.01 * \frac{1}{4}$.

b. We use Algorithm 3.13 to solve the problem. For Step 1-Step 3 see Exercise 3.6.

Step 4.: Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [1 - 2\lambda x - \lambda y \;\; 1 - \lambda x - 2\lambda y \;\; x^2 + xy + y^2 - 4] = [0\ 0\ 0].$$

Notice $\lambda = 0$ does not yield a solution $\rightarrow [L_1^{'}\ L_2^{'}] = [0\ 0]$ yields $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$\rightarrow x = y = \frac{1}{3\lambda} \rightarrow$ Substitution into $L_3^{'} = 0$ yields: $\lambda^2 = \frac{1}{12} \rightarrow \lambda = \frac{1}{\sqrt{12}}$ (with $x = y = \frac{2}{3}\sqrt{3}$) or $\lambda = -\frac{1}{\sqrt{12}}$ (with $x = y = -\frac{2}{3}\sqrt{3}$).

Note: the constraint set is compact. So, by Weierstrass' theorem both a global maximum and minimum exist.

Maximum: $f(\frac{2}{3}\sqrt{3}, \frac{2}{3}\sqrt{3}) = \frac{4}{3}\sqrt{3}$; Minimum: $f(-\frac{2}{3}\sqrt{3}, -\frac{2}{3}\sqrt{3}) = -\frac{4}{3}\sqrt{3}$.

Solutions predicted by Exercise 3.6: for the maximum value, $2 + \frac{1}{3}$; for the minimum value, $-2 - \frac{1}{3}$.

c. We use Algorithm 3.13 to solve the problem. For Step 1-Step 3 see Exercise 3.6.
Step 4.: Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [1 - \frac{1}{2}\frac{\lambda}{\sqrt{K}} \;\; 20 - \lambda \;\; \sqrt{K} + L - 29.9] = [0\ 0\ 0].$$

$L_2^{'} = 0 \rightarrow \lambda = 20$. Substitution into $L_1^{'} = 0$ yields: $K = 100$. $\rightarrow$ Substitution into $L_3^{'} = 0$ yields: $L = 19.9$.
Note: the constraint set is compact. So, by Weierstrass' theorem both a global maximum and minimum exist.
Minimum $Q(100, 19.9) = 498$.
Solution predicted by Exercise 3.6: $500 - 20 * .1 = 498$.

3.4 A Second-Order Criterion for a Constrained Optimum

EXERCISE 3.12 Solve next constrained optimization problems. Optimize
a. $f(x, y) = x^2 - y$ subject to $g(x, y) = x^2 - y^2 = 1$.
b. $f(x, y) = 2x^3 + 6y^3$ subject to $g(x, y) = 3x^2 + 9y^2 = 12$.

ANSWER 3.12
a. Following Algorithm 3.13 we have:
Step 1. $f^{'} = [2x \quad -1]$ and $g^{'} = [2x \quad -2y]$. Clearly, all partial derivatives are continuous. So, $f, g \in C^1$.
Step 2. $g^{'}(x, y) = [0\ 0]$ only in case $(x, y) = (0, 0)$, but $g(0, 0) = 0 \neq 1$. We, therefore, don't find any candidate solutions in this step.
Step 3. The Lagrange function is $L(x, y, \lambda) = x^2 - y - \lambda(x^2 - y^2 - 1)$.
Step 4. Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [2x(1 - \lambda) \quad -1 + 2\lambda y \;\; x^2 - y^2 - 1] = [0\ 0\ 0].$$

$L_1^{'} = 0 \rightarrow x = 0$ or $\lambda = 1$.
Case $\lambda = 1$: $L_2^{'} = 0 \rightarrow y = \frac{1}{2}$. Substitution into $L_3^{'} = 0 \rightarrow x = \pm\frac{1}{2}\sqrt{5} \rightarrow$ Candidate solutions: $(\frac{1}{2}\sqrt{5}, \frac{1}{2})$ and $(-\frac{1}{2}\sqrt{5}, \frac{1}{2})$, with $\lambda = 1$.
Case $x = 0$: $L_3^{'} = 0 \rightarrow$ no solution.
Note: constraint set is not compact. To determine what kind of optimum occurs at the candidate solutions consider the determinant of the function HL (Theorem 3.23) at these points.

$$HL(x, y, \lambda) = \begin{bmatrix} 2(1-\lambda) & 0 & -2x \\ 0 & 2\lambda & 2y \\ -2x & 2y & 0 \end{bmatrix}. \rightarrow$$

$HL(\frac{1}{2}\sqrt{5}, \frac{1}{2}, 1) = \begin{bmatrix} 0 & 0 & -\sqrt{5} \\ 0 & 2 & 1 \\ -\sqrt{5} & 1 & 0 \end{bmatrix}$. $\rightarrow$ $\det(HL) = -10 \rightarrow$ Theorem 3.23: problem has a minimum at $(\frac{1}{2}\sqrt{5}, -\frac{1}{2})$.

$HL(-\frac{1}{2}\sqrt{5}, -\frac{1}{2}, 1) = \begin{bmatrix} 0 & 0 & \sqrt{5} \\ 0 & 2 & 1 \\ \sqrt{5} & 1 & 0 \end{bmatrix}$. $\rightarrow$ $\det(HL) = -10 \rightarrow$ Theorem 3.23: problem has a minimum at $(-\frac{1}{2}\sqrt{5}, -\frac{1}{2})$.

Note this is a global minimum as the problem is equivalent with finding the optima of $F(y) = y^2 - y + 1$.

b. Following Algorithm 3.13 we have:

Step 1. $f^{'} = [4x^2 \;\; 18y^2]$ and $g^{'} = [6x \;\; 18y]$. Clearly, all partial derivatives are continuous. So, $f, g \in C^1$.

Step 2. $g^{'}(x, y) = [0\ 0]$ only in case $(x, y) = (0, 0)$, but $g(0, 0) = 0 \neq 12$. We, therefore, don't find any candidate solutions in this step.

Step 3. The Lagrange function is $L(x, y, \lambda) = 2x^3 + 6y^3 - \lambda(3x^2 + 9y^2 - 12)$.

Step 4. Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [6x(x - \lambda) \;\; 18y(y - \lambda) \;\; 3x^2 + 9y^2 - 12] = [0\ 0\ 0].$$

$L_1^{'} = 0 \rightarrow \lambda = x$ or $x = 0$.

Case $\lambda = x$: $L_2^{'} = 0 \rightarrow y = 0$ or $y = x$.

i. $\lambda = x$, $y = 0$: $L_3^{'} = 0 \rightarrow x = \pm 2$.

$\rightarrow$ candidate solutions: $(2, 0)$ with $\lambda = 2$, and $(-2, 0)$ with $\lambda = -2$.

ii. $\lambda = x$, $y = x$: $L_3^{'} = 0 \rightarrow x = \pm 1$.

$\rightarrow$ candidate solutions: $(1, 1)$ with $\lambda = 1$, and $(-1, -1)$ with $\lambda = -1$.

Case $x = 0$: $L_3^{'} = 0 \rightarrow y = \pm\sqrt{\frac{4}{3}} \rightarrow L_2^{'} = 0$ gives $\lambda = y$.

$\rightarrow$ candidate solutions: $(0, \sqrt{\frac{4}{3}})$ with $\lambda = \sqrt{\frac{4}{3}}$, and $(0, -\sqrt{\frac{4}{3}})$ with $\lambda = -\sqrt{\frac{4}{3}}$.

Note: constraint set is compact and f is continuous. $f(2, 0) = 16$; $f(-2, 0) = -16$; $f(1, 1) = 8$; $f(-1, -1) = -8$; $f(0, \sqrt{\frac{4}{3}}) = 8\sqrt{\frac{4}{3}}$; $f(0, -\sqrt{\frac{4}{3}}) = -8\sqrt{\frac{4}{3}}$. $\rightarrow$ global maximum at $(2, 0)$ and global minimum at $(-2, 0)$.

To assess what kind of optimum occurs at the other candidates we consider the determinant of the function HL (Theorem 3.23) at these points.

$HL(x, y, \lambda) = \begin{bmatrix} 12x - 6\lambda & 0 & -6x \\ 0 & 36y - 18\lambda & -18y \\ -6x & -18y & 0 \end{bmatrix}$. $\rightarrow$

$HL(1, 1, 1) = \begin{bmatrix} 6 & 0 & -6 \\ 0 & 18 & -18 \\ -6 & -18 & 0 \end{bmatrix}$. $\rightarrow$ determinant $HL < 0 \rightarrow$ Theorem 3.23: problem has a

local minimum at $(1,1)$.

$HL(-1,-1,-1) = \begin{bmatrix} -6 & 0 & 6 \\ 0 & -18 & 18 \\ 6 & 18 & 0 \end{bmatrix}$. $\rightarrow \det(HL) > 0 \rightarrow$ Theorem 3.23: problem has a local maximum at $(-1,-1)$.

$HL(0,\sqrt{\frac{4}{3}},\sqrt{\frac{4}{3}}) = \begin{bmatrix} -6\sqrt{\frac{4}{3}} & 0 & 0 \\ 0 & 18\sqrt{\frac{4}{3}} & -18\sqrt{\frac{4}{3}} \\ 0 & -18\sqrt{\frac{4}{3}} & 0 \end{bmatrix}$. $\rightarrow \det(HL) > 0 \rightarrow$ Theorem 3.23: problem has a local maximum at $(0,\sqrt{\frac{4}{3}})$.

$HL(0,-\sqrt{\frac{4}{3}},-\sqrt{\frac{4}{3}}) = \begin{bmatrix} 6\sqrt{\frac{4}{3}} & 0 & 0 \\ 0 & -18\sqrt{\frac{4}{3}} & 18\sqrt{\frac{4}{3}} \\ 0 & 18\sqrt{\frac{4}{3}} & 0 \end{bmatrix}$. $\rightarrow \det(HL) < 0 \rightarrow$ Theorem 3.23: problem has a local minimum at $(0,-\sqrt{\frac{4}{3}})$.

EXERCISE 3.13 Solve next constrained optimization problems.

a. Optimize $f(x,y) = x^2 + 4y^2 - x^2y$ subject to $g(x,y) = x^2 + y^2 = 3$.

b. Optimize $f(x,y,z) = x^2 + y^2 + z^2 - 6x$ subject to $g(x,y) = x + y - z = 0$.

c. Optimize $f(x,y) = xy + y$ subject to $g(x,y) = x^2 + y^2 = 1$.

d. Minimize $f(x,y) = x + 2y$ subject to $g(x,y) = x^2 + y^2 = \frac{5}{4}$, $x \geq 0$, $y \geq 0$.

ANSWER 3.13

a. From the constraint it follows: $x^2 = 3 - y^2$

$\rightarrow F(y) = 3 + 4y^2 - y(3 - y^2) = y^3 + 4y^2 - 3y + 3$, where $y \in [-\sqrt{3}, \sqrt{3}]$.

$\rightarrow F'(y) = 3y^2 + 8y - 3$. $\rightarrow$ Stationary points: $y = -3$ and $y = \frac{1}{3}$.

```
                    -         +              F'(x)
----0------|----------0-------|-------
    -3   -3^(0.5)      1/3     3^(0.5)     x
```

$\rightarrow F$ has a global minimum at $y = \frac{1}{3}$. Note: $F(-\sqrt{3}) = F(\sqrt{3}) = 15$. $\rightarrow F$ has a global maximum at $y = -\sqrt{3}$ and $y = -\sqrt{3}$.

$\rightarrow$ Problem has global minimum at $(\pm\sqrt{\frac{27}{9}}, \frac{1}{3})$ and a global maximum at $(0, \pm\sqrt{3})$.

b. From the constraint it follows: $y = z - x$. $\rightarrow F(x,z) = x^2 + (z-x)^2 + z^2 - 6x$. $\rightarrow$ Stationary points:

$DF = [2x - 2(z-x) - 6 \;\; 2(z-x) + 2z] = [0 \;\; 0]$. $\rightarrow (x,z) = (2,1)$.

$HF = \begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix}$. $\rightarrow F$ is positive definite. $\rightarrow (2,1)$ is global minimum.

$\rightarrow$ Problem has global minimum at $(2,1,1)$.

c. Following Algorithm 3.13 we have:

Step 1. $f^{'} = [y \;\; x{+}1]$ and $g^{'} = [2x \;\; 2y]$. Clearly, all partial derivatives are continuous. So, $f, g \in C^1$.

Step 2. $g^{'}(x,y) = [0\ 0]$ only in case $(x,y) = (0,0)$, but $g(0,0) = 0 \neq 1$. We, therefore, don't find any candidate solutions in this step.

Step 3. The Lagrange function is $L(x,y,\lambda) = xy + y - \lambda(x^2 + y^2 - 1)$.

Step 4. Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [y - 2\lambda x \;\; x + 1 - 2\lambda y \;\; x^2 + y^2 - 1] = [0\ 0\ 0].$$

$L_1^{'} = 0 \rightarrow y = 2\lambda x. \rightarrow L_2^{'} = 0$ yields: $(1 - 4\lambda^2)x + 1 = 0$. Note, $(1 - 4\lambda^2)$ cannot be zero. $\rightarrow$ $x = -\frac{1}{1-4\lambda^2}$ and $y = -\frac{2\lambda}{1-4\lambda^2}$.

$\rightarrow$ Substitution into $L_3^{'} = 0$ yields: $1 + 4\lambda^2 = (1 - 4\lambda^2)^2$. $\rightarrow \lambda^2(16\lambda^2 - 12) = 0$. $\rightarrow \lambda = 0$ or $\lambda = \pm\sqrt{\frac{3}{4}}$. $\rightarrow$

Candidate solutions: $(-1,0)$ with $\lambda = 0$; $(\frac{1}{2}, -\sqrt{\frac{3}{4}})$ with $\lambda = -\sqrt{\frac{3}{4}}$; and $(\frac{1}{2}, \sqrt{\frac{3}{4}})$ with $\lambda = \sqrt{\frac{3}{4}}$.

Note: constraint set is compact, f is continuous. $f(-1,0) = 0$; $f(\frac{1}{2}, -\sqrt{\frac{3}{4}}) = -\frac{3}{2}\sqrt{\frac{3}{4}}$ and $f(\frac{1}{2}, \sqrt{\frac{3}{4}}) = \frac{3}{2}\sqrt{\frac{3}{4}} \rightarrow$ By Weierstrass' Theorem the problem has a global maximum at $(\frac{1}{2}, \sqrt{\frac{3}{4}})$ and a global minimum at $(\frac{1}{2}, -\sqrt{\frac{3}{4}})$.

To assess what kind of optimum occurs at $(-1,0)$ we consider the determinant of the function HL (Theorem 3.23) at this point.

$$HL(x,y,\lambda) = \begin{bmatrix} -2\lambda & 1 & -2x \\ 1 & -2\lambda & -2y \\ -2x & -2y & 0 \end{bmatrix}. \rightarrow HL(-1,0,0) = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}. \rightarrow \det(HL(-1,0,0)) = 0$$

$\rightarrow$ Theorem 3.23 does not provide information about this point.

Note: $f(-\sqrt{1-\epsilon}, \sqrt{\epsilon}) > 0$ for small positive ϵ, whereas $f(-\sqrt{1-\epsilon}, -\sqrt{\epsilon}) < 0$ for small positive ϵ $\rightarrow f$ has a saddle point at $(-1,0)$.

d. Note: constraint set is compact, f is continuous $\rightarrow$ By Weierstrass' Theorem the problem has a global minimum.

First consider the optimization problem without the assumption $x \geq 0$ and $y \geq 0$.

Following Algorithm 3.13 we have:

Step 1. $f^{'} = [1 \;\; 2]$ and $g^{'} = [2x \;\; 2y]$. Clearly, all partial derivatives are continuous. So, $f, g \in C^1$.

Step 2. $g^{'}(x,y) = [0\ 0]$ only in case $(x,y) = (0,0)$, but $g(0,0) = 0 \neq \frac{5}{4}$. We, therefore, don't find any candidate solutions in this step.

Step 3. The Lagrange function is $L(x,y,\lambda) = x + 2y - \lambda(x^2 + y^2 - \frac{5}{4})$.

Step 4. Differentiation of the Lagrange function shows that the candidate solutions must satisfy

$$L^{'} = [1 - 2\lambda x \;\; 2 - 2\lambda y \;\; x^2 + y^2 - \frac{5}{4}] = [0\ 0\ 0].$$

$L_1^{'} = 0 \rightarrow$ Note, $\lambda = 0$ does not provide a solution $\rightarrow x = \frac{1}{2\lambda}$. Similarly, it follows from $L_2^{'} = 0$ that $y = \frac{1}{\lambda} \rightarrow$ Substitution into $L_3^{'} = 0$ shows $\lambda^2 = 1 \rightarrow \lambda = \pm 1 \rightarrow$

Candidate solutions: $(\frac{1}{2}, 1)$ and $\lambda = (-\frac{1}{2}, -1)$.

Note $f(\frac{1}{2}, 1) = 2.5$ and $f(-\frac{1}{2}, -1) = -2.5 \rightarrow$ Problem, with the constraint set $g(x, y) = \frac{5}{4}$ has a global minimum at $(-\frac{1}{2}, -1)$. However, $(-\frac{1}{2}, -1)$ does not satisfy $x \geq 0$ and $y \geq 0 \rightarrow$ Constraint problem attains its global minimum at the boundary of the constraint set $\rightarrow$ either at $(0, 1)$ or $(1, 0)$. $f(0, 1) = 2$, $f(1, 0) = 1 \rightarrow$ Constraint problem attains its global minimum at $(1, 0)$.

3.5 The Method of Lagrange: the general case

EXERCISE 3.14

a. Consider the equation

$$g(x, s, t, u) = e^{x+s+t+u} - x^2 - s^2 - t^2 - u^2 = 1 \text{ at } (x, s, t, u) = (0, 0, 0, 0).$$

Show that from all points (x, s, t, u) satisfying $g(x, s, t, u) = 1$, the variable x is a differentiable function of $s,\ t$ and u in a neighborhood of $(0, 0, 0)$. Determine its derivative. In particular calculate the derivative at $(0, 0, 0)$.

b. Consider the equations

$$g_1(x, y, z) = xy + yz + xz = 11 \text{ and } g_2(x, y, z) = x^2 - xyz = -5.$$

Show from all points satisfying both these equations the variables x and y are a differentiable function of z in a neighborhood of $z = 3$ and $(x, y) = (1, 2)$. Determine both derivatives. In particular calculate both derivatives at $z = 3$.

c. Reconsider item a). Can variable u also be written as a differentiable function $u = U(x, s, t)$ in a neighborhood of $(x, s, t) = (0, 0, 0)$? Motivate your answer.

ANSWER 3.14

a. The derivative of function g is

$$Dg = \left[\begin{array}{cccc} e^{x+s+t+u} - 2x & e^{x+s+t+u} - 2s & e^{x+s+t+u} - 2t & e^{x+s+t+u} - 2u \end{array} \right].$$

Since all partial derivatives are continuous, $g \in C^1$. Moreover, $g(0, 0, 0, 0) = 1$ and $D_x g(0, 0, 0, 0) = 1$. Hence, by the Implicit Function Theorem, $x = h(s, t, u)$ in a ball around $(0, 0, 0)$. Furthermore,

$$h^{'}(s, t, u) = -\frac{1}{e^{x+s+t+u} - 2x}[e^{x+s+t+u} - 2s \ \ e^{x+s+t+u} - 2t \ \ e^{x+s+t+u} - 2u].$$

In particular $h^{'}(0, 0, 0) = [1\ 1\ 1]$.

b. With $g := [g_1\ g_2]^T$, $Dg = \begin{bmatrix} y+z & x+z & x+y \\ 2x-yz & -xz & -xy \end{bmatrix}$. Since all partial derivatives are continuous $g \in C^1$. Furthermore, $g(1,2,3) = [11\ -5]^T$ and $D_{x,y}g(1,2,3) = \begin{bmatrix} 5 & 4 \\ -4 & -3 \end{bmatrix}$, which determinant differs from zero. $\rightarrow$ According the Implicit Function Theorem $x = h_1(z)$ and $y = h_2(z)$ in a ball around $z = 3$. Moreover, $\begin{bmatrix} h_1'(z) \\ h_2'(z) \end{bmatrix} = -\begin{bmatrix} y+z & x+z \\ 2x-yz & -xz \end{bmatrix}^{-1} \begin{bmatrix} x+y \\ -xy \end{bmatrix}$. In particular $\begin{bmatrix} h_1'(3) \\ h_2'(3) \end{bmatrix} = -\begin{bmatrix} 5 & 4 \\ -4 & -3 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$.

c. Yes, $D_u g(0,0,0) \neq 0$.

EXERCISE 3.15 Solve the next constrained optimization problems.

a. Optimize $f(x,y,z) = x - y + z$ subject to the constraints $g_1(x,y,z) = x^2 + 2y^2 + z^2 = 40$ and $g_2(x,y,z) = x + y + z = 0$. Moreover, calculate by how much the global maximum will, approximately, change if, $b_1 = 40$, is decreased to, $b_1 = 39$.

b. Optimize $f(x,y,z) = x^2 - y + z$ subject to the constraints $g_1(x,y,z) = x^2 + y^2 = 5$ and $g_2(x,y,z) = x^2 + z = 0$. Moreover, calculate by how much the global minimum will, approximately, change if, $b_1 = 5$, is decreased to, $b_1 = 4.9$.

c. Optimize $f(x,y,z) = x^2 + y - z$ subject to the constraints $g_1(x,y,z) = x^2 + y - z = 0$ and $g_2(x,y,z) = \frac{2}{3}x^3 + \frac{1}{2}y^2 + z = 0$.

ANSWER 3.15

a. Note that f is continuous and that the constraint set is compact (g_1 describes already a compact set) $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.34 we have:

Step 1. $Df = [1\ -1\ 1]$ and $Dg = \begin{bmatrix} 2x & 4y & 2z \\ 1 & 1 & 1 \end{bmatrix}$. Note all partial derivatives are continuous $\rightarrow$ $f, g \in C^1$.

Step 2. Dg has not full row rank iff. simultaneously $2x - 4y = 2x - 2z = 4y - 2z = 0$. This only occurs if $(x,y,z) = (0,0,0)$. But $(0,0,0)$ is not located in the constraint set.

Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x,y,z,\lambda_1,\lambda_2) := x - y + z - \lambda_1(x^2 + 2y^2 + z^2 - 40) - \lambda_2(x + y + z)$.

Step 4. $DL = [1 - 2\lambda 1x - \lambda_2\ -1 - 4\lambda_1 y - \lambda_2\ 1 - 2\lambda_1 z - \lambda_2\ g_1 - 40\ g_2] \rightarrow$ necessary conditions: $i)\ 1 - 2\lambda 1x - \lambda_2 = 0;\ ii)\ -1 - 4\lambda_1 y - \lambda_2 = 0;\ iii)\ 1 - 2\lambda_1 z - \lambda_2 = 0;\ iv)\ g_1 = 40;\ v)\ g_2 = 0$. From i) and ii) we have that either $\lambda_1 = 0$ or $x = z$. In case $\lambda_1 = 0$, ii) and iii) yield a contradiction. If $x = z$, by v), $y = -2x$. Substitution of this into i) and ii) gives: $\lambda_2 = 1 - 2\lambda_1 x = -1 + 8\lambda_1 x \rightarrow x = z = \frac{1}{5\lambda_1}$ and $y = \frac{-2}{5\lambda_1}$. Substitution of this into iv) yields: $\lambda_1 = \pm\frac{1}{10}$. $\rightarrow$ Stationary points: $(2,-4,2)$ with $\lambda_1 = \frac{1}{10}$, $\lambda_2 = \frac{3}{5}$; and $(-2,4,-2)$ with $\lambda_1 = \frac{-1}{10}$, $\lambda_2 = \frac{3}{5}$.

$f(2,-4,2) > f(-2,4,-2) \rightarrow f$ has a global maximum at $(2,-4,2)$ and a global minimum at $(-2,4,-2)$.

If $b_1 = 40$ is reduced by 1 unit, the maximum 8 decreases approximately by $\lambda_1 = \frac{1}{10}$ to $7\frac{9}{10}$.

b. Note that f is continuous and that the constraint set is compact (By g_1, x and y are bounded, from g_2 it follows next that z must be bounded too) $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.34 we have:

Step 1. $Df = [2x \ -1 \ 1]$ and $Dg = \begin{bmatrix} 2x & 2y & 0 \\ 2x & 0 & 1 \end{bmatrix}$. Note all partial derivatives are continuous $\rightarrow$ $f, g \in C^1$.

Step 2. Dg has not full row rank iff. simultaneously $-4xy = 2x = 2y = 0$. This only occurs if $(x,y,z) = (0,0,0)$. But $(0,0,0)$ is not located in the constraint set.

Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x,y,z,\lambda_1,\lambda_2) := x^2 - y + z - \lambda_1(x^2+y^2-5) - \lambda_2(x^2+z)$.

Step 4. $DL = [2x - 2\lambda 1x - 2\lambda_2 x \ -1 - 2\lambda_1 y \ 1 - \lambda_2 \ g_1 - 5 \ g_2] \rightarrow$ necessary conditions: $i)\ 2x(1 - \lambda 1 - \lambda_2) = 0;\ ii)\ -1 - 2\lambda_1 y = 0;\ iii)\ 1 - \lambda_2 = 0;\ iv)\ g_1 = 5;\ v)\ g_2 = 0$.

From iii) we have that $\lambda_2 = 1$. From i) we have that either $\lambda_1 = 0$ or $x = 0$. In case $\lambda_1 = 0$, ii) gives a contadiction. In case $x = 0$, by v), $z = 0$ and by iv), $y = \pm\sqrt{5} \rightarrow$ Stationary points: $(0,\sqrt{5},0)$ with $\lambda_1 = \frac{-1}{2\sqrt{5}}$, $\lambda_2 = 1$;and $(0,-\sqrt{5},0)$ with $\lambda_1 = \frac{1}{2\sqrt{5}}$, $\lambda_2 = 1$.

$f(0,\sqrt{5},0) = -\sqrt{5}$ and $f(0,-\sqrt{5},0) = \sqrt{5} \rightarrow f$ has a global maximum at $(0,-\sqrt{5},0)$ and a global minimum at $(0,\sqrt{5},0)$.

If $b_1 = 40$ is reduced by 0.1 unit, the maximum decreases approximately by $\lambda_1 * 0.1 = \frac{1}{20\sqrt{5}}$.

c. Following Algorithm 3.34 we have:

Step 1. $Df = [2x \ 1 \ -1]$ and $Dg = \begin{bmatrix} 2x & 1 & -1 \\ 2x^2 & y & 1 \end{bmatrix}$. Note all partial derivatives are continuous $\rightarrow f, g \in C^1$.

Step 2. Dg has not full row rank iff. simultaneously $2xy - 2x^2 = 2x + 2x^2 = 1 + y = 0$. This occurs if $(x,y,z) = (0,-1,z)$ for arbitrary z. But substitution into the constraints of this shows that there exists no point $(0,-1,z)$ satisfying both of them simultaneously.

Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x,y,z,\lambda_1,\lambda_2) := x^2 - y + z - \lambda_1(x^2 - y + z) - \lambda_2(\frac{2}{3}x^3 + \frac{1}{2}y^2 + z)$.

Step 4. $DL = [2x - 2\lambda 1x - 2\lambda_2 x^2 \ 1 - \lambda_1 - \lambda_2 y \ -1 + \lambda_1 - \lambda_2 \ g_1 \ g_2] \rightarrow$ necessary conditions: $i)\ 2x(1 - \lambda 1 - \lambda_2 x) = 0;\ ii)\ 1 - \lambda_1 - \lambda_2 y = 0;\ iii)\ -1 + \lambda_1 - \lambda_2 = 0;\ iv)\ g_1 = 0;\ v)\ g_2 = 0$.

From i) we have that either $1 - \lambda 1 - \lambda_2 x = 0$ or $x = 0$. In case $x = 0$, by iv) and v), $y = z$ and either $z = 0$ or $z = -2$. From ii) and iii) it follows that $\lambda_1 = 1$ and $\lambda_2 = 0$. $\rightarrow$ Stationary points: $(0,0,0)$ with $\lambda_1 = 1$, $\lambda_2 = 0$;and $(0,-2,-2)$ with $\lambda_1 = 1$, $\lambda_2 = 0$.

In case $1 - \lambda 1 - \lambda_2 x = 0$ it follows from ii) that either $x = y$ or, $\lambda_2 = 0$. If $\lambda_2 = 0$, by iii), $\lambda_1 = 1$. $\rightarrow$ all points satisfying iv) and v) with $\lambda_1 = 1$ and $\lambda_2 = 0$ are stationary points.

In case $x = y$, iv) gives $x = 0$ and we obtain the same stationary points as before.

Clearly, all points satisfying the constraints are optima. They all give rise to the same function value: 0.

Mixed Exercises

EXERCISE 3.16 Determine all optima (location, value, kind (local/global)) of the next constrained optimization problems.

a. Optimize $f(x,y) = xy^2$, subject to the constraint $g(x,y) = x + y = 1$.
b. Optimize $f(x,y) = (1-y)e^x$, subject to the constraint $g(x,y) = x^3 + y = 1$.
c. Optimize $f(x,y,z) = \frac{1}{2}x^2y + z$ subject to the constraints $g_1(x,y,z) = \frac{1}{2}x^2 + 1 - y = 0$ and $g_2(x,y,z) = \frac{4}{3}x - \frac{4}{3}xy - z = 0$.
d. Optimize $f(x,y,z) = x^2 + y^2 + z^2$, subject to the constraints $g_1(x,y,z) = x + y + z = 0$ and $g_2(x,y,z) = x - y + z = 2$.
e. Optimize $f(x,y) = xy^2$, subject to the constraint $g(x,y) = x + y^2 = 1$.
f. Optimize $f(x,y,z) = e^{xyz}$, subject to the constraint $g(x,y,z) = xy - z = 0$.
g. Optimize $f(s,t,u,v) = s(u+v) + t(u-v)$, subject to the constraints $g_1(s,t,u,v) = s - u - v = 0$ and $g_2(s,t,u,v) = s + t - 2u = 0$.
h. Optimize $f(s,t,u,v) = (s-t)(u-v)$, subject to the constraint $g_1(s,t,u,v) = s - t - u + v = 0$.

ANSWER 3.16

a. $y = 1 - x \rightarrow F(x) = x(1-x)^2 \rightarrow F^{'}(x) = (1-x)^2 - 2x(1-x) = (1-x)[1-3x] \rightarrow$ Stationary points: $x = 1$ and $x = \frac{1}{3}$.

```
   +           -           +     F'(x)
-------0---------------0-------
      1/3             1        x
```

$\rightarrow F$ has a local maximum at $x = \frac{1}{3}$ and a local minimum at $x = 1$. Not global, as $F(100) > F(\frac{1}{3})$ and $F(-100) < F(1)$.

$\rightarrow$ Problem has a local maximum at $(\frac{1}{3}, \frac{2}{3})$ and a local minimum at $(1,0)$.

b. $y = 1 - x^3 \rightarrow F(x) = x^3e^x \rightarrow F^{'}(x) = (3x^2 + x^3)e^x \rightarrow$ Stationary points: $x^2e^x(3+x) = 0 \rightarrow x = 0$ and $x = -3$.

```
   -           +           +     F'(x)
-------0---------------0-------
      -3              0        x
```

$\rightarrow F$ has a global minimum at $x = -3$ and a saddle point at $x = 0$.

$\rightarrow$ Problem has a global minimum at $(-3, 28)$ and a saddle point at $(0,1)$.

c. $y = 1 + \frac{1}{2}x^2$ and $z = \frac{4}{3}x - \frac{4}{3}x(1 + \frac{1}{2}x^2) \rightarrow F(x) = \frac{1}{2}x^2y + z = \frac{1}{4}x^4 - \frac{2}{3}x^3 + \frac{1}{2}x^2 \rightarrow F^{'}(x) = x^3 - 2x^2 + x = x(x^2 - 2x + 1) = x(x-1)^2 \rightarrow$ Stationary points: $\rightarrow x = 0$ and $x = 1$.

```
   -           +           +      F'(x)
-------0---------------0-------
       0               1        x
```

$\rightarrow F$ has a global minimum at $x = 0$ and a saddle point at $x = 1$.

$\rightarrow$ Problem has a global minimum at $(0, 1, 0)$ and a saddle point at $(1, \frac{3}{2}, -\frac{2}{3})$.

d. $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} -x \\ 2-x \end{bmatrix}$

$\rightarrow \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -x \\ 2-x \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -x \\ 2-x \end{bmatrix} = \begin{bmatrix} -1 \\ 1-x \end{bmatrix}$.

$\rightarrow y = 1$ and $z = 1 - x$

$\rightarrow F(x) = x^2 + y^2 + z^2 = x^2 + 1 + (x-1)^2 \rightarrow F^{'}(x) = 2x + 2(x-1) \rightarrow$ Stationary points: $\rightarrow x = \frac{1}{2}$.

```
   -                   +      F'(x)
---------------0---------------
               1/2              x
```

$\rightarrow F$ has a global minimum at $x = \frac{1}{2}$.

$\rightarrow$ Problem has a global minimum at $(\frac{1}{2}, 1, \frac{1}{2})$.

e. $x = 1 - y^2 \rightarrow F(y) = (1-y^2)y^2 = y^2 - y^4 \rightarrow F^{'}(y) = 2y - 4y^3 = 2y(1-2y^2)$

$\rightarrow$ Stationary points: $y = 0$ and $y = \pm\sqrt{\frac{1}{2}}$.

```
   +           -           +           -      F'(x)
-------0---------------0-----------0---------
  -(1/2)^(0.5)         0        (1/2)^(0.5)     x
```

$\rightarrow F$ has a local maximum at $y = \pm\sqrt{\frac{1}{2}}$ and a local minimum at $y = 0$.

$F(100) < F(0) \rightarrow F$ has not a global minimimum at $y = 0$. $F(\sqrt{\frac{1}{2}}) = F(-\sqrt{\frac{1}{2}}) \rightarrow$ (see sign scheme derivative) F has a global maximum at these points.

f. $z = xy \rightarrow F(x, y) = e^{x^2y^2} \rightarrow F^{'}(x, y) = e^{x^2y^2} \begin{bmatrix} 2xy^2 & 2yx^2 \end{bmatrix}$

$\rightarrow$ Stationary points: $(x, 0)$ and $(0, y)$.

$HF = \begin{bmatrix} 2y^2 & 4xy \\ 4xy & 2x^2 \end{bmatrix}$. $\rightarrow HF(x, 0)$ and $HF(0, y)$ are positive semi-definite $\rightarrow$ No conclusions possible.

However, clearly $F(x, y) \geq 1$ $(x^2y^2 \geq 0)$, $F(0, x) = F(y, 0) = 1 \rightarrow F$ has global minima at the stationary points $\rightarrow$ Problem has a global minimum at all points $(0, y, 0)$ and $(x, 0, 0)$.

g. From g_1: $s = u + v$; from g_2: $t = 2u - s = u - v \rightarrow F(u, v) = (u+v)^2 + (u-v)^2$

$\rightarrow F^{'}(u, v) = [2(u+v) + 2(u-v) \;\; 2(u+v) - 2(u-v)] = [4u \;\; 4v]$

$\rightarrow$ Stationary point: $(0, 0)$.

$HF = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$. $\rightarrow HF(0, 0)$ is positive definite $\rightarrow F$ has a local minimum at $(0, 0)$.

Note $HF(x, y)$ is positive definite $\rightarrow F$ has a global minimum at the stationary point $\rightarrow$ Problem has a global minimum at $(0, 0, 0, 0)$.

h. From g_1: $s = t + u - v \rightarrow F(t,u,v) = (u-v)^2 \rightarrow F^{'}(t,u,v) = [0 \;\; 2(u-v) \;\; -2(u-v)]$
$\rightarrow$ Stationary points: (t,u,u).
$HF = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & 2 \end{bmatrix}$. $\rightarrow HF(t,u,v)$ is positive semi-definite $\rightarrow F$ is convex $\rightarrow F$ has a global minimum at the stationary points $\rightarrow$ Problem has a global minimum at all points (t,t,u,u).

EXERCISE 3.17 Determine all optima (location, value, kind (local/global)) of the next constrained optimization problems.

a. Optimize $f(x,y) = xy^2$, subject to the constraint $g(x,y) = x^2 + y^2 = 12$.
b. Optimize $f(x,y) = e^x y^2$, subject to the constraint $g(x,y) = x + y^2 = 1$.
c. Optimize $f(x,y) = x^2 y$, subject to the constraint $g(x,y) = x^2 - y^2 = 1$.
d. Optimize $f(x,y,z) = x^2 z$ subject to the constraints $g_1(x,y,z) = x^2 + y^2 + z = 1$ and $g_2(x,y,z) = x^2 + y = 1$.
e. Optimize $f(x,y,z) = x^2 + y^2 + z^2$, subject to the constraints $g_1(x,y,z) = x^2 + y^2 = 4$ and $g_2(x,y,z) = y^2 + z^2 = 2$.
f. Optimize $f(x,y,z) = x^2 + \frac{4}{3}y^3 + 16y - 8z^2$, subject to the constraints $g_1(x,y,z) = x^2 + y^4 = 16$ and $g_2(x,y,z) = (y-2)^2 + z^2 = 9$.
g. Optimize $f(x,y) = e^{xy} z - 3(1-x-y)$, subject to the constraint $g_1(x,y,z) = e^{xy} - z = -1$ and $g_2(x,y,z) = x + y + z = 1$.
h. Optimize $f(x,y,z) = e^{xyz}$, subject to the constraint $g_1(x,y,z) = xy - z = 0$ and $g_2(x,y,z) = x^2 + z^2 = 1$.

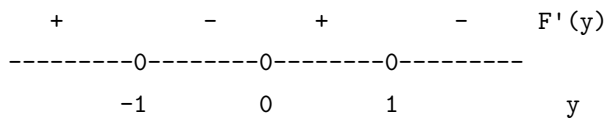
ANSWER 3.17

a. $y^2 = 12 - x^2 \geq 0 \rightarrow F(x) = x(12 - x^2)$ where $x \in [-\sqrt{12}, \sqrt{12}]$
$\rightarrow F$ is continuous on a compact domain $\rightarrow$ (Weierstrass) F has both a global maximum and a global minimum.
$F^{'}(x) = 12 - 3x^2 \rightarrow$ Stationary points: $x = 2$ and $x = -2$.

```
        -                +                -          F'(x)
   |----------0----------------0-----------|
-12^(0.5)     -2                2         12^(0.5)   x
```

$\rightarrow F$ has a local minimum at $x = -2$ and $x = \sqrt{12}$; a local maximum at $x = -\sqrt{12}$ and $x = 2$.
$F(-2) = -16 < 0 = F(\sqrt{12}) \rightarrow F$ has a global minimum at $x = -2$ and a local minimum (which is not global) at $x = \sqrt{12}$.
$F(2) = 16 > 0 = F(\sqrt{12}) \rightarrow F$ has a global maximum at $x = 2$ and a local maximum (which is not global) at $x = -\sqrt{12}$.
$\rightarrow$ Problem has a global maximum at $(2, \sqrt{8})$ and $(2, -\sqrt{8})$; a local maximum (which is not global) at $(-\sqrt{12}, 0)$; a global minimum at $(-2, \sqrt{8})$ and $(-2, -\sqrt{8})$ and a local minimum (which is not global) at $(\sqrt{12}, 0)$.

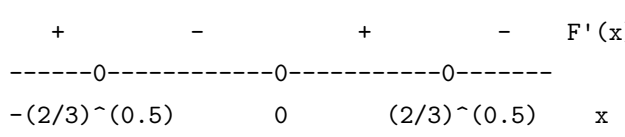
b. $x = 1 - y^2 \geq 0 \rightarrow F(x) = e^{1-y^2}y^2 \rightarrow F^{'}(y) = 2ye^{1-y^2} - 2y^3e^{1-y^2} = 2y(1-y^2)e^{1-y^2}$
$\rightarrow$ Stationary points: $2y(1-y^2) = 0 \rightarrow y = 0$; $y = 1$ and $y = -1$.



$\rightarrow F$ has a local maximum at $y = -1$ and $y = 1$; a local minimum at $y = 0$.
$F(-1) = F(1) \rightarrow F$ has a global maximum at $y = \pm 1$; $F(0) = 0 \leq F(y) \rightarrow$ F has a global minimum at $y = 0$
$\rightarrow$ Problem has a global maximum at $(0, -1)$ and $(0, 1)$; a global minimum at $(1, 0)$.

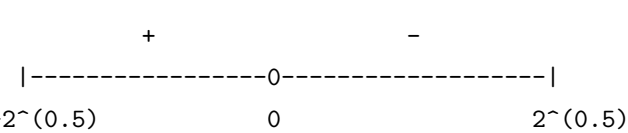
c. $x^2 = 1 + y^2 \geq 0 \rightarrow F(y) = (1 + y^2)y$ where $y \in \mathbb{R}$ ($y^2 \geq -1$ always holds)
$F^{'}(y) = 1 + 3y^2 \rightarrow$ No extremal points.

d. $g_2 : y = 1 - x^2 \rightarrow g_1 : z = 1 - x^2 - y^2 = 1 - x^2 - (1 - x^2)^2 \rightarrow F(x) = x^2(1 - x^2 - (1 - x^2)^2) = x^4 - x^6$.
$\rightarrow F^{'}(x) = 4x^3 - 6x^5 = 2x^3(2 - 3x^2) \rightarrow$ Stationary points: $x = 0$ and $x = \pm\sqrt{\frac{2}{3}}$.



$\rightarrow F$ has a local minimum at $x = 0$; a local maximum at $x = \pm\sqrt{\frac{2}{3}}$.
$F(0) > F(100) \rightarrow F$ has a local minimum (which is not global) at $x = 0$.
$F(-\sqrt{\frac{2}{3}}) = F(\sqrt{\frac{2}{3}}) \rightarrow F$ has a global maximum at $x = \pm\sqrt{\frac{2}{3}}$.
$\rightarrow$ Problem has a local minimum (which is not global) at $(0, 1, 0)$ and a global maximum at $(\pm\sqrt{\frac{2}{3}}, \frac{1}{3}, \frac{2}{9})$.

e. $g_1 : x^2 = 4 - y^2 \geq 0 \rightarrow y^2 \leq 4$; $g_2 : z^2 = 2 - y^2 \geq 0 \rightarrow y^2 \leq 2$
$\rightarrow F(y) = 4 - y^2 + y^2 + 2 - y^2 = 6 - y^2$ where $y \in [-\sqrt{2}, \sqrt{2}]$
$\rightarrow F$ is continuous on a compact domain $\rightarrow$ (Weierstrass) F has both a global maximum and a global minimum.
$F^{'}(y) = -2y \rightarrow$ Stationary point: $y = 0$.

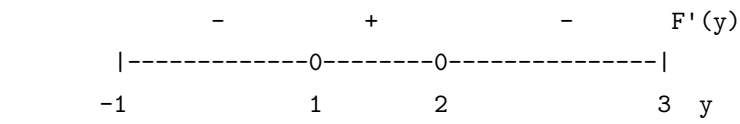


$\rightarrow F$ has a local minimum at $y = \pm\sqrt{2}$; a global maximum at $y = 0$.
$F(-\sqrt{2}) = F(\sqrt{2}) \rightarrow F$ has a global minimum at $y = \pm\sqrt{2}$.
$\rightarrow$ Problem has a global maximum at $(\pm 2, 0, \pm\sqrt{2})$ and a global minimum at $(\pm\sqrt{2}, \pm\sqrt{2}, 0)$.

f. $g_1 : x^2 = 81 - y^4 \geq 0 \rightarrow y^4 \leq 81 \rightarrow y \in [-3, 3]$; $g_2 : z^2 = 9 - (y - 2)^2 \geq 0 \rightarrow (y - 2)^2 \leq 9$
$\rightarrow y - 2 \in [-3, 3] \rightarrow y \in [-1, 5]$.
$\rightarrow F(y) = 81 - y^4 + \frac{4}{3}y^3 + 16y - 72 + 8(y - 2)^2 = -y^4 + \frac{4}{3}y^3 + 8y^2 - 16y + 41$ where $y \in [-1, 3] \rightarrow$ F is continuous on a compact domain $\rightarrow$ (Weierstrass) F has both a global maximum and a global minimum.

$F^{'}(y) = -4y^3 + 4y^2 + 16y - 16 = -4(y-1)(y+2)(y-2)$

$\rightarrow$ Stationary points: $y = 1$, $y = -2$ and $y = 2$.



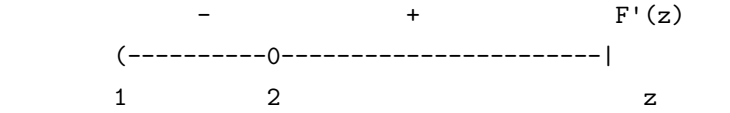
$\rightarrow F$ has a local maximum at $y = -1$ and $y = 2$; a local minimum at $y = 1$ and $y = 3$.

$F(-1) = 62\frac{2}{3} > 35\frac{2}{3} = F(2) \rightarrow F$ has a global maximum at $y = 2$.

$F(1) = 33\frac{1}{3} > 20 = F(3) \rightarrow F$ has a global minimum at $y = 3$. $\rightarrow$ Problem has a global maximum at $(\pm\sqrt{65}, 2, \pm 3)$; a local maximum (which is not global) at $(\pm\sqrt{80}, -1, \pm\sqrt{8})$; a global minimum at $(0, 3, \pm\sqrt{8})$ and a local minimum at $(\pm\sqrt{80}, 1, \pm\sqrt{8})$ (which is not global).

g. $g_1 : e^{xy} = -1 + z > 0$, $g_2 : x + y - 1 = z \rightarrow F(z) = (z-1)z - 3z$, where $z > 1$.

$F^{'}(z) = 2z - 4 \rightarrow$ Stationary points: $z = 2$.



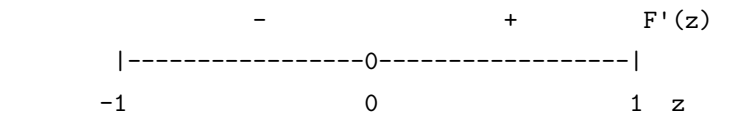
$\rightarrow F$ has a global minimum at $z = 2$.

$\rightarrow$ Problem has a global minimum at $(0, -1, 2)$ and $(-1, 0, 2)$.

h. $g_1 : z = xy$, $g_2 : z \in [-1, 1] \rightarrow F(z) = e^{z^2}$, where $z \in [-1, 1]$.

$\rightarrow F$ is continuous on a compact domain $\rightarrow$ (Weierstrass) F has both a global maximum and a global minimum.

$F^{'}(z) = 2ze^{z^2} \rightarrow$ Stationary point: $z = 0$.



$\rightarrow F$ has a global minimum at $z = 0$; a local maximum at $z = -1$ and $z = 1$.

$F(-1) = F(1) \rightarrow F$ has a global maximum at $z = \pm 1$

$\rightarrow$ Problem has a global minimum at $(\pm 1, 0, 0)$ and a global maximum at $(1, -1, 1)$, $(-1, 1, 1)$, $(1, 1, 1)$ and $(-1, -1, 1)$.

EXERCISE 3.18 Use Lagrange's method to determine all points that qualify for an optimum of the next constrained optimization problems. Use, next, second-order criteria; Weierstrass' theorem; or convexity of the Lagrange function (see Exercise 3.24, below) to determine the optima (location, value, kind (local/global)) of the problems.

a. Optimize $f(x, y) = xy$, subject to the constraint $g(x, y) = x^2 + y^2 = 2$.

b. Optimize $f(x, y) = x^2 y$, subject to the constraint $g(x, y) = x^4 + y^4 = 64$.

c. Optimize $f(x, y) = 4x + 2y$, subject to the constraint $g(x, y) = x^{\frac{2}{3}} y^{\frac{1}{3}} = 10,\ x > 0,\ y > 0$.

d. Optimize $f(x, y) = x^2 + y^2$, subject to the constraint $g(x, y) = x^2 + xy + y^2 = 3$.

e. Optimize $f(x,y) = x^2 + (y-1)^2$, subject to the constraint $g(x,y) = 4x^2 + y^2 = 16$.
f. Optimize $f(x,y) = x^2 + \frac{3}{2}\ln(2+y^2)$, subject to the constraints $g(x,y) = 2x + y^2 = 2$.
g. Optimize $f(x,y) = e^{xy}$, subject to the constraint $g(x,y) = x^2 + y^2 = 1$.
h. Optimize $f(x,y) = xye^{\frac{1}{2}x^2+y^2}$, subject to the constraint $g(x,y) = x^2 + y^2 = 4$.

ANSWER 3.18

a. Note that f is continuous and that the constraint set is compact $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.13 we have:
Step 1. $Df = [y \;\; x]$ and $Dg = [2x \;\; 2y]$. Note all partial derivatives are continuous $\rightarrow f, g \in C^1$.
Step 2. Dg has not full row rank iff. simultaneously $2x = 2y = 0$. This only occurs if $(x,y) = (0,0)$. But $(0,0)$ is not located in the constraint set.
Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x,y,\lambda) := xy - \lambda(x^2 + y^2 - 2)$.
Step 4. $DL = [y - 2\lambda x \;\; x - 2\lambda y \;\; g - 2] \rightarrow$ necessary conditions: $i)\; y - 2\lambda x = 0;\; ii)\; x - 2\lambda y = 0;\; iii)\; g = 2$.
Substitution from i) into ii) shows $x(1 - 4\lambda^2) = 0 \rightarrow x = 0$ or $\lambda = \pm\frac{1}{2}$.
In case $x = 0$, i) shows $y = 0$, which violates iii).
If $\lambda = \frac{1}{2}$, i) shows $y = x$. Substitution into iii) gives then $x^2 = 1 \rightarrow$ Candidate extremal points: $(1,1)$ and $(-1,-1)$ with $\lambda = \frac{1}{2}$.
If $\lambda = -\frac{1}{2}$, i) shows $y = -x$. Substitution into iii) gives then $x^2 = 1 \rightarrow$ Candidate extremal points: $(1,-1)$ and $(-1,1)$ with $\lambda = -\frac{1}{2}$.
Note: $f(1,1) = f(-1,-1) = 1$; $f(-1,1) = f(1,-1) = -1 \rightarrow$ Problem has a global maximum at $(1,1)$ and $(-1,-1)$; and a global minimum at $(-1,1)$ and $(1,-1)$.

b. Note that f is continuous and that the constraint set is compact $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.13 we have:
Step 1. $Df = [2xy \;\; x^2]$ and $Dg = [4x^3 \;\; 4y^3]$. Note all partial derivatives are continuous $\rightarrow f, g \in C^1$.
Step 2. Dg has not full row rank iff. simultaneously $4x^3 = 4y^3 = 0$. This only occurs if $(x,y) = (0,0)$. But $(0,0)$ is not located in the constraint set.
Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x,y,\lambda) := x^2y - \lambda(x^4 + y^4 - 64)$.
Step 4. $DL = [2xy - 4\lambda x^3 \;\; x^2 - 4\lambda y^3 \;\; g - 64] \rightarrow$ necessary conditions: $i)\; 2xy - 4\lambda x^3 = 0;\; ii)\; x^2 - 4\lambda y^3 = 0;\; iii)\; g = 64$.
From i): $2x(y - 2\lambda x^2) = 0 \rightarrow x = 0$ or $y = 2\lambda x^2$.
1. Case $x = 0$: Substitution into ii) shows $y = 0$ or $\lambda = 0 \rightarrow (0,0)$ violates iii). Candidates: $(0, \pm(64)^{1/4})$ with $\lambda = 0$.
2. Case $y = 2\lambda x^2$: Substitution into ii) shows $x^2(1 - 32\lambda^4 x^4) = 0 \rightarrow x = 0$ or $1 - 32\lambda^4 x^4 = 0$. Case $x = 0$: $\rightarrow y = 0$, but $(0,0)$ violates iii).
3. Case $1 - 32\lambda^4 x^4 = 0$: $\lambda x = \pm\frac{1}{2} \rightarrow x = \pm\frac{1}{2\lambda}$. Substitution into $y = 2\lambda x^2$ shows $y = \frac{1}{2\lambda}$. Substitution of $x = \pm\frac{1}{2\lambda}$ and $y = \frac{1}{2\lambda}$ into g shows: $(\frac{1}{2\lambda})^4 = 32 \rightarrow \frac{1}{2\lambda} = \pm 2 \rightarrow \lambda = \pm\frac{1}{4}$.

$\rightarrow$ Candidates: $(\pm 2, 2)$ with $\lambda = \frac{1}{4}$ and $(\pm 2, -2)$ with $\lambda = -\frac{1}{4}$.

Evaluating f at the candidate solutions shows that the problem has a global maximum at $(\pm 2, 2)$; and a global minimum at $(\pm 2, -2)$.

Furthermore, $f(0, \pm(64)^{1/4}) = 0$. Notice $f(\epsilon, (64-\epsilon)^{1/4}) > 0$ for ϵ close to zero $\rightarrow$ f has a local minimum at $(0, (64)^{1/4})$; and $f(\epsilon, -(64+\epsilon)^{1/4}) < 0$ if ϵ is close to zero $\rightarrow$ f has a local maximum at $(0, -(64)^{1/4})$.

c. Note that f and g are continuous on the domain. Following Algorithm 3.13 we have:

Step 1. $Df = [4 \;\; 2]$ and $Dg = [\frac{2}{3}x^{-\frac{1}{3}}y^{\frac{1}{3}} \;\; \frac{1}{3}x^{\frac{2}{3}}y^{-\frac{2}{3}}]$. Note all partial derivatives are continuous $\rightarrow$ $f, g \in C^1$.

Step 2. Dg has not full row rank iff. simultaneously $x = y = 0$. But $(0,0)$ is not located in the constraint set.

Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x, y, \lambda) := 4x + 2y - \lambda(x^{\frac{2}{3}}y^{\frac{1}{3}} - 10)$.

Step 4. $DL = [4 - \frac{2}{3}\lambda x^{-\frac{1}{3}}y^{\frac{1}{3}} \;\; 2 - \frac{1}{3}\Lambda x^{\frac{2}{3}}y^{-\frac{2}{3}} \;\; g - 10] \rightarrow$ necessary conditions: $i)\; 4 - \frac{2}{3}\lambda x^{-\frac{1}{3}}y^{\frac{1}{3}} = 0;\; ii)\; 2 - \frac{1}{3}x^{\frac{2}{3}}y^{-\frac{2}{3}} = 0;\; iii)\; g = 10$.

Rewriting i) and ii) and next dividing both resulting equations shows $\frac{2y}{x} = \frac{4}{2} \rightarrow x = y$. Substitution into iii) gives then $x = 10 \rightarrow$ Candidate extremal point: $(10, 10)$ with $\lambda_0 := \frac{3}{2}10^{\frac{1}{3}}$.

Following Exercise 23, consider $L(x, y) = 4x + 2y - \lambda_0(x^{\frac{2}{3}}y^{\frac{1}{3}} - 10) \rightarrow DL = [4 - \frac{2}{3}\lambda_0 x^{-\frac{1}{3}}y^{\frac{1}{3}} \;\; 2 - \frac{1}{3}\lambda_0 x^{\frac{2}{3}}y^{-\frac{2}{3}} \;\; g - 10] \rightarrow HL(x, y) = \begin{bmatrix} \frac{2}{9}\lambda_0 x^{-\frac{4}{3}}y^{\frac{1}{3}} & -\frac{2}{9}\lambda_0 x^{-\frac{1}{3}}y^{-\frac{2}{3}} \\ -\frac{2}{9}\lambda_0 x^{-\frac{1}{3}}y^{-\frac{2}{3}} & \frac{2}{9}\lambda_0 x^{\frac{2}{3}}y^{-\frac{5}{3}} \end{bmatrix} \rightarrow HL(x, y)$ is positive semi-definite $\rightarrow L(x, y, \lambda_0)$ is convex $\rightarrow$ Problem has a global minimum at $(10, 10)$.

d. Note that f is continuous and that the constraint set can be rewritten as $(x + \frac{1}{2}y)^2 + \frac{3}{4}y^2 = 3$. So the constraint set is compact $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.13 we have:

Step 1. $Df = [2x \;\; 2y]$ and $Dg = [2x + y \;\; x + 2y]$. Note all partial derivatives are continuous $\rightarrow$ $f, g \in C^1$.

Step 2. Dg has not full row rank iff. simultaneously $2x + y = x + 2y = 0$. This only occurs if $(x, y) = (0, 0)$. But $(0,0)$ is not located in the constraint set.

Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x, y, \lambda) := x^2 + y^2 - \lambda(x^2 + xy + y^2 - 3)$.

Step 4. $DL = [2x - 2\lambda x - \lambda y \; 2y - \lambda x - 2\lambda y \; g - 3] \rightarrow$ necessary conditions: $i)\; 2x - 2\lambda x - \lambda y = 0;\; ii)\; 2y - \lambda x - 2\lambda y] = 0;\; iii)\; g = 3$.

From i): $x = \frac{\lambda}{2(1-\lambda)}$. Note that $\lambda = 1$ yields $x = y = 0$, which does not satisfy the constraint. Substitution of $x = \frac{\lambda}{2(1-\lambda)}$ into ii) shows $2y(1 - \lambda) = \frac{\lambda^2}{2(1-\lambda)}y \rightarrow y = 0$ (with $x = 0$ again, violating the constraint) or $4(1-\lambda)^2 = \lambda^2$. From the last equality we have either $\lambda = 2$ or $\lambda = \frac{2}{3}$. With $\lambda = 2$, corresponds $y = -x$. Substitution into the constraint yields $(x, y) = (\sqrt{3}, -\sqrt{3})$ or $(x, y) = (-\sqrt{3}, \sqrt{3})$.

With $\lambda = \frac{2}{3}$ corresponds $y = x$. Substitution into the constraint yields $(x, y) = (1, 1)$ or $(x, y) = (-1, -1)$.

Evaluating f at the candidate solutions shows that the problem has a global maximum at $(\sqrt{3}, -\sqrt{3})$ and $(x, y) = (-\sqrt{3}, \sqrt{3})$; and a global minimum at $(1, 1)$ and $(-1, -1)$.

e. Note that f is continuous and that the constraint set is compact $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.13 we have:

Step 1. $Df = [2x \;\; 2(y-1)]$ and $Dg = [8x \;\; 2y]$. Note all partial derivatives are continuous $\rightarrow$ $f, g \in C^1$.

Step 2. Dg has not full row rank iff. simultaneously $2x = 2(y-1) = 0$. This only occurs if $(x, y) = (0, 1)$. But $(0, 1)$ is not located in the constraint set.

Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x, y, \lambda) := x^2 + (y-1)^2 - \lambda(4x^2 + y^2 - 16)$.

Step 4. $DL = [2x - 4\lambda x \; 2(y-1) - 2\lambda y \; g - 16] \rightarrow$ necessary conditions: $i)\; 2x(1 - 2\lambda) = 0;\; ii)\; 2(y-1) - 2\lambda y = 0;\; iii)\; g = 16$.

From i): $x = 0$ or $\lambda = \frac{1}{2} \rightarrow (x, y) = (0, 4)$, $(x, y) = (0, -4)$ or, with $\lambda = \frac{1}{2}$, $(x, y) = (\sqrt{3}, 2)$ and $(x, y) = (-\sqrt{3}, 2)$.

Evaluating f at the candidate solutions shows that the problem has a global maximum at $(0, 4)$ and $(0, -4)$; and a global minimum at $(\sqrt{3}, 2)$ and $(-\sqrt{3}, 2)$.

f. Note that f and g are continuous on the domain. Following Algorithm 3.13 we have:

Step 1. $Df = [2x \;\; \frac{3y}{2+y^2}]$ and $Dg = [2 \; 2y]$. Note all partial derivatives are continuous $\rightarrow$ $f, g \in C^1$.

Step 2. Dg has always full row rank. So, no additional points qualify for being an optimum.

Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x, y, \lambda) := x^2 + \frac{3}{2}\ln(2 + y^2) - \lambda(2x + y^2 - 2)$.

Step 4. $DL = [2x - 2\lambda \; \frac{3y}{2+y^2} - 2y\lambda \; g - 2] \rightarrow$ necessary conditions: $i)\; 2x - 2\lambda = 0;\; ii)\; \frac{3y}{2+y^2} - 2y\lambda = 0;\; iii)\; g = 2$.

From i): $x = \lambda$; from ii): $y(\frac{1.5}{2+y^2} - \lambda) = 0 \rightarrow y = 0$ or $\lambda = \frac{1.5}{2+y^2}$. With $y = 0$ we obtain from the constraint $x = 1$. So, one candidate point is $(x, y, \lambda) = (1, 0, 1)$. Substitution of $x = \frac{1.5}{2+y^2}$ into the constraint yields: $\frac{3}{2+y^2} + y^2 = 2$. So, $(x, y, \lambda) = (\frac{1}{2}, 1, \frac{1}{2})$ or $(x, y, \lambda) = (\frac{1}{2}, -1, \frac{1}{2})$.

Next we calculate the determinant of the bordered Hessian in all three points. Notice:

$$HL(x, y, \lambda) = \begin{bmatrix} 2 & 0 & -2 \\ 0 & \frac{6(1-y^2)}{(2+y^2)^2} - 2\lambda & -2y \\ -2 & -2y & 0 \end{bmatrix}.$$

So, at $(\frac{1}{2}, 1, \frac{1}{2})$ $\det HL = -\frac{16}{3} < 0$ and at $(\frac{1}{2}, -1, \frac{1}{2})$ $\det HL = -\frac{16}{3} < 0$. So, at both points f has a local minimum. Whereas at $(1, 0, 1)$ $\det HL = 2 > 0$. So, here f has a local maximum. Note that at, e.g., $(x, y) = (-49, 10)$, $f(x, y) > f(1, 0)$. So, f has not a global maximum. Without going into further detail we notice that the minima are global (this can be shown, e.g., by substitution of the constraint $y^2 = 2 - 2x$, with $x \leq 1$ into f).

g. Note that f is continuous and that the constraint set is compact $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.13 we have:

Step 1. $Df = [ye^{xy} \;\; xe^{xy}]$ and $Dg = [2x \;\; 2y]$. Note all partial derivatives are continuous $\rightarrow$ $f, g \in C^1$.

Step 2. Dg has not full row rank iff. simultaneously $2x = 2y = 0$. This only occurs if $(x, y) = (0, 0)$. But $(0, 0)$ is not located in the constraint set.

Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x, y, \lambda) := e^{xy} - \lambda(x^2 + y^2 - 1)$.

Step 4. $DL = [ye^{xy} - 2\lambda x \; xe^{xy} - 2\lambda y \; g - 1] \rightarrow$ necessary conditions: $i)$ $ye^{xy} - 2\lambda x = 0$; ii) $xe^{xy} - 2\lambda y = 0$; iii) $g = 1$.

If $x = 0$ it follows from i) that $y = 0$, which violates iii). So it follows from i) that $\lambda = \frac{y}{2x}e^{xy}$. Substitution of this into ii) yields $(x - \frac{y^2}{x})e^{xy} = 0$. That is, $x^2 = y^2$. Substitution of this into iii) gives $2x^2 = 1$. So, candidate optimal points are: $(\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2})$, $(\frac{1}{2}\sqrt{2}, -\frac{1}{2}\sqrt{2})$, $(-\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2})$ and $(-\frac{1}{2}\sqrt{2}, -\frac{1}{2}\sqrt{2})$.

Evaluating f at the candidate solutions shows that the problem has a global maximum at $(\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2})$ and $(-\frac{1}{2}\sqrt{2}, -\frac{1}{2}\sqrt{2})$; and has a global minimum at $(-\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2})$ and $(\frac{1}{2}\sqrt{2}, -\frac{1}{2}\sqrt{2})$.

h. Note that f is continuous and that the constraint set is compact $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.13 we have:

Step 1. $Df = [y(x^2 + 1)e^{\frac{1}{2}x^2+y^2} \;\; x(2y^2 + 1)e^{\frac{1}{2}x^2+y^2}]$ and $Dg = [2x \;\; 2y]$. Note all partial derivatives are continuous $\rightarrow f, g \in C^1$.

Step 2. Dg has not full row rank iff. simultaneously $2x = 2y = 0$. This only occurs if $(x, y) = (0, 0)$. But $(0, 0)$ is not located in the constraint set.

Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x, y, \lambda) := xye^{\frac{1}{2}x^2+y^2} - \lambda(x^2 + y^2 - 4)$.

Step 4. $DL = [y(x^2 + 1)e^{\frac{1}{2}x^2+y^2} - 2\lambda x \; x(2y^2 + 1)e^{\frac{1}{2}x^2+y^2} - 2\lambda y \; g - 4] \rightarrow$ necessary conditions: i) $y(x^2 + 1)e^{\frac{1}{2}x^2+y^2} - 2\lambda x = 0$; ii) $x(2y^2 + 1)e^{\frac{1}{2}x^2+y^2} - 2\lambda y] = 0$; iii) $g = 4$.

If $x = 0$ it follows from i) that $y = 0$, which violates iii). So w.l.o.g. it follows from i) that $\lambda = \frac{y}{2x}e^{\frac{1}{2}x^2+y^2}$. Substitution of this into ii) yields $(y^2(x^2 + 1) - x^2(2y^2 + 1))e^{\frac{1}{2}x^2+y^2} = 0$. That is, $y^2(x^2 + 1) - x^2(2y^2 + 1) = 0$. Substitution from iii) of $y^2 = 4 - x^2$ into this gives $x^4 - 6x^2 + 4 = 0$. So, candidate optimal points are: $(\sqrt{3 - \sqrt{5}}, \sqrt{1 + \sqrt{5}})$, $(\sqrt{3 - \sqrt{5}}, -\sqrt{1 + \sqrt{5}})$, $(-\sqrt{3 - \sqrt{5}}, \sqrt{1 + \sqrt{5}})$ and $(-\sqrt{3 - \sqrt{5}}, -\sqrt{1 + \sqrt{5}})$.

Evaluating f at the candidate solutions shows that the problem has a global maximum at $(\sqrt{3 - \sqrt{5}}, \sqrt{1 + \sqrt{5}})$ and $(-\sqrt{3 - \sqrt{5}}, -\sqrt{1 + \sqrt{5}})$; and has a global minimum at $(-\sqrt{3 - \sqrt{5}}, \sqrt{1 + \sqrt{5}})$ and $(\sqrt{3 - \sqrt{5}}, -\sqrt{1 + \sqrt{5}})$.

EXERCISE 3.19 Use Lagrange's method to determine all points that qualify for an optimum of the next constrained optimization problems. Use, next, second-order criteria; Weierstrass' theorem; or convexity of the Lagrange function (see Exercise 3.24, below) to determine the optima (location, value, kind (local/global)) of the problems.

a. Optimize $f(x, y, z) = (x-1)^2+(y-1)^2+(z-1)^2$, subject to the constraint $g(x, y) = x^2+y^2+z^2 = 1$.

b. Optimize $f(x, y, z) = x + y + z$, subject to the constraints $g_1(x, y, z) = x^2 + y^2 + z^2 = 12$ and $g_2(x, y, z) = x - y = 0$.

c. Optimize $f(x,y,u,v) = \ln(1+xy)$, subject to the constraint $g_1(x,y,u,v) = x^2+y^2 = 2$, $g_2(x,y,u,v) = xu^2 = 1$ and $g_3(x,y,u,v) = yv^2 = 1$.
d. Optimize $f(x,y) = x^2+y^2$, subject to the constraint $g(x,y) = x^2+xy+y^2 = 3$.
e. Optimize $f(x,y) = x^2+(y-1)^2$, subject to the constraint $g(x,y) = 4x^2+y^2 = 16$.
f. Optimize $f(x,y) = x^2+\ln(2+y^2)$, subject to the constraints $g(x,y) = 2x+y^2 = 2$.
g. Optimize $f(x,y) = e^{xy}$, subject to the constraint $g(x,y) = x^2+y^2 = 1$.
h. Optimize $f(u,s,t,v) = u+s+t+v$, subject to the constraint $g(u,s,t,v) = u^2+s^2+t^2+v^2 = 4$.

ANSWER 3.19

a. Note that f is continuous and that the constraint set is compact $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.34 we have:
Step 1. $Df = [2(x-1)\ 2(y-1)\ 2(z-1)]$ and $Dg = [2x\ 2y\ 2z]$. Note all partial derivatives are continuous $\rightarrow f, g \in C^1$.
Step 2. Dg has not full row rank iff. simultaneously $2x = 2y = 2z = 0$. This only occurs if $(x,y,z) = (0,0,0)$. But $(0,0,0)$ is not located in the constraint set.
Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x,y,\lambda) := (x-1)^2+(y-1)^2+(z-1)^2-\lambda(x^2+y^2+z^2-1)$.
Step 4. $DL = [2(x-1)-2\lambda x\ 2(y-1)-2\lambda y\ 2(z-1)-2\lambda z\ g-1] \rightarrow$ necessary conditions: i) $2(x-1)-2\lambda x = 0$; ii) $2(y-1)-2\lambda y = 0$; iii) $2(z-1)-2\lambda z = 0$; iv) $g = 1$.
From i)-iii) we have $x = y = z = \frac{1}{1-\lambda}$. Substitution into iv) yields: $\frac{3}{(1-\lambda)^2} = 1 \rightarrow \lambda = 1 \pm \sqrt{3}$.
So the candidate solutions are: $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$ and $(-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}})$.
Evaluating f at the candidate solutions shows that the problem has a global maximum at $(-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}})$; and has a global minimum at $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$.

b. Note that f is continuous and that the constraint set is compact $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.34 we have:
Step 1. $Df = [1\ 1\ 1]$ and, with $g := \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}$, $Dg = \begin{bmatrix} 2x & 2y & 2z \\ 1 & -1 & 0 \end{bmatrix}$. Note all partial derivatives are continuous $\rightarrow f, g_i \in C^1$.
Step 2. Note rank $Dg = \text{rank} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2x+2y & 2z \end{bmatrix}$. $\rightarrow Dg$ has not full row rank iff. $2x+2y = 2z = 0$. This only occurs if $(x,y,z) = (s,-s,0)$ with $s \in \mathbb{R}$. From g_2 it follows that this only may occur at $s = 0$, but $(0,0,0)$ is not located in g_1.
Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x,y,\lambda_1,\lambda_2) := x+y+z-\lambda_1(x^2+y^2+z^2-12)-\lambda_2(x-y)$.
Step 4. $DL = [1-2\lambda_1 x-\lambda_2,\ 1-2\lambda_1 y+\lambda_2,\ 1-2\lambda_1 z\ g_1-12\ g_2] \rightarrow$ necessary conditions: i) $1-2\lambda_1 x-\lambda_2 = 0$; ii) $1-2\lambda_1 y+\lambda_2 = 0$; iii) $1-2\lambda_1 z = 0$; iv) $g_1 = 12$; v) $g_2 = 0$.
From i), ii) and v) we have that $\lambda_2 = 0$ and $1-2\lambda_1 x = 0$. So, $x = y = z = \frac{1}{2\lambda_1}$. Substitution of this into g_1 yields: $\frac{3}{4\lambda_1^2} = 12 \rightarrow \lambda_1 = \pm\frac{1}{4}$.
So the candidate solutions are: $(2,2,2)$ and $(-2,-2,-2)$.

Evaluation of f at all these candidate points shows that the problem has a global maximum at $(2,2,2)$; and has a global minimum at $(-2,-2,-2)$.

c. Note that f is continuous. From g_2 we have $x > 0$ and from g_3 we have $y > 0$. So the constraint set where we have to find the optima is bounded but not closed and therefore not compact $\rightarrow$ Weierstrass' theorem is not applicable. One might solve now the optimization problem with just the constraint g_1 and in the end consider the constraints that $x > 0$ and $y > 0$ must be satisfied. Following Algorithm 3.34 for the original problem statement we have:

Step 1. $Df = [\frac{y}{1+xy},\ \frac{x}{1+xy},\ 0,\ 0]$ and, with $g := \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix}$, $Dg = \begin{bmatrix} 2x & 2y & 0 & 0 \\ u^2 & 0 & 2ux & 0 \\ 0 & v^2 & 0 & 2vy \end{bmatrix}$. Note all partial derivatives are continuous $\rightarrow f, g_i \in C^1$.

Step 2. Note from g_2 and g_3 that $u \neq 0$ and $v \neq 0$. $\rightarrow$ rank $Dg =$ rank $\begin{bmatrix} u^2 & 0 & 2ux & 0 \\ 0 & v^2 & 0 & 2vy \\ 0 & 2y & -\frac{4x^2}{u} & 0 \end{bmatrix} =$ rank $\begin{bmatrix} u^2 & 0 & 2ux & 0 \\ 0 & v^2 & 0 & 2vy \\ 0 & 0 & -\frac{4x^2}{u} & -\frac{4y^2}{v} \end{bmatrix}$. $\rightarrow Dg$ has not full row rank iff. $x = y = 0$. But this violates g_1.

Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x,y,u,v,\lambda_1,\lambda_2,\lambda_3) := \ln(1+xy) - \lambda_1(x^2+y^2-2) - \lambda_2(xu^2-1) - \lambda_3(yv^2-1)$.

Step 4. $DL = [\frac{y}{1+xy} - 2\lambda_1 x - \lambda_2 u^2,\ \frac{x}{1+xy} - 2\lambda_1 y - \lambda_3 v^2,\ -2\lambda_2 xu,\ -2\lambda_3 yv,\ g_1 - 2,\ g_2 - 1,\ g_3 - 1]$ $\rightarrow$ necessary conditions: i) $\frac{y}{1+xy} - 2\lambda_1 x - \lambda_2 u^2 = 0$; ii) $\frac{x}{1+xy} - 2\lambda_1 y - \lambda_3 v^2 = 0$; iii) $\lambda_2 xu = 0$; iv) $\lambda_3 yv = 0$; v) $g_1 = 2$; vi) $g_2 = 1$; vii) $g_3 = 1$.

As already noticed above it follows from g_2 that $x > 0$ and $u \neq 0$. So, from iii) it follows that $\lambda_2 = 0$. Similarly it follows from g_3 and iv) that $\lambda_3 = 0$. From i) and ii) we have then that $\lambda_1 = \frac{y}{2x(1+xy)} = \frac{x}{2y(1+xy)}$. So, $x^2 = y^2$, i.e., $x = \pm y$. Notice, as both $x > 0$ and $y > 0$, the possibility of $x = -y$ does not apply. Substitution of $x = y$ into g_1 shows $x^2 = 1 \rightarrow x = 1 \rightarrow y = 1 \rightarrow u = \pm 1$ and $v = \pm 1$.

So the candidate solutions are: $(x,y,u,v) = (1,1,\pm 1,\pm 1)$ with $\lambda_1 = \frac{1}{4}$ and $\lambda_2 = \lambda_3 = 0$.

Next, consider $H\tilde{L} := Hf - \sum_{i=1}^{3} \lambda_i Hg_i =$

$$\begin{bmatrix} -\frac{y^2}{(1+xy)^2} & -\frac{xy}{(1+xy)^2} & 0 & 0 \\ -\frac{xy}{(1+xy)^2} & -\frac{x^2}{(1+xy)^2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \lambda_1 \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \lambda_2 \begin{bmatrix} 0 & 0 & 2u & 0 \\ 0 & 0 & 0 & 0 \\ 2u & 0 & x & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \lambda_3 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2v \\ 0 & 0 & 0 & 0 \\ 0 & 2v & 0 & y \end{bmatrix}.$$

So, at the candidate solutions we get: $H\tilde{L} = \begin{bmatrix} -0.75 & -.25 & 0 & 0 \\ -.25 & -0.75 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

So, with $B := \begin{bmatrix} 2y & 0 & 0 \\ 0 & 2xu & 0 \\ v^2 & 0 & 2yv \end{bmatrix}^{-1} \begin{bmatrix} 2x \\ u^2 \\ 0 \end{bmatrix}$, we see that by evaluating $[1B^T] * H\tilde{L} \begin{bmatrix} 1 \\ B \end{bmatrix}$ at the stationary point, this always yields a negative number. So, this stationary point is a local maximum. Notice that since $(x-y)^2 \geq 0$, it follows that $2xy \leq x^2 + y^2 = 2$. So, $xy \leq 1$. From the fact that the function $\ln z$ is a monotonic increasing function it follows then that the function has a global maximum at the stationary point.

Another way to arrive at this conclusion is to show that $\tilde{L}(x, y, u, v, \frac{1}{4}, 0, 0)$ is a concave function on the constraint set (see Exercise 3.24). To that end consider the first 2×2-block of matrix $-H\tilde{L}(x, y, u, v, \frac{1}{4}, 0, 0)$. That is, matrix

$$A := \begin{bmatrix} \frac{y^2}{(1+xy)^2} + \frac{1}{2} & \frac{xy}{(1+xy)^2} \\ \frac{xy}{(1+xy)^2} & \frac{x^2}{(1+xy)^2} + \frac{1}{2} \end{bmatrix}.$$

The leading principal minors of this matrix are: $\frac{y^2}{(1+xy)^2} + \frac{1}{2}$, $\frac{x^2}{(1+xy)^2} + \frac{1}{2}$ and $(\frac{y^2}{(1+xy)^2} + \frac{1}{2})(\frac{x^2}{(1+xy)^2} + \frac{1}{2}) - \frac{x^2y^2}{(1+xy)^4}$. As already noticed above, $x > 0$ and $y > 0$ on the constraint set. Using this, some elementary analysis shows that all three principal minors of matrix A are positive. This implies that matrix $-H\tilde{L}(x, y, u, v, \frac{1}{4}, 0, 0)$ is positive semi-definite. So, all stationary points of $\tilde{L}(x, y, u, v, \frac{1}{4}, 0, 0)$ provide a global maximum.

d. Note that f is continuous and that the constraint set is compact since $g(x, y) = (x + \frac{1}{2}y)^2 + \frac{3}{4}y^2 = 3$ (so $|y| \leq 2$ and $|x + \frac{1}{2}y| \leq 3$) $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. We use Algorithm 3.34 to find all candidate solutions.

Step 1. $Df = [2x\ 2y)]$ and $Dg = [2x + y\ 2y + x]$. Note all partial derivatives are continuous $\rightarrow$ $f, g \in C^1$.

Step 2. Note Dg has not full row rank iff. $2x + y = 2y + x = 0$. This only occurs if $(x, y) = (0, 0)$. But $(0, 0)$ is not located on g.

Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x, y, \lambda) := x^2 + y^2 - \lambda(x^2 + xy + y^2 - 3)$.

Step 4. $DL = [(2-2\lambda)x - \lambda y\ -\lambda x + (2-2\lambda)y\ g - 3] \rightarrow$ necessary conditions: i) $(2-2\lambda)x - \lambda y = 0$; ii) $-\lambda x + (2-2\lambda)y = 0$; iii) $g = 3$.

From i) and ii) we have that $x = y = 0$ provided $\det A := \det \begin{bmatrix} 2-2\lambda & -\lambda \\ -\lambda & 2-2\lambda \end{bmatrix} \neq 0$. But $g(0, 0) \neq 3$. So, $(0, 0)$ does not qualify as a solution.

$\det A = 0$ iff. $(2-2\lambda)^2 - \lambda^2 = 0$, i.e., if $\lambda = 2$ or $\lambda = \frac{2}{3}$. In that case the equations reduce to the system of equations $i)$ and $g = 3$. In case $\lambda = 2$, $i)$ yields $y = -x$. This yields the candidate solutions $(\sqrt{3}, -\sqrt{3})$ and $(-\sqrt{3}, \sqrt{3})$. In case $\lambda = \frac{2}{3}$, $i)$ yields $y = x$. This yields the additional candidate solutions $(1, 1)$ and $(-1, -1)$.

Evaluation of f at all these candidate points shows that the problem has a global maximum at $(\sqrt{3}, -\sqrt{3})$ and $(-\sqrt{3}, \sqrt{3})$; and has a global minimum at $(-1, -1)$ and $(1, 1)$.

e. Note that f is continuous and that the constraint set is compact. $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.34 we have:
Step 1. $Df = [2x\ 2(y-1))]$ and $Dg = [8x\ 2y]$. Note all partial derivatives are continuous $\rightarrow$ $f, g \in C^1$.
Step 2. Note Dg has not full row rank iff. $(x, y) = (0, 0)$. But $(0, 0)$ is not located on g.
Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x, y, \lambda) := x^2 + (y-1)^2 - \lambda(4x^2 + y^2 - 16)$.
Step 4. $DL = [(2 - 8\lambda)x\ (2 - 2\lambda)y - 2\ g - 16] \rightarrow$ necessary conditions: i) $(2 - 8\lambda)x = 0$; ii) $(2 - 2\lambda)y - 2 = 0$; iii) $g = 16$.
From i) we have that $x = 0$ or $\lambda = \frac{1}{4}$. In case $x = 0$, by iii) $y = \pm 4$. From ii) one can derive then the corresponding values of λ. If $\lambda = \frac{1}{4}$, we obtain from ii) that $y = \frac{4}{3}$. Then, by iii), $x = \pm\frac{4}{3}\sqrt{2}$.
Evaluation of f at all these candidate points shows that the problem has a global minimum at $(\frac{4}{3}\sqrt{2}, \frac{4}{3})$ and $(-\frac{4}{3}\sqrt{2}, \frac{4}{3})$; and has a global maximum at $(0, 4)$ and $(0, -4)$.

f. Note that f is continuous and that the constraint set is compact. $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following the algorithm:
Step 1: $Df = [2x\ \frac{2y}{2+y^2}]$ and $Dg = [4x\ 2y]$. Note all partial derivatives are continuous $\rightarrow f, g \in C^1$.
Step 2: Note Dg has not full row rank iff. $(x, y) = (0, 0)$. But $(0, 0)$ is not located on $g \rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x, y, \lambda) := x^2 + \ln(2 + y^2) - \lambda(2x^2 + y^2 - 2)$.
Step 3: $DL = [(2-4\lambda)x\ \frac{2y}{2+y^2} - 4\lambda y\ g - 2] \rightarrow$ necessary conditions: i) $(2-4\lambda)x = 0$; ii) $\frac{2y}{2+y^2} - 4\lambda y = 0$; iii) $g = 2$.
From i) we have that $x = 0$ or $\lambda = \frac{1}{2}$. In case $x = 0$, by iii) $y = \pm\sqrt{2}$. From ii) one can derive then the corresponding values of λ. If $\lambda = \frac{1}{2}$, we obtain from ii) that $y = 0$. Then, by iii), $x = \pm 1$.
Evaluation of f at all these candidate points shows that the problem has a global maximum at $(0, \sqrt{2})$ and $(0, -\sqrt{2})$; and has a global minimum at $(1, 0)$ and $(-1, 0)$.

g. Note that f is continuous and that the constraint set is compact. $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.34 candidate solutions of the problem are obtained as:
Step 1. $Df = [ye^{xy}\ xe^{xy}]$ and $Dg = [2x\ 2y]$. Note all partial derivatives are continuous $\rightarrow f, g \in C^1$.
Step 2. Note Dg has not full row rank iff. $(x, y) = (0, 0)$. But $(0, 0)$ is not located on g.
Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x, y, \lambda) := e^{xy} - \lambda(x^2 + y^2 - 1)$.
Step 4. $DL = [ye^{xy} - 2\lambda x\ xe^{xy} - 2\lambda y\ g - 1] \rightarrow$ necessary conditions: i) $ye^{xy} - 2\lambda x = 0$; ii) $xe^{xy} - 2\lambda y = 0$; iii) $g = 1$. From i) we have that if $y \neq 0$, $e^{xy} = \frac{2\lambda x}{y}$. Substitution of this into ii) shows that $2\lambda(y^2 - x^2) = 0$. So, either $\lambda = 0$ or $x = \pm y$ should hold. We next elaborate all cases.
1. Case $y = 0$. Then by i) $x = 0$ or $\lambda = 0$. $g(0, 0) \neq 1$, so this possibility drops out. If $\lambda = 0$, by ii) also $x = 0$ results again. Which, as argued above, together with $y = 0$ does not provide an appropriate solution.

2. Case $\lambda = 0$. Again, by i) and ii), we get $(0, 0)$ as a candidate solution, but this does not satisfy iii).
3. Case $x = y$. Then, by e.g. i), $x(e^{x^2} - 2\lambda) = 0$. So, again the inappropriate solution $(0, 0)$ results, or, by iii) the solutions $(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$ and $(-\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$ are obtained.
4. Case $x = -y$. Then, similarly as above, the solutions $(-\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$ and $(\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$ are obtained.

Evaluation of f at all these candidate points shows that the problem has a global maximum at $(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$ and $(-\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$; and has a global minimum at $(-\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$ and $(\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$.

h. Note that f is continuous and that the constraint set is compact. $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.34 we obtain candidate solutions as follows:

Step 1. $Df = [1\ 1\ 1\ 1]$ and $Dg = [2u\ 2s\ 2t\ 2v]$. Note all partial derivatives are continuous $\rightarrow$ $f, g \in C^1$.

Step 2. Note Dg has not full row rank iff. $(u, s, t, v) = (0, 0, 0, 0)$. But $(0, 0, 0, 0)$ is not located on g.

Step 3. $\rightarrow$ All candidate solutions are stationary points of the associated Lagrange function $L(x, y, \lambda) := u + s + t + v - \lambda(u^2 + s^2 + t^2 + v^2 - 4)$.

Step 4. $DL = [1 - 2\lambda u\ 1 - 2\lambda s\ 1 - 2\lambda t\ 1 - 2\lambda v\ g - 4] \rightarrow$ necessary conditions: i) $1 - 2\lambda u = 0$; ii) $1 - 2\lambda s = 0$; iii) $1 - 2\lambda t = 0$; iv) $1 - 2\lambda v = 0$; v) $g = 4$.

As (see e.g. i)) $\lambda \neq 0$ from i)-iv) we have that $u = s = t = v = \frac{1}{2\lambda}$. Then, by v), $u = \pm 1$. So, the only candidate solutions are: $(1, 1, 1, 1)$ and $(-1, -1, -1, -1)$.

Evaluation of f at all these candidate points shows that the problem has a global maximum at $(1, 1, 1, 1)$ and has a global minimum at $(-1, -1, -1, -1)$.

Advanced, Deepening and Broadening Exercises

EXERCISE 3.20 *(solving a constrained optimization problem under linear equality constraints)*

In this exercise we indicate two ways how one may pursue to solve a constrained optimization problem subject to linear equality constraints.

Consider the optimization of $f(x)$ subject to $Ax = b$, where $A \in \mathbb{R}^{m \times n}$ is full row rank.

a. Show that by a simple relabeling the columns of matrix A can be "reshuffled" as $[A_1\ A_2]$, where matrix A_1 is invertible. Deduct from this that the constrained optimization problem can be solved by finding the optima of the unconstrained function $f(x_2)$, where $[x_1^T\ x_2^T]^T$ is the corresponding "reshuffled" vector x with $x_2 \in \mathbb{R}^{n-m}$ and $x_1 = A_1^{-1}(b - A_2 x_2)$.

b. Let $A^+ := A^T(AA^T)^{-1}$. Show that the solution set of $\{x \mid Ax = b\}$ equals $\{A^+b + (I - A^+A)y \mid y \in \mathbb{R}^n\}$. Deduct from this that the constrained optimization problem can be solved by finding the optima of the unconstrained function $f(A^+b + (I - A^+A)y),\ y \in \mathbb{R}^n$.

c. Use both methods from item a. and b. to find the optima of the constrained optimization problem: $f(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 - x_3^2 - x_4^2$ subject to $x_2 + x_3 = 2$ and $x_3 + x_4 = 1$.

ANSWER 3.20

a. Since matrix A has full row rank, by a "reshuffling" of columns (which implies the order of the variables x_i in vector x are changed) we can arrange that matrix A can be rewritten as $[A_1\ A_2]$, where matrix A_1 is invertible. Or, stated in mathematical terms: there exists a permutation matrix P (wich has the property that $P^2 = I$) such that $[x_1^T\ x_2^T]^T = Px$ and $[A_1\ A_2] = AP$. That is, $Ax = b$ can be rewritten as $APPx = b$, or as: $A_1x_1 + A_2x_2 = b$. Since by construction A_1 is invertible, we have from the last equality that $x_1 = A_1^{-1}(b - A_2x_2)$. So, the constrained optimization problem can be rewritten as the optimization of $f(x) = f(P^2x) = f(P[x_1^T\ x_2^T]^T)$, with $x_1 = A_1^{-1}(b - A_2x_2)$, and $x_2 \in \mathbb{R}^{n-m}$ arbitrary.

b. Let $x = A^+b + (I - A^+A)y$, with $y \in \mathbb{R}^{n-m}$ arbitrary. Then

$$Ax = AA^+b + A(I - A^+A)y = AA^T(AA^T)^{-1}b + Ay - AA^T(AA^T)^{-1}Ay = b + Ay - Ay = b.$$

So, every advertized solution satisfies indeed $Ax = b$.

Next we show that every solution of $Ax = b$ can be rewritten as $A^+b + (I - A^+A)y$, for some $y \in \mathbb{R}^{n-m}$.

Let x_p solve $Ax = b$. Then

$$x_p = A^+Ax_p + (I - A^+A)x_p = A^+b + (I - A^+A)x_p.$$

That is, with $y = x_p$ the solution can be rewritten into the above mentioned form.

Consequently, the constrained optimization problem can be rewritten into the advertized unconstrained form.

c. Notice the linear constraint set is described by $Ax = b$, with $A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$ and $b = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

First, consider the method described in item a.

Notice that by "reshuffling" the first column of A towards the third position, $A = [A_1\ A_2]$ with $A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. That is, with the permutation matrix

$P := \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ we have $AP = [A_1\ A_2]$ and $[y_1^T\ y_2^T]^T = Px = [x_2\ x_3\ x_1\ x_4]$.

So, the constrained set can be rewritten as $Ax = APPx = A_1y_1 + A_2y_2 = b$. From which we have:

$$y_1 = A_1^{-1}(b - A_2y_2) = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 2 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} y_2 \right).$$

That is:

$$\begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 2 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} \right) = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 - x_4 \end{bmatrix} = \begin{bmatrix} 1 + x_4 \\ 1 - x_4 \end{bmatrix}.$$

Substitution of this into f yields then the unconstrained optimization problem to find all extremal points of $f(x_1, x_4) = x_1^2 + (1 + x_4)^2 - (1 - x_4)^2 - x_4^2 = x_1^2 + 4x_4 - x_4^2$. This unconstrained problem has one stationary point $(x_1, x_4) = (0, 2)$, which is a saddle-point. So, the original constrained optimization problem has a saddle-point at $x = (0, 3, -1, 2)$.

Next, consider the method described in item b.

Notice $AA^T = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Therefore, $A^+ = A^T(AA^T)^{-1} = \frac{1}{3} \begin{bmatrix} 0 & 0 \\ 2 & -1 \\ 1 & 1 \\ -1 & 2 \end{bmatrix}$. So, $A^+b = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$ and $I - A^+A = \frac{1}{3} \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & -1 & 1 & -1 \\ 0 & 1 & -1 & 1 \end{bmatrix}$. Thus, with $z := \frac{1}{3}(y_2 - y_3 + y_4)$, the unconstrained optimization problem becomes to optimize $f(y_1, z) = y_1^2 + (1 + z)^2 - (1 - z)^2 - z^2 = y_1^2 + 4z - z^2$. As before we obtain a saddle-point at $(y_1, z) = (0, 2)$. This implies that the original constrained optimization problem has a saddle-point at $x = A^+b + (I - A^+A)y$, where $y = (0\ y_2\ y_3\ y_4)^T$, with $y_2 - y_3 + y_4 = 6$. That is, at $x = (0, 3, -1, 2)$.

EXERCISE 3.21 *(Step 2 optimum)*

Consider the problem to find the minimum of $f(x, y) = (x+2)^2+y^2$ subject to the constraint $g(x, y) = x(x + 1)^2 - y^2 = 0$.

a. Find all candidate extremal points using Lagrange's method.

b. Show that $(-1, 0)$ yields the global minimum of the constrained optimization problem.

ANSWER 3.21

a. Note that f is continuous. Using Algorithm 3.13 we find all candidate solutions as follows:

Step 1. $Df = [2(x + 2)\ 2y]$ and $Dg = [(x + 1)(3x + 1)\ \ -2y]$. Note all partial derivatives are continuous $\rightarrow f, g \in C^1$.

Step 2. Note Dg has not full row rank at either $(-1, 0)$ or $(-\frac{1}{3}, 0)$. However, notice $g(-\frac{1}{3}, 0) \neq 0$. As $g(-1, 0) = 0$, this is a candidate solution.

Step 3. All other candidate solutions are stationary points of the associated Lagrange function: $L(x, y, \lambda) := (x + 2)^2 + y^2 - \lambda(x(x + 1)^2 - y^2)$.

Step 4. $DL = [2(x+2) - \lambda(x+1)(3x+1)\ 2y(1+\lambda)\ g] \rightarrow$ necessary conditions: i) $2(x+2) - \lambda(x+1)(3x+1) = 0$; ii) $2y(1+\lambda) = 0$; iii) $g = 0$.

From ii) we have that either $\lambda = -1$ or $y = 0$ should hold. We next elaborate both cases.

1. Case $y = 0$. Then by iii) $x = 0$ or $x = -1$. However $(-1, 0)$ does not satisfy i), so this possibility drops out. If $x = 0$, λ follows from i). So, $(0, 0)$ is a candidate solution.
2. Case $\lambda = -1$. Then, by i), no solutions result.

Notice $f(-1, 0) < f(0, 0)$. So, the problem has not a global minimum at $(0, 0)$.

b. From the constraint it follows that $y^2 = x(x+1)^2$. Substitution of this into f gives:

$$f(x) = (x+2)^2 + x(x+1)^2, \text{ with } x \in \{-1\} \cup [0, \infty)$$

Notice $f^{'}(x) \geq 0$ for $x \geq 0$. Therefore, f has a global minimum at either $x = 0$ or $x = -1$. As $f(-1) < f(0)$, it follows that this optimization problem has a global minimum and is attained at $(-1, 0)$.

EXERCISE 3.22 *(Infinite number of optimum locations)*

Consider the problem to find the minimum of $f(x, y, z) = (x-1)^2 + (y-1)^2 + (z-1)^2$ subject to the constraints $g_1(x, y, z) = x^2 + y^2 + z^2 = 3$ and $g_2(x, y, z) = x + y + z = 0$.

a. Find all extremal points by a direct use of Lagrange's method.

b. Use the constraint g_2 to reduce the problem to a constrained optimization problem in two variables. Solve this reduced problem using Lagrange's method. Compare your answer with item a.

ANSWER 3.22

a. Note that f is continuous and that the constraint set is compact. $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.34 we obtain all candidate extremal points as follows:

Step 1. $Df = [2(x-1)\ 2(y-1)\ 2(z-1)]$ and, with $g := \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}$, $Dg = \begin{bmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{bmatrix}$. Note all partial derivatives are continuous $\rightarrow f, g_i \in C^1$.

Step 2. Note Dg has not full row rank iff $x = y = z$. From $g_2 = 0$, we have then that $x = y = z = 0$. However, $g_1(0, 0, 0) \neq 0$. So $(0, 0, 0)$, is not a candidate solution.

Step 3. All other candidate solutions are stationary points of the associated Lagrange function

$$L(x, y, z, \lambda_1, \lambda_2) := (x-1)^2 + (y-1)^2 + (z-1)^2 - \lambda_1(x^2 + y^2 + z^2 - 3) - \lambda_2(x + y + z).$$

Step 4. $DL = [2(x-1) - 2\lambda_1 x - \lambda_2\ 2(y-1) - 2\lambda_1 y - \lambda_2\ 2(z-1) - 2\lambda_1 z - \lambda_2\ g_1 - 3\ g_2] \rightarrow$ necessary conditions: i) $2(x-1) - 2\lambda_1 x - \lambda_2 = 0$; ii) $2(y-1) - 2\lambda_1 y - \lambda_2 = 0$ iii) $2(z-1) - 2\lambda_1 z - \lambda_2 = 0$; iv) $g_1 = 3$; and v) $g_2 = 0$.

From i)-iii) we have that $\lambda_2 + 2 = 2x(1 - \lambda_1) = 2y(1 - \lambda_1) = 2z(1 - \lambda_1)$. From this it is easily

verified that with $\lambda_1 = 1$ and $\lambda_2 = -2$ all points satisfying iv) and v) satisfy all equations. So all points in the constrained set are candidate solutions.

b. From g_2 it follows that $z = -x - y$. Substitution of this into f gives the equivalent minimization problem:

$$f(x,y) = (x-1)^2 + (y-1)^2 + (x+y+1)^2, \text{ subject to } g(x,y) = x^2 + y^2 + (x+y)^2 = 3.$$

Note that f is continuous and that the constraint set is compact. $\rightarrow$ (Weierstrass' theorem) The problem has both a global maximum and a global minimum. Following Algorithm 3.34, we obtain all candidate extremum locations as follows:

Step 1. $Df = [2(x-1)+2(x+y+1)\ 2(y-1)+2(x+y+1)]$ and $Dg = [2x+2(x+y)\ 2y+2(x+y)]$. Note all partial derivatives are continuous $\rightarrow f, g \in C^1$.

Step 2. Note Dg has not full row rank iff $x = y = 0$, but $g(0,0) \neq 3$. So $(0,0)$ is not a candidate solution.

Step 3. All other candidate solutions are stationary points of the associated Lagrange function

$$L(x,y,\lambda) := (x-1)^2 + (y-1)^2 + (x+y+1)^2 - \lambda(x^2 + y^2 + (x+y)^2 - 3).$$

Step 4. $DL = [2(2x+y) - 2\lambda(2x+y)\ 2(x+2y) - 2\lambda(x+2y)\ g-3] \rightarrow$ necessary conditions: i) $2(2x+y) - 2\lambda(2x+y) = 0$; ii) $2(x+2y) - 2\lambda(x+2y) = 0$; and iii) $g = 3$.

From i)-ii) we have that $(2x+y)(1-\lambda) = (x+2y)(1-\lambda) = 0$. So, again it is easily verified that with $\lambda = 1$ all points that satisfy the constraint satisfy all equations. So all points in the constrained set are still candidate solutions.

Finally, notice $f(x,y) = 2(x^2+y^2+xy) + 3$, whereas $g(x,y) = 2(x^2+y^2+xy)$. So, it is clear that all points on g are optimal and the global minimum (and maximum) of the optimization problem is 6.

EXERCISE 3.23 *(Reformulation of strict inequality constraint problem as an equality constraint problem)* Consider the problem to find optima of $f(x,y) = \ln(1+xy)$, subject to the equality constraint $g(x,y) = x^2 + y^2 = 2$ and inequality constraints $x > 0$ and $y > 0$.

a. Show that $x > 0$ if and only if there exists u such that $xu^2 = 1$.

b. Show that the optimization problem is equivalent with solving next optimization problem $f(x,y,u,v) = \ln(1+xy)$, subject to the equality constraints $g_1(x,y,u,v) = x^2 + y^2 = 2$, $g_2(x,y,u,v) = xu^2 = 1$ and $g_3(x,y,u,v) = yv^2 = 1$.

c. Solve the optimization problem.

ANSWER 3.23

a. $\Rightarrow$ If $x > 0$ it follows the equation $xu^2 = 1$ has the solutions $u = \frac{1}{\sqrt{x}}$ and $u = -\frac{1}{\sqrt{x}}$.
$\Leftarrow$ If $xu^2 = 1$ has a solution u for some x, it follows that $x \neq 0$ and $\frac{1}{x} = u^2 > 0$. So, $x > 0$.

b. This follows directly from item a.

c. Consider the optimization problem as reformulated in item b.

Note f is continuous on its domain. Following Algorithm 3.34, with $g := \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix}$ we find all candidate solutions as follows:

Step 1. $Df = [\frac{y}{1+xy} \ \frac{x}{1+xy}]$ and $Dg = \begin{bmatrix} 2x & 2y & 0 & 0 \\ u^2 & 0 & 2xu & 0 \\ 0 & v^2 & 0 & 2yv \end{bmatrix}$. Note all partial derivatives are continuous $\rightarrow f, g \in C^1$.

Step 2. Note, as all variables differ from zero, Dg has always full row rank.

Step 3. Therefore all candidate solutions are stationary points of the associated Lagrange function

$$L(x, y, \lambda) := \ln(1 + xy) - \lambda_1(x^2 + y^2 - 2) - \lambda_2(xu^2 - 1) - \lambda_3(yv^2 - 1).$$

Step 4. $DL = [\frac{y}{1+xy} - 2\lambda_1 x - \lambda_2 u^2 \ \frac{x}{1+xy} - 2\lambda_1 y - \lambda_3 v^2 \ -2\lambda_2 xu \ -2\lambda_3 yv \ g^T] \rightarrow$ necessary conditions: i) $\frac{y}{1+xy} - 2\lambda_1 x - \lambda_2 u^2 = 0$; ii) $\frac{x}{1+xy} - 2\lambda_1 y - \lambda_3 v^2 = 0$; iii) $\lambda_2 xu = 0$; iv) $\lambda_3 yv = 0$; v) $g_1 = 2$; vi) $g_2 = 1$; and vii) $g_3 = 1$.

As x, y, u and v differ from zero it follows from iii) and iv) that $\lambda_2 = \lambda_3 = 0$. By multiplying i) with y and ii) by x and next subtracting both resulting equalities we have that $\lambda_1(y^2 - x^2) = 0$. So we have that either $\lambda_1 = 0$ or $y^2 = x^2$ should hold. We next elaborate both cases.

1. Case $\lambda_1 = 0$. Then by i) $y = 0$, which violates $vii)$.
2. Case $y^2 = x^2$. Then, by v): $x^2 = 1$ So, we just obtain one candidate solution, i.e $(1, 1)$.

To show that $f(1, 1)$ is a global maximum, one may pursue in two ways.

1. First, one could consider the related optimization problem on the constraint set $g_1 = 2$, $x \geq \epsilon$, $y \geq \epsilon$, where $\epsilon > 0$ is arbitrarily small. By Weierstrass' theorem, this problem has both a global minimum and maximum. One easily verifies that the global maximum is attained at $(1, 1)$, irrespective of ϵ. This implies that also if ϵ converges to zero, the corresponding optimization problem has a global maximum at $(1, 1)$.
2. Second, recall that $(x - y)^2 \geq 0$. From this it follows, using the fact that $g_1 = 2$, that $xy \leq 1$. Since $\ln(1 + z) \leq \ln 2$ if $z \leq 1$, it follows then immediately that our constrained optimization problem has a global maximum at $(1, 1)$.

EXERCISE 3.24 *(Stationary point of a convex Lagrange function provides global constraint minimum)* Below we show next property. Assume f and g are differentiable; for some fixed λ, the Lagrange function $L(x) = f(x) - \lambda(g(x) - b)$ has a stationary point at $x = a$; $g(a) = b$; and $L(x)$ is convex on an open convex set S containing the constraint set. Then $f(x)$, subject to the constraint $g(x) = b$, attains a global minimum at $x = a$.

a. Show that $L(a) \leq L(x)$.

b. Use item a to show that $f(a) \leq f(x)$ for all x that satisfy $g(x) = b$.

c. Formulate a similar sufficient condition under which $x = a$ yields a global maximum of the constrained optimization problem.

ANSWER 3.24

a. Recall that if a differentiable convex function has a stationary point a on an open convex set S, then this function has a global minimum at a (see Corollary 2.38). By assumption L has these properties. From this it follows then directly that $L(x) \geq L(a)$.

b. As $g(x) = b$ for all x in the constraint set (and in particular this holds for $x = a$), it follows from the inequality $L(x) \geq L(a)$ directly that $f(x) \leq f(a)$ for all x in the constraint set.

c. Assume f and g are twice differentiable; for some fixed λ, the Lagrange function $L(x) = f(x) - \lambda(g(x) - b)$ has a stationary point at $x = a$; $g(a) = b$; and $L(x)$ is concave. Then $f(x)$, subject to the constraint $g(x) = b$, attains a global maximum at $x = a$.

EXERCISE 3.25 *(Optimization (homogeneous) function under linear equality constraints)*
Assume $f(x)$ is twice differentiable and the optimization of f subject to the set of linear equality constraints $A^T x = p$ has a solution.

a. Show that this solution $x^*(A, p)$ is a homogeneous function of degree zero.

b. Assume f is a homogeneous function of some degree $k \geq 0$. Show that the solution of the problem to optimize f subject to the set of linear equality constraints $A^T x = \alpha p,\ \alpha \geq 1$, has a solution x^N that satisfies $A^T x^N = \alpha p$ and $x^N = \alpha x^*$.

ANSWER 3.25

a. If the entries of matrix A and vector p increase with the same factor t, the solution set of the linear constraints does not change. Consequently the solution of the optimization problem does not change either. So, $x^*(tA, tp) = x^*(A, p)$.

b. First notice that from the definition of homegeneity of f we have $f(sx) = s^k f(x)$. So by differentiation of both sides of this equality w.r.t. x_i it follows that $D_i f(sx) = s^{k-1} D_i f(x)$.
Consider next the Lagrange function $L(x, \lambda) = f(x) - \lambda^T (A^T x - \alpha p)$. Then, by assumption, for $\alpha = 1$ the corresponding set of first-order equations has a solution $\lambda^*,\ x^*$ and, moreover, $A^T x^* = p$. In particular we obtain then that at $x^N := \alpha x^*$, $A^T x^N = \alpha p$ and

$$\begin{aligned} L_x^{'}(x^N, \alpha^{k-2}\lambda^*) &= f^{'}(\alpha x^*) - \alpha {\lambda^*}^T A^T x^* \\ &= \alpha^{k-1} f^{'}(x^*) - \alpha^{k-2} {\lambda^*}^T A^T \alpha x^* \\ &= \alpha^{k-1}(f^{'}(x^*) - {\lambda^*}^T A^T x^*) = 0. \end{aligned}$$

So, $x^N,\ \alpha^{k-2}\lambda^*$ satisfy the first-order conditions of the Lagrange function. Furthermore, $f(x^N) = \alpha^k f(x^*) \geq \alpha^k f(x) \geq f(x)$ for an arbitrary x.
The advertized result follows now directly from the definition of x^N.

EXERCISE 3.26 *(Optimization linear function under linear equality constraints)*
In this exercise we consider the optimization of a linear function $f(x) = a^T x$ subject to linear equality constraints $g(x) := Bx = b$, where $a,\ x \in \mathbb{R}^n$ and $B \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Without loss of generality assume both a and B have full row rank (which implies in particular that the number of constraints at most equals n). Let $D := \begin{bmatrix} a^T \\ B \end{bmatrix}$.

a. Consider the case $n = 2$ and $m = 1$. Show the problem has an infinite number of solutions iff rank $D = 1$. Show that in that case f has a constant value on the constrained set.
Show that, otherwise, the problem has an infimum of $-\infty$ and a supremum of ∞.
b. Consider the case $n = m = 2$. Show that the problem has both a global maximum and minimum that is realized at the same unique point.
c. Show that for arbitrary $m \leq n$ the problem has a global minimum iff either $m = n$ or the rank of matrix D equals m. In both situations also the problem has a global maximum, that coincides with the minimum. Show that in the firstmentioned case the location where the optimum is attained is unique and in the lastmentioned case f attains the same value at all points in the constraint set.
Show next that if none of both cases applies the problem has an infimum of $-\infty$ and a supremum of ∞.
d. Show the problem has a solution iff rank $D =$ rank B.

ANSWER 3.26

a. Consider $L(x, \lambda) = a^T x - \lambda(Bx - b)$. From this, next first-order conditions for the problem result:

$$\begin{aligned} a^T - \lambda B &= 0 \\ Bx &= b \end{aligned}$$

This set of equations only has a solution if vector a^T is a multiple of vector B.
That is, if rank $\begin{bmatrix} a^T \\ B \end{bmatrix} = 1$.
Now assume above rank condition applies. Then it follows that $a^T x = \lambda B x = \lambda b$, for all x satisfying the constraint. That is, f has both a global minimum and maximum on the constraint set which value, moreover, coincides and is attained at any point satisfying the constraint.
In case the rank condition is not met, it follows that rank $\begin{bmatrix} a^T \\ B \end{bmatrix} = 2$. Consequently the equation $\begin{bmatrix} a^T \\ B \end{bmatrix} x = \begin{bmatrix} \mu \\ b \end{bmatrix}$ has a solution for every $\mu \in \mathbb{R}$ (viz $x = \begin{bmatrix} a^T \\ B \end{bmatrix}^{-1} \begin{bmatrix} \mu \\ b \end{bmatrix}$). That is on the constraint set it is possible to realize both arbitrary negative and arbitrary positive values for $f(x) = a^T x$.
b. If $n = m = 2$ it follows that matrix B is invertible. So the constrained set consists just of the single point $x = B^{-1}b$. From this observation the stated claims follow immediately.

c. If $m = n$ it follows that matrix B is invertible. Therefore the constraint set consists just of the point $x = B^{-1}b$. From which the corresponding claims follow directly.
Next consider the case that $m < n - 1$ and rank D =rank $B = m$. Consider the with this problem corresponding Lagrangian function $L(x, \lambda) = a^T x - \lambda^T (Bx - b)$ and resulting first-order conditions:

$$\begin{aligned} a^T - \lambda^T B &= 0 \\ Bx &= b \end{aligned}$$

From the rank condition rank D = rank B it follows that the first equation above has a solution λ. Moreover, since matrix B has full row rank it follows that this λ is unique. Consequently $a^T x = \lambda^T Bx = \lambda^T b$, for all x in the constraint set. From which again immediately the stated results follow.
In case the rank condition is not met, it follows that rank $\begin{bmatrix} a^T \\ B \end{bmatrix} = m + 1$. Consequently the equation $\begin{bmatrix} a^T \\ B \end{bmatrix} x = \begin{bmatrix} \mu \\ b \end{bmatrix}$ has a solution for every $\mu \in \mathbb{R}$ (viz $x = \begin{bmatrix} a^T \\ B \end{bmatrix}^+ \begin{bmatrix} \mu \\ b \end{bmatrix}$). That is on the constraint set it is possible to realize both arbitrary negative and arbitrary positive values for $f(x) = a^T x$.

d. This follows directly from item c.

EXERCISE 3.27 *(Least-squares optimization under equality constraints)*
The Least-squares approximation problem is the problem to find a vector $x \in \mathbb{R}^m$ such that the distance between a given vector $b \in \mathbb{R}^n$ and Ax is as small as possible. Here A is a given matrix in $\mathbb{R}^{n \times m}$, with $n > m$, which has full column rank. In this exercise you are asked to find the solution to this problem under the additional constraint that vector x should satisfy the set of linear constraints $Ex = d$, where $E \in \mathbb{R}^{k \times m}$ and $d \in \mathbb{R}^k$.

a. Use Lagrange's theorem to show that necessary conditions for solvability of the problem:

$$\min_x (Ax - b)^T Q (Ax - b) \text{ subject to } Ex = d, \tag{3.1}$$

where $Q > 0$ is some weighting matrix, are

$$\begin{bmatrix} A^T QA & -E^T \\ E & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} A^T Qb \\ d \end{bmatrix}. \tag{3.2}$$

b. Use Exercise 3.24 to show that the advertized conditions in item a. are sufficient too for the existence of a solution.

c. Recall that a square matrix B is invertible if the only solution to $By = 0$ is $y = 0$. Show that matrix $A^T QA$ is invertible.

d. Show that the least-squares problem has a solution only if the equation $E(A^TQA)^{-1}E^T\lambda = d - E(A^TQA)^{-1}A^TQb$ has a solution λ.

e. Assume $k \leq m$ and E has full row rank. Show that the least squares problem under equality constraints has the unique solution:

$$x = (A^TQA)^{-1}\{A^TQb + E^T(E(A^TQA)^{-1}E^T)^{-1}(d - E(A^TQA)^{-1}A^TQb)\}.$$

You may use the fact that if a matrix Q is positive definite, also its inverse is positive definite.

f. Reconsider Example 3.15 to find a Mean-Variance Optimal Portfolio. That is, to find the solution of the optimization problem:

$$\begin{aligned}&\text{minimize}_{\omega_b,\omega_s}\sigma^2(\omega_b,\omega_s) := (1-\omega_b-\omega_s)^2\sigma_0^2 + \omega_b^2\sigma_b^2 + \omega_s^2\sigma_s^2\\ &\text{subject to } \mu(\omega_b,\omega_s) := (1-\omega_b-\omega_s)(1+r)X + \omega_b(1+r_b)X + \omega_s(1+r_s)X = \bar{X}.\end{aligned}$$

Let $x := [\omega_0\ \omega_b\ \omega_s]$, with $\omega_0 = 1 - \omega_b - \omega_s$. Use above theory to show its solution is: $x = Q^{-1}E^T(EQ^{-1}E^T)^{-1}d$, for appropriately chosen matrices Q and E and vector d. Find the solution if $r = 0.02$, $r_b = 0.05$, $r_s = 0.15$, $X = 100$, $\bar{X} = 105$, $\sigma_0^2 = 0.01$, $\sigma_b^2 = 0.05$ and $\sigma_s^2 = 0.2$.

g. Consider the unconstrained case, i.e., $E = 0$ and $d = 0$.
Two vectors $y,\ z \in \mathbb{R}^m$ are called perpendicular if their inner product $y^Tz = 0$.
Show that $(A\hat{x})^TQ(b - A\hat{x}) = 0$ if $\hat{x}$ solves (3.2). Interpret this result geometricaly, consider first the case $Q = I$.

ANSWER 3.27

a. With $L(x,\lambda) := (Ax-b)^TQ(Ax-b) - (Ex-d)^T\lambda$ we obtain by straightforward differentiation w.r.t. $x_i,\ i = 1,\cdots,m$, the first-order conditions:

$$1.\ A^TQ(Ax-b) - E^T\lambda = 0 \text{ and } 2.\ Ex = d.$$

Which can be rewritten simultaneously as the matrix equation (3.2).

b. Consider a fixed λ. By twice differentiation of $L(x,\lambda)$ we obtain $HL(x,\lambda) = Q > 0$. So, $L(x,\lambda)$ is convex on $\mathbb{R}^m$ and therefore, by Exercise 3.24, the conditions derived in item a. are sufficient for existence of a solution for the least squares problem too.

c. Consider $A^TQAx = 0$. Then also $x^TA^TQAx = 0$, or, with $y = Ax$, $y^TQy = 0$. Since Q is positive definite by definition $y^TQy = 0$ only if $y = 0$. That is, if $Ax = 0$. Since A is a matrix that has full column rank it follows then that necessarily $x = 0$. That is, A^TQA is invertible.

d. From condition 1 in item a. and item c. we have that $x = (A^TQA)^{-1}(A^TQb + E^T\lambda)$. Substitution of this into condition 2 from item a. yields the advertized result.

e. From item c. we know that A^TQA is positive definite. Consequently, $(A^TQA)^{-1}$ is also positive definite. Since E has full row rank, and therefore E^T has full column rank, it follows - by the same reasoning as in item c. - that $E^T(A^TQA)^{-1}E$ is invertible as well.

From item d. it then follows that

$$\lambda = (E^T(A^TQA)^{-1}E)^{-1}(d - E(A^TQA)^{-1}A^TQb).$$

Substituting this into

$$x = (A^TQA)^{-1}(A^TQb + E^T\lambda)$$

(see proof item d.) yields the stated result.

f. Let $x := [\omega_0\ \omega_b\ \omega_s]^T$, with $\omega_0 = 1 - \omega_b - \omega_s$. We can then rewrite the optimization problem as

$$\min_x \sigma_0^2 x^T[1\ 0\ 0]^T[1\ 0\ 0]x + \sigma_b^2 x^T[0\ 1\ 0]^T[0\ 1\ 0]x + \sigma_s^2 x^T[0\ 0\ 1]^T[0\ 0\ 1]x$$

subject to

$$\begin{bmatrix} (1+r)X & (1+r_b)X & (1+r_s)X \\ 1 & 1 & 1 \end{bmatrix} x = \begin{bmatrix} \bar{X} \\ 1 \end{bmatrix}.$$

So, with $Q := \begin{bmatrix} \sigma_0^2 & 0 & 0 \\ 0 & \sigma_b^2 & 0 \\ 0 & 0 & \sigma_s^2 \end{bmatrix}$, $A := I$, $E := \begin{bmatrix} (1+r)X & (1+r_b)X & (1+r_s)X \\ 1 & 1 & 1 \end{bmatrix}$, $b := 0$ and $d := \begin{bmatrix} \bar{X} \\ 1 \end{bmatrix}$, we can rewrite the problem as advertized in item a.

From the construction of matrix Q it is clear that Q is positive definite. In particular it is therefore invertible. Furthermore, matrix E has full row rank ($r_s > r_b > r > 0$, by assumption). Consequently, matrix $EQ^{-1}E^T$ is invertible. So, by item e., the solution to the minimization problem is: $x = Q^{-1}E^T(EQ^{-1}E^T)^{-1}d$.

Using the advertized parameters we obtain, with $Q = \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.05 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}$, $E = \begin{bmatrix} 102 & 105 & 115 \\ 1 & 1 & 1 \end{bmatrix}$ and $d = \begin{bmatrix} 105 \\ 1 \end{bmatrix}$, $x = [0.5778\ 0.2489\ 0.1733]^T$. So, approximately 58% of the capital should be invested in the risk free asset, 25% into bonds and 17% into stocks in order to realize an expected return of 105 with minimal risk.

g. If $\hat{x}$ solves (3.2) it follows directly from this equation that $(A\hat{x})^TQ(b - A\hat{x}) = \hat{x}^TA^TQb - \hat{x}^TA^TQA\hat{x} = 0$.

If $Q = I$ this implies that $\hat{x}$ is such that the vector $A\hat{x}$ is perpendicular to the vector $b - A\hat{x}$. Or, stated differently, $A\hat{x}$ is the orthogonal projection of vector b onto the column space of matrix A. For an arbitrary Q, the result can be interpreted as that b is via a weighted orthogonal projection projected onto the column space of matrix A, where the weights are defined by matrix Q.

Problems with an Economic Content

EXERCISE 3.28 *(optimal production by a producer)*

Consider a producer who wants to minimize his spending on production cost, which consist of labor (L) and capital (K). Assume his cost per unit of labor and capital is ω and r, respectively. His production function is given by the Cob-Douglas function $P(K,L) = \alpha K^{\gamma} L^{1-\gamma}$ ($\alpha > 0, \gamma \in (0,1), K > 0,\ L > 0$).

a. Assume the producer wants to produce an amount of goods G. Calculate the amount of labor and capital he should use if he wants to minimize his production cost.
b. What will be approximately the additional cost if the producer wants to increase his production by one unit.
c. Determine the production of goods at the stationary point of the problem where the producer wants to maximize his profit, wages are flexible and the price he gets on the market for his good is $p = p_0 + \beta\omega - \delta G$.
d. Formulate an additional necessary condition for production to occur in this model. What is remarkable about the optimal wage set by the producer? Does this make sense?
e. Consider the second-order condition for the stationary point and conclude.
f. Calculate the optimum for the profit maximization problem considered in item c, if wages are fixed.
g. What happens with the maximal profit if the producer increases wages?

ANSWER 3.28

a. The producer's production cost are given by $C(K,L) := \omega L + rK$. Thus, the producer must solve the constrained optimization problem of minimizing $C(K,L)$ subject to $P(K,L) = G$.
   Note that C is continuous on its domain. Following Algorithm 3.13 we get:
   Step 1. $DC = [r\ \omega]$ and $DP = [\alpha\gamma K^{\gamma-1} L^{1-\gamma}\ \alpha(1-\gamma)L^{-\gamma}]$. Note all partial derivatives are continuous $\to C, P \in C^1$.
   Step 2. Note, as all variables differ from zero, DP has always full row rank.
   Step 3. Therefore all candidate solutions are stationary points of the associated Lagrange function

$$L(K,L,\lambda) := \omega L + rK - \lambda(\alpha K^{\gamma} L^{1-\gamma} - G).$$

   Step 4. $DL = [r - \lambda\alpha\gamma K^{\gamma-1} L^{1-\gamma}\ \omega - \lambda\alpha(1-\gamma)K^{\gamma}L^{-\gamma}] \to$ necessary conditions:
   i) $r - \lambda\alpha\gamma K^{\gamma-1} L^{1-\gamma} = 0$; ii) $\omega - \lambda\alpha(1-\gamma)K^{\gamma}L^{-\gamma} = 0$; and iii) $P = G$.
   From i) it follows that $\lambda \neq 0$ and $\lambda\alpha\gamma K^{\gamma-1} L^{1-\gamma} = r$.
   Similarly we have from ii) that $\lambda\alpha(1-\gamma)K^{\gamma}L^{-\gamma} = \omega$. Dividing both these equations yields:

$$\frac{\gamma K^{\gamma-1} L^{1-\gamma}}{(1-\gamma)K^{\gamma}L^{-\gamma}} = \frac{r}{\omega}, \text{ or } \frac{\gamma L}{(1-\gamma)K} = \frac{r}{\omega}, \text{ or } L = \frac{r(1-\gamma)}{\omega\gamma}K.$$

Substitution of this relationship into the production function $P(K,L)=G$ yields then

$$\alpha K^{\gamma}(\frac{r(1-\gamma)}{\omega\gamma}K)^{1-\gamma}=G.$$

That is,

$$K=\frac{G}{\alpha}\left(\frac{r(1-\gamma)}{\omega\gamma}\right)^{\gamma-1},\ \text{and}\ L=\frac{r(1-\gamma)}{\omega\gamma}K=\frac{G}{\alpha}\left(\frac{r(1-\gamma)}{\omega\gamma}\right)^{\gamma}.$$

To show that above found candidate solution yields indeed a global minimum, consider the Lagrangian $\tilde{L}(K,L):=\omega L+rK-\lambda^{*}(\alpha K^{\gamma}L^{1-\gamma}-G)$. One easily verifies that $H\tilde{L}(K,L)$ is positive semi-definite on the open convex set $K>0,\ L>0$. Therefore, see Exercise 3.24, the stationary point provides a global minimum. So, the minimal production cost for the producer are:

$$C=\omega\frac{G}{\alpha}\left(\frac{r(1-\gamma)}{\omega\gamma}\right)^{\gamma}+r\frac{G}{\alpha}\left(\frac{r(1-\gamma)}{\omega\gamma}\right)^{\gamma-1}.$$

b. The additional cost by increasing production by one unit is approximately $\lambda^{*}(K^{*},L^{*})$. From item a, e.g. equality ii), it follows (dropping for notational convenience the $'*'$ everywhere) that

$$\lambda=\frac{\omega}{\alpha(1-\gamma)}\left(\frac{L}{K}\right)^{\gamma}=\frac{\omega}{\alpha(1-\gamma)}\left(\frac{r(1-\gamma)}{\omega\gamma}\right)^{\gamma}=\frac{1}{\alpha}\left(\frac{r}{\gamma}\right)^{\gamma}\left(\frac{\omega}{1-\gamma}\right)^{1-\gamma}.$$

Notice one can also obtain this value by a direct evaluation of $C^{'}(G)$.

c. The profit function of the producer is:

$$\pi(K,L,\omega)=pG-\omega L-rK=(p_0+\beta\omega-\delta G)G-\omega L-rK,$$

where $G=\alpha K^{\gamma}L^{1-\gamma}$.

The first-order conditions are:

i) $D_1\pi=(p_0+\beta\omega)D_1G-2\delta GD_1G-r=0$,

ii) $D_2\pi=(p_0+\beta\omega)D_2G-2\delta GD_2G-\omega=0$,

iii) $D_3\pi=\beta G-L=0$.

Straightforward elaboration of i)-iii) yields:

i) $(p_0+\beta\omega-2\delta G)\left(\frac{K}{L}\right)^{\gamma-1}=\frac{r}{\alpha\gamma}$,

ii) $(p_0+\beta\omega-2\delta G)\left(\frac{K}{L}\right)^{\gamma}=\frac{\omega}{\alpha(1-\gamma)}$, and

iii) $\left(\frac{K}{L}\right)^{\gamma}=\frac{1}{\alpha\beta}$.

Dividing ii) by i) yields: $\frac{K}{L}=\frac{\omega\gamma}{r(1-\gamma)}$. Substitution of iii) into this equality gives:

$$iii')\ \omega=\frac{r(1-\gamma)}{\gamma}\left(\frac{1}{\alpha\beta}\right)^{1/\gamma}.$$

Substitution of iii) into ii) gives after some rewriting: $2\delta G = p_0 - \beta\frac{\gamma}{1-\gamma}\omega$. Since $L = \beta G$, it follows that $L = \frac{\beta}{2\delta}(p_0 - \beta\frac{\gamma}{1-\gamma}\omega)$. As (by iii)) $K = \left(\frac{1}{\alpha\beta}\right)^{1/\gamma} L$, we get

$$ii') \qquad L = \frac{\beta}{2\delta}(p_0 - \beta r\left(\frac{1}{\alpha\beta}\right)^{1/\gamma}) \text{ and}$$

$$i') \qquad K = \frac{\beta}{2\delta}(p_0 - \beta r\left(\frac{1}{\alpha\beta}\right)^{1/\gamma})\left(\frac{1}{\alpha\beta}\right)^{1/\gamma}.$$

So, the optimal production is $G = \frac{1}{2\delta}(p_0 - \beta r\left(\frac{1}{\alpha\beta}\right)^{1/\gamma})$ and (after some elementary calculations) the maximal profit is δG^2.

d. Note that production will only take place if $p_0 \geq \beta r\left(\frac{1}{\alpha\beta}\right)^{1/\gamma}$.
The optimal wage in this model does not depend on the parameters p_0 and δ. So, apparently in this model, the extent consumers react on the amount of the product brought to the market does not impact their wage.

e. Some elementary calculations show that the second-order derivative of π has next structure

$$\pi" = \begin{bmatrix} a & b & \beta D_1 G \\ b & c & d \\ \beta D_1 G & d & 0 \end{bmatrix}.$$

From this one easily verifies that with $x := [1\ 0\ s]^T$, s large, $x^T\pi"x = a + 2s\beta D_1 G > 0$ and with $y := [1\ 0\ -s]$, $y^T\pi"y = a - 2s\beta D_1 G < 0$. So, the stationary point is a saddle point.

f. Following item c, we reobtain next first-order conditions:
i) $D_1\pi = (p_0 + \beta\omega)D_1 G - 2\delta G D_1 G - r = 0$,
ii) $D_2\pi = (p_0 + \beta\omega)D_2 G - 2\delta G D_2 G - \omega = 0$,
or
i) $(p_0 + \beta\omega - 2\delta G)\left(\frac{K}{L}\right)^{\gamma-1} = \frac{r}{\alpha\gamma}$ and
ii) $(p_0 + \beta\omega - 2\delta G)\left(\frac{K}{L}\right)^{\gamma} = \frac{\omega}{\alpha(1-\gamma)}$.
Dividing ii) by i) yields: $\frac{K}{L} = \frac{\omega\gamma}{r(1-\gamma)}$. Let $s := \left(\frac{\omega\gamma}{r(1-\gamma)}\right)^{\gamma}$.
Notice $G = \alpha L\left(\frac{K}{L}\right)^{\gamma}$. Substitution of this into, e.g., ii) yields:

$$2\delta\alpha s^2 L \quad = \quad (p_0 + \beta\omega)s - \frac{\omega}{\alpha(1-\gamma)}, \text{ or, } L = \frac{(p_0 + \beta\omega)s - \frac{\omega}{\alpha(1-\gamma)}}{2\delta\alpha s^2}.$$

So the stationary point is:

$$(K_S, L_S) = (\frac{\omega\gamma}{r(1-\gamma)}\frac{(p_0 + \beta\omega)s - \frac{\omega}{\alpha(1-\gamma)}}{2\delta\alpha s^2}, \frac{(p_0 + \beta\omega)s - \frac{\omega}{\alpha(1-\gamma)}}{2\delta\alpha s^2}).$$

And

$$\begin{aligned}\pi(K_S, L_S) &= (p_0 + \beta\omega - \delta G)G - \omega L_S - rK_S \\ &= \delta G^2 + (p_0 + \beta\omega - 2\delta G)G - \omega L_S - rK_S \\ &= \delta G^2 + \frac{\omega}{\alpha(1-\gamma)}\alpha L_S - \omega L_S - rK_S \\ &= \delta G^2 + \frac{\omega}{(1-\gamma)}L_S - \omega L_S - \frac{\omega\gamma}{(1-\gamma)}L_S \\ &= \delta G^2.\end{aligned}$$

Elementary calculations show that

$$\pi^{"} = \begin{bmatrix} (p_0 + \beta\omega - 2\delta G)D_{11}G - 2\delta(D_1G)^2 & (p_0 + \beta\omega - 2\delta G)D_{21}G - 2\delta D_1GD_2G \\ (p_0 + \beta\omega - 2\delta G)D_{21}G - 2\delta D_1GD_2G & (p_0 + \beta\omega - 2\delta G)D_{22}G - 2\delta(D_2G)^2 \end{bmatrix}.$$

Note that $D_1G = \frac{\gamma G}{K}$, $D_2G = \frac{(1-\gamma)G}{L}$, $D_{11}G = \frac{-\gamma(1-\gamma)G}{K^2}$, $D_{12}G = \frac{\gamma(1-\gamma)G}{KL}$ and $D_{22}G = \frac{-\gamma(1-\gamma)G}{L^2}$ and $t := p_0 + \beta\omega - 2\delta G = \frac{rK}{\alpha\gamma G}$.
It then follows that

$$(p_0 + \beta\omega - 2\delta G)D_{11}G - 2\delta(D_1G)^2 = -\frac{r(1-\gamma)}{\alpha K} - \frac{2\delta\gamma^2G^2}{K^2} < 0,$$

and similarly,

$$(p_0 + \beta\omega - 2\delta G)D_{22}G - 2\delta(D_2G)^2 < 0.$$

The determinant of $\pi^{"}$ is given (with slight abuse of notation and omitting arguments) by

$$t(tD_{11}D_{22} - 2\delta(D_1^2D_{22} + D_2^2D_{11}) - tD_{12}^2 + 4\delta D_{21}D_1D_2).$$

Substituting the expressions for the partial derivatives yields

$$d = \frac{t}{K^2L^2}(2\delta(\gamma^3(1-\gamma)G^3 + (1-\gamma)^3\gamma G^3) + 4\delta\gamma^2(1-\gamma)^2G^3) > 0.$$

This shows that profit attains a maximum at the stationary point.

g From item f we have that maximum profit is given by δG^2, where

$$\begin{aligned}G &= \alpha(\frac{K}{L})^\gamma L = \alpha sL = \frac{\alpha s}{\alpha(1-\gamma)}\frac{(p_0 + \beta\omega)\alpha(1-\gamma)s - \omega}{2\delta\alpha s^2} \\ &= \frac{p_0 + \beta\omega}{2\delta} - \frac{\omega}{2\alpha(1-\gamma)\delta s} = \frac{p_0 + \beta\omega}{2\delta} - (\frac{r^(1-\gamma)}{\gamma})^\gamma\frac{1}{2\alpha(1-\gamma)\delta}\omega^{1-\gamma}.\end{aligned}$$

So in this model there is an incentive for producers to have wages as large as possible. If wages are increased, the producers will increase their production by increasing the amount of used capital and labor, which results in a higher total profit for the producer.

EXERCISE 3.29 *(optimal stock acquisition)*

Consider a firm that sells two similar products A and B (e.g. iphone's). The amount he can sell from product A, q_A, and B, q_B, is given by, respectively:

$$\begin{aligned} q_A &= -p_A + p_B + 100, \text{ and} \\ q_B &= -2p_B + \frac{3}{2}p_A + 150. \end{aligned}$$

Here p_A and p_B are the prices he asks for product A and B, respectively. The firm has a budget b to spend. The buying-in price for product A is p_1 and for product B is p_2.

a. Assume the firm wants to maximize its sales. Calculate the prices the firm should ask and its optimal stock.
b. For which buying-in prices does the firm not make any profit?
c. Consider the case that the firm makes profit. Under which conditions of the buying-in prices should the firm try to increase its budget?

ANSWER 3.29

a. The optimization problem reads as follows:
   find p_A and p_B that maximize $p_A q_A + p_B q_B$ subject to $p_1 q_A + p_2 q_B = b$.
   Following Algorithm 3.13, consider

$$L(p_A, p_B, \lambda) := p_A(-p_A + p_B + 100) + p_B(-2p_B + \frac{3}{2}p_A + 150) - \lambda(p_1 q_A + p_2 q_B - b).$$

   Differentiation of L yields next first-order conditions:

$$\begin{aligned} -2p_A + \frac{5}{2}p_B + 100 - \lambda p_1 &= 0 \\ \frac{5}{2}p_A - 4p_B + 150 - \lambda p_2 &= 0 \\ p_1 q_A + p_2 q_B &= b. \end{aligned}$$

   Notice that for all λ, L is concave on $\mathbb{R}^2$. So, according Exercise 3.24, any solution of the first-order conditions yields a maximum.
   Eliminating λ from the first two equations above yields next reduced system in p_A and p_B:

$$\begin{aligned} (2p_2 + \frac{5}{2}p_1)p_A - (\frac{5}{2}p_2 + 4p_1)p_B &= 100p_2 - 150p_1 \\ (\frac{3}{2}p_2 - p_1)p_A + (p_1 - 2p_2)p_B &= b - 100p_1 - 150p_2. \end{aligned}$$

That is,

$$\begin{aligned}\begin{bmatrix} p_A \\ p_B \end{bmatrix} &= -\begin{bmatrix} 2p_2 + \frac{5}{2}p_1 & -(\frac{5}{2}p_2 + 4p_1) \\ \frac{3}{2}p_2 - p_1 & p_1 - 2p_2 \end{bmatrix}^{-1} \begin{bmatrix} 100p_2 - 150p_1 \\ b - 100p_1 - 150p_2 \end{bmatrix} \\ &= \frac{4}{(p_1-p_2)^2 + 5p_1^2} \begin{bmatrix} p_1 - 2p_2 & \frac{5}{2}p_2 + 4p_1 \\ p_1 - \frac{3}{2}p_2 & 2p_2 + \frac{5}{2}p_1 \end{bmatrix} \begin{bmatrix} 100p_2 - 150p_1 \\ b - 100p_1 - 150p_2 \end{bmatrix} \\ &= \frac{4}{(p_1-p_2)^2 + 5p_1^2} \begin{bmatrix} (\frac{5}{2}p_2 + 4p_1)b - 450p_1p_2 - 550p_1^2 - 575p_2^2 \\ (\frac{5}{2}p_1 + 2p_2)b - 250p_1p_2 - 400p_1^2 - 450p_2^2 \end{bmatrix}.\end{aligned}$$

The optimal stock for the firm follows then directly by substitution of above relations into the equations $Q_A = -p_A + p_B + 100$ and $Q_B = -2p_B + \frac{3}{2}p_A + 150$.

b. The profit equals $\Pi(b, p_1, p_2) := p_A q_A + p_B q_B - b$, with p_A and q_A as advertized above. So, for a fixed budget b, the firm will not make profit if p_1 and p_2 are such that $\Pi(b, p_1, p_2) < 0$.

c. From the first-order conditions we have in particular that $\lambda = \frac{2p_A - \frac{5}{2}p_B - 100}{p_1}$. So, in case profit is positive and $\lambda > 0$, an increase in its budget will increase profit. So, if

$$-2p_A + \frac{5}{2}p_B + 100 = -\frac{7}{4}p_1 b + 275p_1p_2 + 100p_1^2 + 25p_2^2 + 25((p_1 - p_2)^2 + 5p_1^2) > 0,$$

the firm can increase profits by increasing its budget.

EXERCISE 3.30 *(derivation of the demand function from a utility function)*
Consider a twice differentiable concave utility function $U(x_1, x_2, x_3)$ that is a strictly increasing function for $x_i > 0$, where x_i is the quantity consumed by the consumer of good i. Assume the consumer has a total budget of $m > 0$ to spend and the price of good i is $p_i > 0$.

a. Show that the consumer's utility is maximal if the marginal rate of substitutions, $\frac{U'_{x_i}}{U'_{x_j}}$, equal the corresponding price relatives.

b. Consider $U(x_1, x_2, x_3) = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}$, with $\alpha_i < 1$. Show that in that case the demand for good i, $d_i(p_1, p_2, p_3) = \frac{\alpha_i}{\sum_{i=1}^3 \alpha_i} \frac{m}{p_i}$.

c. Calculate by how much maximum utility will change if the budget changes by Δm.

d. Let $v(p, m) := U(d_1(p, m), d_2(p, m), d_3(p, m))$. Show that $d_i(p, m) = -\frac{v_{p_i}(p,m)}{v_m(p,m)}$ (*Roy's identity*).

ANSWER 3.30 Denote by S the (open, convex) domain of U, $\{x \in \mathbb{R}^3 \mid x_i > 0\}$.

a. With $x := [x_1\ x_2\ x_3]^T$ and $p := [p_1\ p_2\ p_3]^T$ the maximization problem reads as follows: $\max_x U(x)$ s.t. $p^T x = m$. Following Lagrange's algorithm, let

$$L(x, \lambda) := U(x) - \lambda(p^T x - m).$$

The first-order conditions yield then: $U^{'}_{x_i} - \lambda p_i = 0,\ i = 1, 2, 3.$

As $U(x)$ is concave on S it follows (see Exercise 3.24) that a stationary point yields a maximum location. From which we directly deduce that at the maximum location $\frac{U^{'}_{x_i}}{U^{'}_{x_j}} = \frac{p_i}{p_j}$.

b. Elementary differentiation shows $U^{'}_{x_1} = \alpha_1 x_1^{\alpha_1 - 1} x_2^{\alpha_2} x_3^{\alpha_3}$, whereas similar expressions are obtained for $U^{'}_{x_2}$ and $U^{'}_{x_3}$. Hence,

$$\frac{U^{'}_{x_i}}{U^{'}_{x_j}} = \frac{\alpha_i x_j}{\alpha_j x_i}.$$

Not that U is concave on its domain. Therefore, the Lagrange function $\tilde{L}(x) = L(x, \lambda^*)$, where λ^* corresponds by the stationary point, is also concave on S. By Exercise 3.24, the stationary point thus yields a maximum.

From item a., it follows that at this stationary point,

$$\frac{\alpha_i x_j}{\alpha_j x_i} = \frac{p_i}{p_j}, \text{ and } p^T x = m.$$

The first condition implies that $p_j x_j = \frac{\alpha_j}{\alpha_i} p_i x_i$. Substituting these relationships into the budget-constraint gives

$$p_1 x_1 + p_2 x_2 + p_3 x_3 = p_1 x_1 + \frac{\alpha_2}{\alpha_1} p_1 x_1 + \frac{\alpha_3}{\alpha_1} p_1 x_1 = m.$$

Hence

$$x_1 = \frac{m \alpha_1}{p_1 \sum_{i=1}^{3} \alpha_i}.$$

Similarly, one obtains the demand functions for x_2 and x_3.

c. The change in utility will approximately be $\lambda * \Delta m$. From item a. we have $\lambda = \frac{U^{'}_{x_i}}{p_i}$. Substitution of x_i from item b. into $U^{'}_{x_i}$ yields after some rewriting: $\lambda = \left(\frac{\sum_{i=1}^{3} \alpha_i}{m}\right)^{1 - \sum_{i=1}^{3} \alpha_i} \Pi_{i=1}^{3} \left(\frac{\alpha_i}{p_i}\right)^{\alpha_i}$.

d. Note that

$$v^{'}_{p_j} = \sum_{i=1}^{3} U^{'}_{x_i} \frac{\partial d_i}{\partial p_j}$$

From item a., we have $U^{'}_{x_i} = \lambda p_i$, where, by Lagrange's theorem, $\lambda = v^{'}_{m}$. Hence, from above equality we obtain

$$v^{'}_{p_j} = \sum_{i=1}^{3} \lambda p_i \frac{\partial d_i}{\partial p_j} = \lambda \sum_{i=1}^{3} p_i \frac{\partial d_i}{\partial p_j} = v^{'}_{m} \sum_{i=1}^{3} p_i \frac{\partial d_i}{\partial p_j} \quad (\mathbf{1}).$$

On the other hand, the budgetconstraint implies that $\sum_{i=1}^{3} p_i d_i(p, m) = m$.

Differentiating this constraint with respect to p_j gives

$$d_j(p, m) + \sum_{i=1}^{3} p_i \frac{\partial d_i(p, m)}{\partial p_j} = 0.$$

That is, $\sum_{i=1}^{3} p_i \frac{\partial d_i(p,m)}{\partial p_j} = -d_j(p, m)$. Substitution of this into (**1.**) yields then Roy's identity.

EXERCISE 3.31 *(Engel curves and Income consumption curves)*
Consider a consumer who wants to maximize his utility function $U(x,y,z) \in \mathbb{C}^2$ given his budget I and prices p_i for good $i,\ i=1,2,3$, from which he consumes a positive amount of x, y and z, respectively.

a. Use the Lagrange approach to determine the first-order conditions for this maximization problem.
b. Consider $U(x,y,z) = x^{\alpha_1}y^{\alpha_2}z^{\alpha_3}$, where $0<\alpha_i<1,\ i=1,2,3$. Determine the optimal amount of consumption of the goods x, y and z as an explicit function of the budget. This are the so-called *Engel curves*.
c. Show that in the optimal situation, the optimal amount of consumption of product 1, x, is linearly related to both the optimal amount of consumption of product 2, y, and product 3, z, respectively. The functions $x(y)$ and $x(z)$ are called *income consumption curves*.

ANSWER 3.31

a. the maximization problem reads as follows: $\max_{x,y,z} U(x,y,z)$ s.t. $p_1x+p_2y+p_3z=I$. Following Algorithm 3.13, let

$$L(x,y,z,\lambda) := U(x,y,z) - \lambda(p_1x+p_2y+p_3z-I).$$

The first-order conditions yield then: $U_x^{'}-\lambda p_1=0$, $U_y^{'}-\lambda p_2=0$ and $U_z^{'}-\lambda p_3=0$, together with the budget constraint. Or, by eliminating λ, the following equivalent equations are obtained:

$$\frac{U_x^{'}}{U_y^{'}}=\frac{p_1}{p_2},\ \frac{U_x^{'}}{U_z^{'}}=\frac{p_1}{p_3},\ \text{and } p_1x+p_2y+p_3z=I.$$

b. With $U(x,y,z)=x^{\alpha_1}y^{\alpha_2}z^{\alpha_3}$, above equations reduce to the system of equations:

$$\frac{\alpha_1}{\alpha_2}\frac{y}{x}=\frac{p_1}{p_2},\ \frac{\alpha_1}{\alpha_3}\frac{z}{x}=\frac{p_1}{p_3},\ \text{and } p_1x+p_2y+p_3z=I.$$

Recall that U is concave on its domain S. So, for every λ, the Lagrange function is also concave on S. By Exercise 3.24, any stationary point thus yields a global maximum.
Moreover, the first two equations imply that $\alpha_1p_2y=\alpha_2p_1x$ and $\alpha_1p_3z=\alpha_3p_1x$, respectively. Substituting these into the budget equation gives

$$x=\frac{\alpha_1}{\sum_{i=1}^3\alpha_i}\frac{1}{p_1}I \text{ (i)}.$$

Similarly, using $\frac{\alpha_2}{\alpha_3}\frac{z}{y}=\frac{p_2}{p_3}$, we obtain

$$y=\frac{\alpha_2}{\sum_{i=1}^3\alpha_i}\frac{1}{p_2}I \text{ (ii)},\ \text{and } z=\frac{\alpha_3}{\sum_{i=1}^3\alpha_i}\frac{1}{p_3}I \text{ (iii)}.$$

c. From (i) we have $I = \frac{\sum_{i=1}^{3}\alpha_i}{\alpha_1}p_1x$ and from (ii) $I = \frac{\sum_{i=1}^{3}\alpha_i}{\alpha_2}p_2y$. From which we obtain that the income consumption curve of x w.r.t. y is given by: $x = \frac{\alpha_1}{\alpha_2}\frac{p_2}{p_1}y$. Similarly we obtain from (i) and (iii) that the income consumption curve of x w.r.t. z is given by: $x = \frac{\alpha_1}{\alpha_3}\frac{p_3}{p_1}z$.

EXERCISE 3.32 *(Deriving the production function for a producer)*
Consider a producer who wants to minimize his spending on production cost, which consist of labor (L) and capital (K). Assume his cost per unit of labor and capital is ω and r, respectively. His production function is given by a concave function $P(K, L)$.

a. Assume the producer wants to produce an amount of goods y. Show that the amount of labor and capital he should use if he wants to minimize his production cost is determined by the two equations:
$$P(K, L) = y, \text{ and } \frac{P_L^{'}}{P_K^{'}} = \frac{\omega}{r}.$$

b. Consider $P(K, L) = \mu(\alpha K^{\rho} + (1-\alpha)L^{\rho})^{\frac{\beta}{\rho}}$, where $0 < \alpha < 1$, $\mu > 0$, $\beta > 0$ and $\rho < 1$ (a production function that has constant elasticity of substitution, a so-called CES-production function). Show that the with this production function corresponding long run cost function is: $C(y) = \theta y^{\frac{1}{\rho}}$, for some constant θ.

c. Show, using the second-order criterion that $\det HL < 0$, that the solution obtained in item b indeed yields a minimum.

d. Assume the producer will be a monopolist in the market and the price of his product is given by $p(y) = \frac{c}{\sqrt{y}}$ and $0 < \rho < 1$. Calculate the amount of labor and capital the producer should invest to maximize his profit.

ANSWER 3.32

a. the minimization problem reads as follows: $\min_{K,L} \omega L + rK$ s.t. $P(K, L) = y$. Following Algorithm 3.13, let

$$L(K, L, \lambda) := \omega L + rK - \lambda(P(K, L) - y).$$

The first-order conditions yield then: $r - \lambda P_K^{'} = 0$, $\omega - \lambda P_L^{'} = 0$ together with the production constraint $P(K, L) = y$. Or by elimination of λ, equivalently, next equations:

$$\frac{P_L^{'}}{P_K^{'}} = \frac{\omega}{r}, \text{ and } P(K, L) = y.$$

b. With $P(K, L) = \mu(\alpha K^{\rho} + (1-\alpha)L^{\rho})^{\frac{\beta}{\rho}}$, $P_L^{'} = (1-\alpha)L^{\rho-1}\rho\frac{\beta}{\rho}\mu(\alpha K^{\rho} + (1-\alpha)L^{\rho})^{\frac{\beta}{\rho}-1}$ and $P_K^{'} = \alpha K^{\rho-1}\rho\frac{\beta}{\rho}\mu(\alpha K^{\rho} + (1-\alpha)L^{\rho})^{\frac{\beta}{\rho}-1}$. Therefore above equations reduce to the system of equations:

$$\frac{1-\alpha}{\alpha}\left(\frac{L}{K}\right)^{\rho-1} = \frac{\omega}{r}, \text{ and } \mu(\alpha K^{\rho} + (1-\alpha)L^{\rho})^{\frac{\beta}{\rho}} = y.$$

From the first equation we have: $L = \left(\frac{\alpha\omega}{(1-\alpha)r}\right)^{\frac{1}{\rho-1}} K$. Substitution of this into the second equation shows: $[\alpha + (1-\alpha)\left(\frac{\alpha\omega}{(1-\alpha)r}\right)^{\frac{\rho}{\rho-1}}]K^\rho = y$. Or, $K = [\alpha + (1-\alpha)\left(\frac{\alpha\omega}{(1-\alpha)r}\right)^{\frac{\rho}{\rho-1}}]^{\frac{-1}{\rho}} y^{\frac{1}{\rho}}$. So, the long run cost for the producer are

$$C(y) = \omega L + rK = [\omega\left(\frac{\alpha\omega}{(1-\alpha)r}\right)^{\frac{1}{\rho-1}} + r][\alpha + (1-\alpha)\left(\frac{\alpha\omega}{(1-\alpha)r}\right)^{\frac{\rho}{\rho-1}}]^{\frac{-1}{\rho}} y^{\frac{1}{\rho}} =: \theta y^{\frac{1}{\rho}}.$$

c. Notice that $HL = \begin{bmatrix} -\lambda P_{KK} & -\lambda P_{KL} & -P_K \\ -\lambda P_{LK} & -\lambda P_{LL} & -P_L \\ -P_K & -P_L & 0 \end{bmatrix}$, where with $S := \beta\mu(\alpha K^\rho + (1-\alpha)L^\rho)^{\frac{\beta}{\rho}-1}$,

$$\begin{aligned} P_K &= \alpha K^{\rho-1}S; P_L = (1-\alpha)L^{\rho-1}S,\ P_{KK} = \alpha(\rho-1)K^{\rho-2}S + \alpha K^{\rho-1}S_K, \\ P_{LL} &= (1-\alpha)(\rho-1)L^{\rho-2}S + (1-\alpha)L^{\rho-1}S_L, \text{ and } P_{KL} = \alpha K^{\rho-1}S_L; P_{LK} = (1-\alpha)L^{\rho-1}S_K. \end{aligned}$$

In particular we obtain from this that at the point where the first-order conditions apply (denoted with an asterisk)

$$\begin{aligned} \lambda^* P^*_{KK} &= \alpha(\rho-1){K^*}^{-1}\lambda^*{K^*}^{\rho-1}S + \alpha\lambda^*{K^*}^{\rho-1}S^*_K = (\rho-1)r{K^*}^{-1} + \alpha\lambda^*{K^*}^{\rho-1}S^*_K \\ \lambda^* P^*_{LL} &= (1-\alpha)(\rho-1){L^*}^{-1}\lambda^*{L^*}^{\rho-1}S + (1-\alpha)\lambda^*{L^*}^{\rho-1}S^*_L = (\rho-1)\omega{L^*}^{-1} + (1-\alpha)\lambda^*{L^*}^{\rho-1}S^*_L. \end{aligned}$$

Notice, using the first-order conditions and the fact that $\lambda \neq 0$, that:

$$\begin{aligned} \det HL^* &= -P^*_K(\lambda^* P^*_{LK}P^*_L - \lambda^* P^*_{LL}P^*_K + P^*_L(\lambda^* P^*_{KK}P^*_L - \lambda^* P^*_{KL}P^*_K) \\ &= -\lambda^* P^*_K(P^*_{LK}P^*_L - P^*_{LL}P^*_K) + \lambda^* P^*_L(P^*_{KK}P^*_L - P^*_{KL}P^*_K) \\ &= -r(P^*_{LK}P^*_L - P^*_{LL}P^*_K) + \omega(P^*_{KK}P^*_L - P^*_{KL}P^*_K) \\ &= \frac{1}{\lambda^*}(-r(\omega P^*_{LK} - rP^*_{LL}) + \omega(\omega P^*_{KK} - rP^*_{KL})). \end{aligned}$$

Using next the above derived expressions for $\lambda^* P^*_{KK}$ and $\lambda^* P^*_{LL}$ respectively we obtain from the last equation that

$$\begin{aligned} \det HL^* &= \frac{1}{\lambda^{*2}}(-r(\omega\lambda^* P^*_{LK} - r\lambda^* P^*_{LL}) + \omega(\omega\lambda^* P^*_{KK} - r\lambda^* P^*_{KL})) \\ &= \frac{1}{\lambda^{*2}}(r^2\omega(\rho-1){L^*}^{-1} + r^2(1-\alpha)\lambda^*{L^*}^{\rho-1}S^*_L - \omega r\alpha\lambda^*{K^*}^{\rho-1}S^*_L - \omega r(1-\alpha)\lambda^*{L^*}^{\rho-1}S^*_K \\ &+ \omega^2{K^*}^{-1}(\rho-1)r + \omega^2\alpha\lambda^*{K^*}^{\rho-1}S^*_K). \end{aligned}$$

Next observe that, again using the first-order conditions and the fact that $\frac{S_K}{S_L} = \frac{\alpha}{1-\alpha}$:

$$\begin{aligned} r^2(1-\alpha)\lambda^* L^{*^{\rho-1}} S_L^* - \omega r\alpha\lambda^* K^{*^{\rho-1}} S_L^* - \omega r(1-\alpha)\lambda^* L^{*^{\rho-1}} S_K^* + \omega^2\alpha\lambda^* K^{*^{\rho-1}} S_K^* = \\ r^2(1-\alpha)\lambda^* L^{*^{\rho-1}} S_L^* - \omega r\alpha\lambda^* K^{*^{\rho-1}} S_L^* - \omega r\alpha\lambda^* L^{*^{\rho-1}} S_L^* + \omega^2 \frac{\alpha^2}{1-\alpha}\lambda^* K^{*^{\rho-1}} S_L^* = \\ ((1-\alpha)r - \alpha\omega)\lambda^* S_L^* (rL^{*^{\rho-1}} - \frac{\alpha\omega}{1-\alpha} K^{*^{\rho-1}}) = 0. \end{aligned}$$

Consequently, $\det HL = \frac{1}{\lambda^{*^2}}(r^2\omega(\rho-1)L^{*^{-1}} + \omega^2 K^{*^{-1}}(\rho-1)r) < 0$. So, by Theorem 3.23, we can conclude that at the stationary point we have a minimum location.

d. Profit is given by

$$\Pi(y) = p(y)y - C(y) = c\sqrt{y} - \theta y^{\frac{1}{\rho}}.$$

Thus,

$$\Pi'(y) = \frac{c}{2\sqrt{y}} - \frac{\theta}{\rho} y^{\frac{1}{\rho}-1} \text{ and } \Pi"(y) = -\frac{c}{4} y^{-\frac{3}{2}} - \frac{\theta}{\rho}(\frac{1}{\rho} - 1) y^{\frac{1}{\rho}-2}.$$

From this we straightforwardly obtain that $y^* = (\frac{c}{2\theta})^{\frac{2\rho}{2-\rho}}$ yields maximum profit. Hence, provided profit is nonnegative at this point, the optimal investment in labor and capital are

$$K = [\alpha + (1-\alpha)\left(\frac{\alpha\omega}{(1-\alpha)r}\right)^{\frac{\rho}{\rho-1}}]^{\frac{-1}{\rho}} y^{*\frac{1}{\rho}}, \text{ and } L = \left(\frac{\alpha\omega}{(1-\alpha)r}\right)^{\frac{1}{\rho-1}} K.$$

EXERCISE 3.33 *(Income and substitution effect in optimal consumer choice if price of a good changes)* Consider a consumer who wants to spend his budget I on two goods. The price of good i is p_i, $i = 1, 2$. His utility from consumption of amounts x and y of the goods, respectively, is $U(x, y) \in C^2$.

a. Formulate the optimization problem the consumer has to solve if he wants to maximize his utility. Use the Lagrange approach to formulate the first-order conditions. Assume utility is maximized by a consumption of (x^*, y^*) of the goods.

b. Assume the price of the first good increases from p_1 to p_1^N. Assume utility is maximized in this case by a consumption of x^N and y^N of both goods, respectively.
Assume the utility function is given by $U(x, y) = xy$. Calculate x^N and y^N.

c. Next, some authorities want to compensate the consumer for his loss of utility by increasing income such that with these new prices he can still achieve his old utility $U^* := U(x^*, y^*)$. This results in the optimization problem $\max_{x,y} U(x, y)$ subject to the budget constraint $p_1^N x + p_2 y = I^N$ and $U(x, y) = U^*$.
Determine this drop in real income if the utility function is given by $U(x, y) = xy$ and the corresponding optimal consumption bundles (x^S, y^S). In literature the move from the bundle (x^*, y^*) to the bundle (x^H, y^H) is called the *(Hicks) substitution effect* and the move from the bundle (x^H, y^H) to the bundle (x^N, y^N) the *(Hicks) income effect*.

d. Other authorities want to compensate the consumer by increasing income such that he can purchase his old optimal bundle (x^*, y^*). With this new income, I^S, the consumer can in fact realize a higher utility. Calculate this corresponding maximum utility and consumption bundles (x^S, y^S) if $U(x, y) = xy$. In literature the move from the bundle (x^*, y^*) to the bundle (x^S, y^S) is called the *(Slutsky) substitution effect* and the move from the bundle (x^S, y^S) to the bundle (x^N, y^N) the *(Slutsky) income effect*.

ANSWER 3.33

a. The optimization problem is: $\max_{x,y} U(x, y)$ subject to the budget constraint $p_1x + p_2y = I$. So, with $L(x, y, \lambda) = U(x, y) - \lambda(p_1x + p_2y - I)$, the resulting first-order conditions are:

$$U_x^{'} - \lambda p_1 = 0,\ U_y^{'} - \lambda p_2 = 0,\ \text{and } p_1x + p_2y = I.$$

Which imply

$$\frac{U_x^{'}}{U_y^{'}} = \frac{p_1}{p_2} \text{ together with } p_1x + p_2y = I.$$

b. From above first-order conditions it follows with $U(x, y) = xy$:

$$\frac{y}{x} = \frac{p_1^N}{p_2},\ \text{and } p_1^Nx + p_2y = I.$$

From which we get $x^N = \frac{I}{2p_1^N}$ and $y^N = \frac{I}{2p_2}$.
Some elementary calculations show that

$$HL = \begin{bmatrix} 0 & 1 & -p_1^N \\ 1 & 0 & -p_2 \\ -p_1^N & -p_2 & 0 \end{bmatrix}.$$

The determinant of HL equals $2p_1^Np_2 > 0$. So we have a maximum at (x^N, y^N).

c. First note that, similarly as above, we have $(x^*, y^*) = (\frac{I}{2p_1}, \frac{I}{2p_2})$. The substitution effect is determined by solving the system

$$\frac{y}{x} = \frac{p_1^N}{p_2},\ p_1^Nx + p_2y = I^N,\ \text{and } xy = \frac{I^2}{4p_1p_2}.$$

This gives: $I^N = \sqrt{\frac{p_1^N}{p_1}}I$ and $(x^H, y^H) = (\frac{1}{2}\sqrt{\frac{1}{p_1p_1^N}}I, \frac{1}{2p_2}\sqrt{\frac{p_1^N}{p_1}}I)$.
That is, under new prices, with an income of I^N one could have realized the same utility as before.

d. $I^S := p_1^Nx^* + p_2y^* = \frac{1}{2}(\frac{p_1^N}{p_1} + 1)I$. Next consider the maximization of $U(x, y)$ subject to the budget constraint $p_1^Nx + p_2y = I^S$. Along the lines of item c, we obtain next first-order conditions:

$$\frac{y}{x} = \frac{p_1^N}{p_2},\ \text{and } p_1^Nx + p_2y = I^S.$$

Solving this system gives $(x^S, y^S) = (\frac{I^S}{2p_1^N}, \frac{I^S}{2p_2})$. Along the lines of item c (consider det HL), it follows that this point provides a maximum location too.

CHAPTER 4

Constrained Optimization: Inequality Constraints

4.1 The Directional Derivative

EXERCISE 4.1 Calculate in which directions the function f at the indicated point increases maximally, minimally and almost not, respectively.

a. $f(x,y) = 4xy$ at $(x,y) = (1,2)$.

b. $f(x,y) = x^2 + y^2$ at $(x,y) = (0,0)$.

c. $f(x,y,z) = x^2 - 2y^2 + z^2$ at $(x,y,z) = (2,1,2)$.

ANSWER 4.1

a. $Df(x,y) = [4y\ 4x]$. Therefore $Df(1,2) = [8\ 4]$. So f will increase maximal in the direction $[8\ 4]^T$, minimal in the direction $[-8\ -4]^T$, and almost not change in the directions $[-4\ 8]^T$ and $[4\ -8]^T$.

b. $Df(x,y) = [2x\ 2y]$. Therefore $Df(0,0) = [0\ 0]$. So, in all directions the value of f will almost not change.

c. $Df(x,y,z) = [2x\ -4y\ 2z]$. Therefore $Df(2,1,2) = [4\ -4\ 4]$. So f will increase maximal in the direction $[1\ -1\ 1]^T$, minimal in the direction $[-1\ 1\ -1]^T$, and almost not change in the directions perpendicular to $Df(2,1,2)$ that is all vectors pointing into the plane $x - y + z = 0$. That is

$$\{(x,y,z) \quad | \quad z = y - x,\ x,y \in \mathbb{R}\} =$$
$$\{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad | \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},\ \lambda_i \in \mathbb{R},\ i = 1,2\}.$$

4.2 Two variables and one inequality constraint

EXERCISE 4.2 Solve the next constrained optimization problems.

a. Maximize $f(x,y) = 4xy$ subject to the constraint $g(x,y) := x^2 + 2y^2 \leq 6$.

b. Maximize $f(x,y,z) = -x^2 - y^2 + 4x + 4y$, subject to the constraint $g(x,y) := x + y \leq 2$.

c. Minimize $f(x,y) = x^2 + y^2$, subject to the constraint $g(x,y) := p_x x + p_y y \leq 1$, where p_x and p_y are some positive numbers.

d. Minimize $f(x,y) = x + y$, subject to the constraint $g(x,y) := x^2 + y^2 \geq 8$.

ANSWER 4.2

a. To find all candidate solutions, we proceed along the lines of Algorithm 4.9.

Step 1. $Df = [4y \; 4x]$, $Dg = [2x \; 4y]$. So f, $g \in C^1$.

Step 2. Check for points where the implicit function theorem does not apply.

Consider $G := [2x \; 4y \; x^2 + 2y^2 - 6] = [0 \; 0 \; 0]$. There is no solution to this set of equations. So the Implicit Function Theorem can be applied to all the points in the feasible domain.

Step 3. Consider $\mathcal{L} = 4xy - \lambda(x^2 + 2y^2 - 6)$.

Step 4. The Kuhn-Tucker conditions are:

$$\begin{cases} [4y - 2\lambda x \; 4x - 4\lambda y] = [0 \; 0], \\ \lambda(x^2 + 2y^2 - 6) = 0, \\ \lambda \geq 0, \\ x^2 + 2y^2 - 6 \leq 0. \end{cases}$$

i) When the constraint is not binding, $x^2 + 2y^2 - 6 < 0$, $\lambda = 0$, and consequently $[4y - 2\lambda x \; 4x - 4\lambda y] = [0 \; 0]$ gives $(x,y) = (0,0)$. Obviously, $(x,y) = (0,0)$ satisfies $x^2 + 2y^2 - 6 \leq 0$. So this point satisfy the Kuhn-Tucker conditions and is a candidate point.

ii) When the constraint is binding, i.e. $x^2 + 2y^2 - 6 = 0$, in combination with $[4y - 2\lambda x \; 4x - 4\lambda y] = [0 \; 0]$ this gives $(x^2, y^2, \lambda^2) = (3, \frac{3}{2}, 2)$. Since $\lambda \geq 0$, $\lambda = \sqrt{2}$. Since $x = \lambda y$, x and y must have the same sign. There are two candidate points found in this step: $(\sqrt{3}, \frac{\sqrt{6}}{2}), (-\sqrt{3}, -\frac{\sqrt{6}}{2})$.

So there are in total three candidate points: $(0,0), (\sqrt{3}, \frac{\sqrt{6}}{2}), (-\sqrt{3}, -\frac{\sqrt{6}}{2})$.

Evaluation of f at these points gives:

$$f(0,0) = 0 \text{ and } f(\sqrt{3}, \frac{\sqrt{6}}{2}) = f(-\sqrt{3}, -\frac{\sqrt{6}}{2}) = 6\sqrt{2}.$$

Since there are only three candidate points and the constraint set is compact it follows from Weierstrass' theorem that $(\sqrt{3}, \frac{\sqrt{6}}{2}), (-\sqrt{3}, -\frac{\sqrt{6}}{2})$ are the two points where f attains its global maximum.

b. To find all candidate solutions, we proceed along the lines of Algorithm 4.9.

Step 1. $Df = [-2x + 4 \; -2y + 4]$, $Dg = [1 \; 1]$. So f, $g \in C^1$.

Step 2. Clearly Dg has full row rank everywhere. So the Implicit Function Theorem can be applied to all the points in the feasible domain.

Step 3. Consider $\mathcal{L} := -x^2 - y^2 + 4x + 4y - \lambda(x + y - 2)$.

Step 4. The Kuhn-Tucker conditions are:

$$\begin{cases} [-2x + 4 - \lambda \ -2y + 4 - \lambda] = [0\ 0], \\ \lambda(x + y - 2) = 0, \\ \lambda \geq 0, \\ x + y - 2 \leq 0. \end{cases}$$

i) When the constraint is not binding, i.e. $x+y < 2$ it follows that $\lambda = 0$. So $[-2x+4\ -2y+4] = [0\ 0]$, which gives $(x, y) = (2, 2)$. But $(x, y) = (2, 2)$ does not satisfy $x + y - 2 < 0$. So this point does not satisfy the Kuhn-Tucker conditions and is not a candidate point.

ii) When the constraint is binding, i.e. $x+y = 2$, together with $[-2x+4-\lambda\ -2y+4-\lambda] = [0\ 0]$ gives $(x, y, \lambda) = (1, 1, 2)$ as the only candidate solution.

So, there is only one candidate solution: $(x, y, \lambda) = (1, 1, 2)$.

Note that the second-order derivative of $L(x, y, 2) = -x^2 - y^2 + 2x + 2y + 4$ is $\begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$. So, $HL(x, y, 2) < 0$. Therefore $L(x, y, 2)$ is concave on $\mathbb{R}^2$. Consequently, at $(1, 1)$ the constrained optimization problem has a global maximum.

c. Consider the maximization of $f(x, y) := -x^2-y^2$ subject to the inequality constraint $p_x x+p_y y \leq 1$. To find all candidate solutions, we proceed along the lines of Algorithm 4.9.

Step 1. $Df = [-2x\ -2y]$, $Dg = [p_x\ p_y]$. So f, $g \in C^1$.

Step 2. Clearly Dg has full row rank everywhere. So the Implicit Function Theorem can be applied to all the points in the feasible domain.

Step 3. Consider $\mathcal{L} := -x^2 - y^2 - \lambda(p_x x + p_y y - 1)$.

Step 4. The Kuhn-Tucker conditions are:

$$\begin{cases} [-2x - \lambda p_x \ -2y - \lambda p_y] = [0\ 0], \\ \lambda(p_x x + p_y y - 1) = 0, \\ \lambda \geq 0, \\ p_x x + p_y y - 1 \leq 0. \end{cases}$$

i) When the constraint is not binding, i.e. $p_x x + p_y y < 1$ it follows that $\lambda = 0$.
So $[-2x - \lambda p_x \ -2y - \lambda p_y] = [0\ 0]$ gives $(x, y) = (0, 0)$. $(x, y) = (0, 0)$ also satisfies $p_x x + p_y y - 1 < 0$. So this point satisfies the Kuhn-Tucker conditions and is therefore a candidate point.

ii) When the constraint is binding, i.e. $x + y = 2$, together with $[-2x - \lambda p_x \ -2y - \lambda p_y] = [0\ 0]$ does not provide any solution (add for instance the last two equations and, next, substitute

into the resulting equation $x + y = 2$. Then, as $\lambda \geq 0$, the resulting equation does not have a solution).

So, the only candidate solution is $(x, y) = (0, 0)$.

Note that the second-order derivative of $L(x, y, 0) = -x^2 - y^2$ is $\begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$. So, $HL(x, y, 2) < 0$. Therefore $L(x, y, 0)$ is concave on $\mathbb{R}^2$. Consequently, at $(0, 0)$ the initially formulated constrained optimization problem has a global minimum.

d. Consider the maximization of $f(x, y) := -x - y$ subject to the inequality constraint $-x^2 - y^2 + 8 \leq 0$.

To find all candidate solutions, we proceed along the lines of Algorithm 4.9.

Step 1. $Df = [-1 \ -1]$, $Dg = [-2x \ -2y]$. So $f, g \in C^1$.

Step 2. The only point where Dg has not full row rank is $(0, 0)$. But $(0, 0)$ is not located in the constraint set. So the Implicit Function Theorem can be applied to all the points in the feasible domain.

Step 3. Consider $\mathcal{L} := -x - y - \lambda(-x^2 - y^2 + 8)$.

Step 4. The Kuhn-Tucker conditions are:

$$\begin{cases} [-1 + 2\lambda x \ \ -1 + 2\lambda y] = [0 \ 0], \\ \lambda(x^2 + y^2 - 8) = 0, \\ \lambda \geq 0, \\ x^2 + y^2 \geq 8. \end{cases}$$

i) When the constraint is not binding, i.e. $x^2 + y^2 > 8$ it follows that $\lambda = 0$. In this case $[-1 + 2\lambda x \ \ -1 + 2\lambda y] = [0 \ 0]$ has no solution.

ii) When the constraint is binding, i.e. $x^2 + y^2 = 8$, together with $[-1 + 2\lambda x \ \ -1 + 2\lambda y] = [0 \ 0]$, from the last 2 equations it follows that $x = y$. Substitution into $x^2 + y^2 = 8$ gives $(x, y) = (2, 2)$, with $\lambda = \frac{1}{4}$.

So, the only candidate solution is $(x, y) = (2, 2)$.

Note that the second-order derivative of $L(x, y, \frac{1}{4})$ is $\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$. So, $HL(x, y, \frac{1}{4}) > 0$. Therefore $L(x, y, \frac{1}{4})$ is convex on $\mathbb{R}^2$. Consequently, we cannot conclude from either Weierstrass' Theorem or Proposition 4.11 that the initially formulated problem has a global minimum. In fact, by choosing both x and y arbitrarily negative, we see that we can make $x + y$ arbitrarily negative while still satisfying the inequality constraint $x^2 + y^2 \geq 8$. That is the initially formulated minimization problem does not have a solution.

4.3 The general case

EXERCISE 4.3 Solve the next constrained optimization problems.

a. Maximize $f(x,y) = x + y$, subject to the constraints $x^2 + 2y^2 \leq 6$, $x \geq 0$ and $y \geq 0$.

b. Maximize $f(x,y) = -x^2 - xy - y^2$ subject to the constraints $g_1(x,y) = x - 2y \leq -1$; $g_2(x,y) = 2x + y \leq 2$; $x \geq 0$ and $y \geq 0$.

c. Maximize $f(x,y) = y^2 - x^2$, subject to the constraints $g_1(x,y) = x^2 + y^2 \leq 4$ and $g_2(x,y) = x^2 + (y-2)^2 \leq 4$.

d. Minimize $f(x,y) = x + xy$, subject to the constraints $y \leq 2 - x^2$, $x \geq 0$ and $y \geq 0$.

e. Maximize $f(x,y) = x + y$, subject to the constraint $g(x,y) = y \leq -\frac{1}{2}x^2$.

f. Maximize $f(x,y) = x - \frac{1}{3}xy$, subject to the constraints $y \geq x^2$ and $x \leq 2$.

ANSWER 4.3

a. The feasible domain is compact. So, by Weierstrass' Theorem, the global maximum and global minimum exist. To find all candidate solutions, we proceed along the lines of Algorithm 4.20.

Step 1. $Df = [1\ 1]$, $Dg = \begin{bmatrix} 2x & 4y \\ -1 & 0 \\ 0 & -1 \end{bmatrix}$. Therefore, $f, g \in C^1$.

Step 2. Check invertibility/full row rank

$$G = \begin{bmatrix} 2x & 4y & x^2 + 2y^2 - 6 & 0 & 0 \\ -1 & 0 & 0 & -x & 0 \\ 0 & -1 & 0 & 0 & -y \end{bmatrix}.$$

By construction, G has full row rank at the interior points. So we only need to check the condition at the boundary points.

1. First consider the case when only one constraint is binding. Note that the last two rows of G are linear independent. So we only need to verify whether the first row is linearly independent from the last two rows. When $x^2 + 2y^2 - 6 = 0$, then either $x \neq 0$ or $y \neq 0$. Then, clearly, the first row is linearly independent from the last two rows. When $-x = 0$, since $x^2 + 2y^2 - 6 \neq 0$, the three rows are linearly independent. Similarly, the three rows are linearly independent when $-y = 0$. Therefore, G has full row rank when only one constraint is binding.
2. Next consider the case when two constraints are binding. This leads to five corner points:

$$(0, \pm\sqrt{3}), (\pm\sqrt{6}, 0)$$

and

$$(0,0).$$

Evaluation of G at these points shows that G always has full row rank.

Therefore, no candidate extremal points are found.

Step 3. Consider $\mathcal{L} := x + y - \lambda_1(x^2 + 2y^2 - 6) - \lambda_2(-x) - \lambda_3(-y)$.

Step 4. Kuhn-Tucker Conditions:

$$\begin{cases} [1 - 2\lambda_1 x + \lambda_2,\ 1 - 4\lambda_1 y + \lambda_3] = [0,\ 0]; \\ \lambda_1(x^2 + 2y^2 - 6) = 0; \\ \lambda_2 x = 0; \\ \lambda_3 y = 0; \\ \lambda_1, \lambda_2, \lambda_3, x, y \geq 0; \\ x^2 + 2y^2 \leq 6. \end{cases}$$

i) If no constraint is binding, $\lambda_i = 0,\ i = 1, 2, 3$. Substitution into $[1-2\lambda_1 x+\lambda_2,\ 1-4\lambda_1 y+\lambda_3] = [0,\ 0]$ shows that this is not possible.

ii) If only one constraint is binding:

1. If $x^2 + 2y^2 = 6$, $x > 0$ and $y > 0$. Then $\lambda_2 = \lambda_3 = 0$. Substitution into $[1 - 2\lambda_1 x + \lambda_2,\ 1 - 4\lambda_1 y + \lambda_3] = [0,\ 0]$ yields:
$$\begin{cases} \lambda_1 = \frac{1}{4}; \\ x = 2; \\ y = 1. \end{cases}$$
2. If $x = 0$, $x^2 + 2y^2 < 6$ and $y > 0$. Then $\lambda_1 = \lambda_3 = 0$. Substitution into $[1 - 2\lambda_1 x + \lambda_2,\ 1 - 4\lambda_1 y + \lambda_3] = [0,\ 0]$ yields $\lambda_2 = -1$. Which is not possible.
3. If $y = 0$, $x^2 + 2y^2 < 6$ and $x > 0$. Then $\lambda_1 = \lambda_2 = 0$. Substitution into $[1 - 2\lambda_1 x + \lambda_2,\ 1 - 4\lambda_1 y + \lambda_3] = [0,\ 0]$ yields $\lambda_3 = -1$. Which is not possible.

iii) If two constraints are binding. This is essentially checking the three corner points.

1. When $x^2 + 2y^2 = 6$, $x = 0$ and $y > 0$. Then $\lambda_3 = 0$ and, consequently, $(x, y) = (0, \pm\sqrt{3})$. Substitution into $[1 - 2\lambda_1 x + \lambda_2,\ 1 - 4\lambda_1 y + \lambda_3] = [0,\ 0]$ shows $\lambda_2 = -1$. Which is not possible.
2. When $x^2 + 2y^2 = 6$, $y = 0$ and $x > 0$. Then $\lambda_2 = 0$ and, consequently, $(x, y) = (\pm\sqrt{6}, 0)$. Substitution into $[1 - 2\lambda_1 x + \lambda_2,\ 1 - 4\lambda_1 y + \lambda_3] = [0,\ 0]$ shows $\lambda_3 = -1$. Which is not possible.
3. When $x = 0$, $y = 0$ and $x^2 + 2y^2 < 6$. Then $\lambda_1 = 0$ and $(x, y) = (0, 0)$. Substitution into $[1 - 2\lambda_1 x + \lambda_2,\ 1 - 4\lambda_1 y + \lambda_3] = [0,\ 0]$ shows $\lambda_2 = -1$. Which is not possible.

So, there is just one candidate point $(x, y, \lambda_1, \lambda_2, \lambda_3) = (2, 1, \frac{1}{4}, 0, 0)$.

Weierstrass' theorem tells us that a global maximum exists. Therefore, the maximum is attained at $(x, y) = (2, 1)$.

c. The feasible domain is compact. By Weierstrass' theorem therefore a global maximum and global minimum exist. To find all candidate solutions, we proceed along the lines of Algorithm 4.9.

Step 1. $Df = [2y \ -2x]$, $Dg = \begin{bmatrix} 2x & 2y \\ 2x & 2(y-2) \end{bmatrix}$. Therefore, $f, g \in C^1$.

Step 2. Check invertibility/full row rank of

$$G = \begin{bmatrix} 2x & 2y & x^2+y^2-4 & 0 \\ 2x & 2(y-2) & 0 & x^2+(y-2)^2-4 \end{bmatrix}.$$

By construction G has full row rank at the interior points. So we only need to check boundary points.

1. First consider the case when only one constraint is binding. Assume the first constraint, $x^2 + y^2 = 4$ is the only binding constraint. Then matrix G has not full row rank iff. $x = y = 0$. But this condition violates the first binding constraint. So if only the first constraint is binding G has full row rank. Similarly it follows that if only the second constraint is binding matrix G has full row rank.
2. Finally, both constraints are simultaneously binding at the points: $(\pm\sqrt{3}, 1)$. Evaluation of G at these points shows that G always has full row rank.

Therefore, no candidate points are found.

Step 3. Consider $\mathcal{L} := y^2 - x^2 - \lambda_1(x^2 + y^2 - 4) - \lambda_2(x^2 + (y-2)^2 - 4)$.

Step 4. Kuhn-Tucker Conditions:

$$\begin{cases} [-2x - 2\lambda_1 x - 2\lambda_2 x \ 2y - 2\lambda_1 y - 2\lambda_2(y-2)] = [0 \ 0]; \\ \lambda_1(x^2 + y^2 - 4) = 0; \\ \lambda_2(x^2 + (y-2)^2 - 4) = 0; \\ \lambda_1, \lambda_2 \geq 0; \\ x^2 + y^2 \leq 4; \\ x^2 + (y-2)^2 \leq 4. \end{cases}$$

i) If no constraint is binding, $\lambda_i = 0, \ i = 1, 2$.
Substitution into $[-2x - 2\lambda_1 x - 2\lambda_2 x, \ 2y - 2\lambda_1 y - 2\lambda_2(y-2)] = [0, \ 0]$ yields $(x, y) = (0, 0)$. But $(0, 0)$ is binding in the second constraint. So, we don't get any solutions here.

ii) If only one constraint is binding:

1. If $x^2 + y^2 = 4$ and $x^2 + (y-2)^2 < 4$. Then $\lambda_2 = 0$.
Substitution into $[-2x - 2\lambda_1 x - 2\lambda_2 x, \ 2y - 2\lambda_1 y - 2\lambda_2(y-2)] = [0, \ 0]$ yields:

$$\begin{cases} -2x - 2\lambda_1 x = 0; \\ 2y - 2\lambda_1 y = 0; \\ x^2 + y^2 = 4. \end{cases}$$

From the first two equations we have that either $(x, y) = (0, 0)$ or $(x, \lambda_1) = (0, 1)$. Clearly $(0, 0)$ does not satisfy the third equation. With $(x, \lambda_1) = (0, 1)$, $y = \pm 2$ results from the third equation. However, only $(x, y) = (0, 2)$ satisfies the second constraint. So, the point $(x, y) = (0, 2)$ with $(\lambda_1, \lambda_2) = (0, 1)$ satisfies all conditions.

2. If $x^2 + (y-2)^2 = 4$ and $x^2 + y^2 < 4$. Then $\lambda_1 = 0$.
   Substitution into $[-2x - 2\lambda_1 x - 2\lambda_2 x,\ 2y - 2\lambda_1 y - 2\lambda_2(y-2)] = [0,\ 0]$ yields:

$$\begin{cases} -2x - 2\lambda_2 x = 0; \\ 2y - 2\lambda_2(y-2) = 0; \\ x^2 + (y-2)^2 = 4. \end{cases}$$

   From the first and third equation we derive that either $(x, y) = (0, 0)$ or $(x, y) = (0, 4)$. As $(0, 4)$ violates the first inequality constraint, the only candidate solution that remains is $(x, y) = (0, 0)$ with $(\lambda_1, \lambda_2) = (0, 0)$.

iii) The only points where both constraints are binding are: $(\pm\sqrt{3}, 1)$.
   Substitution into $[-2x - 2\lambda_1 x - 2\lambda_2 x,\ 2y - 2\lambda_1 y - 2\lambda_2(y-2)] = [0,\ 0]$ yields:

$$\begin{cases} 1 + \lambda_1 + \lambda_2 = 0; \\ 1 - \lambda_1 + \lambda_2 = 0. \end{cases}$$

   As $\lambda_i \geq 0$ it is clear from the first equation that there is no appropriate solution in this case.

So we find two candidate solutions: $(x, y) = (0, 0)$ and $(x, y) = (0, 2)$.

Weierstrass' theorem tells us that a global maximum exists. Therefore, the maximum is attained at $(0, 2)$.

d. The feasible domain is compact. So, by Weierstrass' theorem, a global maximum and global minimum exist. To find all candidate solutions, we proceed along the lines of Algorithm 4.9.

Step 1. $Df = [1 + y \ x]$, $Dg = \begin{bmatrix} 2x & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}$. Therefore, $f, g \in C^1$.

Step 2. Check invertibility/full row rank of

$$G = \begin{bmatrix} 2x & 1 & x^2 + y - 2 & 0 & 0 \\ -1 & 0 & 0 & -x & 0 \\ 0 & -1 & 0 & 0 & -y \end{bmatrix}.$$

By construction G has full row rank at the interior points. So we only need to check boundary points.

1. First consider when only one constraint is binding. Assume the first constraint, $x^2 + y = 2$ is the only binding constraint. Then matrix G has full row rank as both x and y differ from

zero. Similarly it follows that if either only the second or third constraint is binding, matrix G has full row rank.

2. Finally, two constraints are simultaneously binding at the points: $(0,0)$, $(0,2)$ and $(\sqrt{2},0)$. Evaluation of G at these points shows that G always has full row rank.

Therefore, no candidate points are found.

Step 3. As we have to minimize $f(x,y) = x + xy$, consider the Lagrange function:

$$\mathcal{L} := -x - xy - \lambda_1(x^2 + y - 2) + \lambda_2 x + \lambda_3 y.$$

Step 4. Kuhn-Tucker Conditions:

$$\begin{cases} [-1 - y - 2\lambda_1 x + \lambda_2,\ -x - \lambda_1 + \lambda_3] = [0,\ 0]; \\ \lambda_1(x^2 + y - 2) = 0; \\ \lambda_2 x = 0; \\ \lambda_3 y = 0; \\ \lambda_i \geq 0,\ i = 1,2,3; \\ x^2 + y \leq 2;\ x \geq 0;\ y \geq 0. \end{cases}$$

i) If no constraint is binding, $\lambda_i = 0,\ i = 1,2,3$.
Substitution into $[-1 - y - 2\lambda_1 x + \lambda_2\ -x - \lambda_1 + \lambda_3] = [0\ 0]$ yields $(x,y) = (0,-1)$. But $y \geq 0$.
So, we don't get any solutions here.

ii) If only one constraint is binding:

1. If $x^2 + y = 2$; $x > 0$ and $y > 0$. Then $\lambda_2 = \lambda_3 = 0$.
Substitution into $[-1 - y - 2\lambda_1 x + \lambda_2,\ -x - \lambda_1 + \lambda_3] = [0,\ 0]$ yields:

$$\begin{cases} -1 - y - 2\lambda_1 x = 0; \\ -x - \lambda_1 = 0; \\ x^2 + y = 2. \end{cases}$$

As $x \geq 0$ and $\lambda_i \geq 0$ one easily verifies that above set of equations has no solution.

2. If $x^2 + y < 2$; $x = 0$ and $y > 0$. Then $\lambda_1 = \lambda_3 = 0$.
Substitution into $[-1 - y - 2\lambda_1 x + \lambda_2,\ -x - \lambda_1 + \lambda_3] = [0,\ 0]$ yields:

$$\begin{cases} -1 - y + \lambda_2 = 0; \\ 0 = 0; \\ x = 0. \end{cases}$$

So, all points $(0,y)$, with $0 < y < 2$, and $(\lambda_1, \lambda_2, \lambda_3) = (0, 1 + y, 0)$ satisfy all conditions.

3. If $x^2 + y < 2$; $x > 0$ and $y = 0$. Then $\lambda_1 = \lambda_2 = 0$.
   Substitution into $[-1 - y - 2\lambda_1 x + \lambda_2,\ -x - \lambda_1 + \lambda_3] = [0,\ 0]$ yields:

$$\begin{cases} -1 - y = 0; \\ x = 0; \\ y = 0. \end{cases}$$

   Clearly there is no solution here.

iii) If two constraints are binding:
   This occurs at the points $(0, 0)$ (with $\lambda_1 = 0$); $(0, 2)$ (with $\lambda_3 = 0$) and $(\sqrt{2}, 0)$ (with $\lambda_2 = 0$). Evaluation of $[-1-y-2\lambda_1 x+\lambda_2,\ -x-\lambda_1+\lambda_3] = [0,\ 0]$ at these points shows that $(0, 0)$ satisfies all KKT conditions too with $(\lambda_1, \lambda_2, \lambda_3) = (0, 1, 0)$, and $(0, 2)$ too with $(\lambda_1, \lambda_2, \lambda_3) = (0, 3, 0)$.

So we find the candidate solutions $(0, y)$, with $0 \leq y \leq 2$, and $(\lambda_1, \lambda_2, \lambda_3) = (0, 1 + y, 0)$.
Weierstrass' theorem tells us that a global minimum exists. Evaluation of f shows that at all these points f attains the same value of zero. So all candidate points are global minimum locations.

e. First note the feasible domain is not bounded, so not compact. To find all candidate solutions, we proceed along the lines of Algorithm 4.9.
Step 1. $Df = [1\ 1]$, $Dg = \begin{bmatrix} x & 1 \end{bmatrix}$. Therefore, $f, g \in C^1$.
Step 2. Clearly, g has always full row rank. Therefore, no candidate points are found.
Step 3. Consider $\mathcal{L} = x + y - \lambda(y + \frac{1}{2}x^2)$.
Step 4. Kuhn-Tucker Conditions:

$$\begin{cases} [1 - \lambda x\ 1 - \lambda] = [0\ 0]; \\ \lambda(y + \frac{1}{2}x^2) = 0; \\ \lambda \geq 0; \\ y + \frac{1}{2}x^2 \leq 0. \end{cases}$$

From above conditions we obtain $(x, y, \lambda) = (1, -\frac{1}{2}, 1)$ is the only candidate solution.
With $\lambda = 1$, the constrained Lagrangian $\tilde{L}(x, y) := x + y - (y + \frac{1}{2}x^2)$. So, $H\tilde{L} = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$. From which we conclude that $\tilde{L}(x, y)$ is concave on $\mathbb{R}^2$. So, by Proposition 4.11, f has a global maximum at $(1, -\frac{1}{2})$.

f. First notice the feasible domain is not compact. Furthermore observe that by choosing x arbitrarily negative and $y = x^2$, $f(x, y) = x - \frac{1}{3}x^3$ becomes arbitrarily large. So the problem has not a global maximum. To find all candidate solutions, we proceed along the lines of Algorithm 4.9.
Step 1. With $g_1 := x^2 - y$ and $g_2 := x - 2$: $Df = [1 - \frac{1}{3}y,\ -\frac{1}{3}x]$; $Dg = \begin{bmatrix} 2x & -1 \\ 1 & 0 \end{bmatrix}$. Therefore, $f, g \in C^1$.

Step 2. Check invertibility/full row rank of

$$G = \begin{bmatrix} 2x & -1 & x^2 - y & 0 \\ 1 & 0 & 0 & x - 2 \end{bmatrix}.$$

By construction, G has full row rank at the interior points. So we only need to check boundary points.

1. First consider when only one constraint is binding. One easily verifies that G has than in both cases rank 2.
2. If both constraints are binding, it follows $(x, y) = (2, 4)$. Again one easily verifies that G has rank 2 then.

So we don't find any candidate locations in this step.

Step 3. Consider $\mathcal{L} = x - \frac{1}{3}xy - \lambda_1(x^2 - y) - \lambda_2(x - 2)$.

Step 4. Kuhn-Tucker Conditions:

$$\begin{cases} [1 - \frac{1}{3}y - 2\lambda_1 x - \lambda_2,\ -\frac{1}{3}x + \lambda_1] = [0,\ 0]; \\ \lambda_1(x^2 - y) = 0; \\ \lambda_2(x - 2) = 0; \\ \lambda_1, \lambda_2 \geq 0; \\ y \geq x^2 \text{ and } x \leq 2. \end{cases}$$

i) If no constraint is binding, $\lambda_i = 0,\ i = 1, 2$.
Substitution into $[1 - \frac{1}{3}y - 2\lambda_1 x - \lambda_2,\ -\frac{1}{3}x + \lambda_1] = [0,\ 0]$ yields $(x, y, \lambda_1, \lambda_2) = (0, 3, 0, 0)$ as a candidate solution.

ii) If only one constraint is binding:
1. If $y = x^2$ and $x < 2$. Then $\lambda_2 = 0$.
Substitution into $[1 - \frac{1}{3}y - 2\lambda_1 x - \lambda_2,\ -\frac{1}{3}x + \lambda_1] = [0,\ 0]$ yields:

$$\begin{cases} 1 - \frac{1}{3}x^2 - 2\lambda_1 x = 0; \\ \lambda_1 = \frac{1}{3}x; \\ y = x^2. \end{cases}$$

Which yields $(x, y, \lambda_1, \lambda_2) = (1, 1, \frac{1}{3}, 0)$ as a candidate solution.
2. If $x = 2$ and $y > x^2$. Then $\lambda_1 = 0$.
Substitution into $[1 - \frac{1}{3}y - 2\lambda_1 x - \lambda_2,\ -\frac{1}{3}x + \lambda_1] = [0,\ 0]$ yields $x = 0$, which violates the assumption $x = 2$.

iii) If both constraints are binding, $(x, y) = (2, 4)$. However, the KKT conditions don't have an appropriate solution then.

So, the only two candidate solutions are $(x, y, \lambda_1, \lambda_2) = (0, 3, 0, 0)$ and $(x, y, \lambda_1, \lambda_2) = (1, 1, \frac{1}{3}, 0)$.

To see whether the problem has a local maximum at either one of these points we consider Theorem 4.19.

1. First consider $(x, y, \lambda_1, \lambda_2) = (0, 3, 0, 0)$.
   As both constraints are not binding we have to consider the hessian of f at this point. Note that $Hf = \begin{bmatrix} 0 & -\frac{1}{3} \\ -\frac{1}{3} & 0 \end{bmatrix}$. From this we conclude that $Hf(0, 3)$ is indefinite. So, the constrained optimization problem has not a local maximum at $(0, 3)$.
2. Next consider $(x, y, \lambda_1, \lambda_2) = (1, 1, \frac{1}{3}, 0)$.
   In this case g_1 is binding. So, following Theorem 4.19, consider next

$$H\tilde{L} := Hf - \frac{1}{3}Hg_1 = \begin{bmatrix} 0 & -\frac{1}{3} \\ -\frac{1}{3} & 0 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } B := [1\ 1]^T.$$

   Then, $B^T H\tilde{L}B = -\frac{4}{3} < 0$. So, the constrained optimization problem has a local maximum at $(1, 1)$.

4.4 Some pathological examples

EXERCISE 4.4 Solve the optimization problem: maximize $f(x, y, z) = x-(z-1)^2$, w.r.t. x, y, z, subject to the constraints $y - (1 - x)^3 \leq 0$, $x \geq 0$, $y \geq 0$ and $0 \leq z \leq 2$.

ANSWER 4.4

As $0 \leq y \leq (1-x)^3$ and $x \geq 0$, it follows $0 \leq x \leq 1$ and $0 \leq y \leq 1$. So the feasible domain is compact. By Weierstrass' Theorem a global maximum and global minimum exist. To find all candidate solutions, we proceed along the lines of Algorithm 4.9.

Step 1. $Df = [1\ 0\ -2(z-1)]$, $Dg = \begin{bmatrix} 3(1-x)^2 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}$. Therefore, $f, g \in C^1$.

Step 2. Check invertibility/full row rank of

$$G = \begin{bmatrix} 3(1-x)^2 & 1 & 0 & y-(1-x)^3 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & -x & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & -y & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & -z & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & z-2 \end{bmatrix}.$$

By construction G has full row rank at the interior points. So we only need to check boundary points.

1. First consider when only one constraint is binding. Then one easily verifies that in all cases matrix G has full row rank.
2. Now assume two constraints are binding.
   1. Case $y-(1-x)^3=0$ and $x=0$. That is: $x=0$ and $y=1$. Then G has full row rank.
   2. Case $y-(1-x)^3=0$ and $y=0$. That is: $x=1$ and $y=0$. Then G has not full row rank.
   3. Case $y-(1-x)^3=0$ and $z=0$. That is: $x=0$ and $y=1$. Then G has full row rank.
   4. Case $y-(1-x)^3=0$ and $z=2$. That is: $x=0$ and $y=1$. Then G has full row rank.
   5. Case $x=0$ and $y=0$. Then G has full row rank.
   6. Case $x=0$ and $z=0$. Then G has full row rank.
   7. Case $x=0$ and $z=2$. Then G has full row rank.
   8. Case $y=0$ and $z=0$. Then G has full row rank.
   9. Case $y=0$ and $z=2$. Then G has full row rank.
3. Now assume three constraints are binding.
   1. Case $y-(1-x)^3=0$, $x=0$ and $z=0$. That is: $x=0$; $y=1$ and $z=0$. Then G has full row rank.
   2. Case $y-(1-x)^3=0$, $x=0$ and $z=2$. That is: $x=0$; $y=1$ and $z=2$. Then G has full row rank.
   3. Case $y-(1-x)^3=0$, $y=0$ and $z=0$. That is: $x=1$; $y=0$ and $z=0$. Then G has not full row rank.
   4. Case $y-(1-x)^3=0$, $y=0$ and $z=2$. That is: $x=1$; $y=0$ and $z=2$. Then G has not full row rank.
   5. Case $x=0$, $y=0$ and $z=0$. Then G has full row rank.
   6. Case $x=0$, $y=0$ and $z=2$. Then G has full row rank.

So, at all points $(1,0,z),\ 0\leq z\leq 2$, G has not full row rank. So all these points are candidate points for a maximum location.

Another way to arrive at this conclusion is to use Gaussian elimination on matrix G. One easily verifies, using this elimination procedure, that the rank of matrix G coincides with the rank of next matrix

$$\tilde{G}=\begin{bmatrix} -1 & 0 & 0 & 0 & -x & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & -y & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & -z & 0 \\ 0 & 0 & 0 & y-(1-x)^3 & 3x(1-x)^2 & -y & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -z & z-2 \end{bmatrix}.$$

From this it is clear that the rank of $\tilde{G}$ is 5 iff not simultaneously $y-(1-x)^3=3x(1-x)^2=-y=0$. So, the rank of $\tilde{G}$ is smaller than 5 if $x=1$, $y=0$ and $0\leq z\leq 2$.

Step 3. Consider $\mathcal{L}=x-(z-1)^2-\lambda_1(y-(1-x)^3)+\lambda_2 x+\lambda_3 y+\lambda_4 z-\lambda_5(z-2)$.

Step 4. Kuhn-Tucker Conditions:

$$\begin{cases} [1-3\lambda_1(1-x)^2+\lambda_2,\ -\lambda_1+\lambda_3,\ -2(z-1)+\lambda_4-\lambda_5]=[0,\ 0,\ 0]; \\ \lambda_1(y-(1-x)^3)=0; \\ \lambda_2 x=0; \\ \lambda_3 y=0; \\ \lambda_4 z=0; \\ \lambda_5(z-2)=0; \\ \lambda_i\geq 0,\ i=1,\cdots,5; \\ y\leq(1-x)^3;\ x\geq 0;\ y\geq 0;\ 0\leq z\leq 2. \end{cases}$$

i) If no constraint is binding, $\lambda_i=0,\ i=1,\cdots,5$. In that case the KKT conditions have no solution.

ii) If only one constraint is binding:

1. If $y=(1-x)^3$ is the only one. Then $\lambda_i=0,\ i=2,\cdots,5$. In that case the KKT conditions have no solution.
2. If $x=0$ is the only one. Then $\lambda_i=0,\ i=1,3,4,5$. In that case the KKT conditions have no solution.
3. If $y=0$ is the only one. Then $\lambda_i=0,\ i=1,2,4,5$. In that case the KKT conditions have no solution.
4. If $z=0$ is the only one. Then $\lambda_i=0,\ i=1,2,3,5$. In that case the KKT conditions have no solution.
5. If $z=2$ is the only one. Then $\lambda_i=0,\ i=1,2,3,4$. In that case the KKT conditions have no solution.

iii) If precisely two constraints are binding:

1. If $y=(1-x)^3$ and $x=0$, $\lambda_3=\lambda_4=\lambda_5=0$. Then also $\lambda_1=0$ and it follows that the KKT conditions have no solution.
2. If $y=(1-x)^3$ and $y=0$, $\lambda_2=\lambda_4=\lambda_5=0$. Then $x=1$ and it follows that the KKT conditions have no solution.
3. If $y=(1-x)^3$ and $z=0$, $\lambda_2=\lambda_3=\lambda_5=0$. Then also $\lambda_1=0$ and it follows that the KKT conditions have no solution.
4. If $y=(1-x)^3$ and $z=2$, $\lambda_2=\lambda_3=\lambda_4=0$. Then also $\lambda_1=0$ and it follows that the KKT conditions have no solution.
5. If $x=0$ and $y=0$, $\lambda_1=\lambda_4=\lambda_5=0$. Then also $\lambda_3=0$ and it follows that the KKT conditions have no solution.
6. If $x=0$ and $z=0$, $\lambda_1=\lambda_3=\lambda_5=0$. Then it follows that the KKT conditions have no solution.
7. If $x=0$ and $z=2$, $\lambda_1=\lambda_3=\lambda_4=0$. Then it follows that the KKT conditions have no solution.

8. If $y = 0$ and $z = 0$, $\lambda_1 = \lambda_2 = \lambda_5 = 0$. Then it follows that the KKT conditions have no solution.
9. If $y = 0$ and $z = 2$, $\lambda_1 = \lambda_2 = \lambda_4 = 0$. Then it follows that the KKT conditions have no solution.

iv) If precisely three constraints are binding:

1. If $y = (1-x)^3$; $x = 0$ and $z = 0$, $\lambda_3 = \lambda_5 = 0$. Then also $\lambda_1 = 0$ and it follows that the KKT conditions have no solution.
2. If $y = (1-x)^3$; $x = 0$ and $z = 2$, $\lambda_3 = \lambda_4 = 0$. Then also $\lambda_1 = 0$ and it follows that the KKT conditions have no solution.
3. If $x = 0$; $y = 0$ and $z = 0$, $\lambda_1 = \lambda_5 = 0$. Then also $\lambda_3 = 0$ and it follows that the KKT conditions have no solution.
4. If $x = 0$; $y = 0$ and $z = 2$, $\lambda_1 = \lambda_4 = 0$. Then also $\lambda_3 = 0$ and it follows that the KKT conditions have no solution.

So, there are no additional candidate solutions.

So we find the candidate solutions points $(1, 0, z),\ 0 \leq z \leq 2$.

Weierstrass' theorem tells us that a global maximum exists. Evaluation of f at these points shows that at $(1, 0, 1)$ a global maximum occurs.

EXERCISE 4.5 Solve the optimization problem: maximize $f(x, y, z) = x-(z-1)^2$, w.r.t. x, y, z, subject to the constraints $y^2 + (x-1)^3 \leq 0$, $x \geq 0$ and $0 \leq z \leq 2$.

ANSWER 4.5

As $x \geq 0$ and $0 \leq y^2 \leq -(x-1)^3$, it follows $0 \leq x \leq 1$ and $-1 \leq y \leq 1$. As also $0 \leq z \leq 2$ it follows that the feasible domain is compact. So, by Weierstrass' Theorem, a global maximum and global minimum exist. To find all candidate solutions, we proceed along the lines of Algorithm 4.9.

Step 1. $Df = [1\ 0\ -2(z-1)]$, $Dg = \begin{bmatrix} 3(x-1)^2 & 2y & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}$. Therefore, $f, g \in C^1$.

Step 2. Check invertibility/full row rank of

$$G = \begin{bmatrix} 3(x-1)^2 & 2y & 0 & y^2+(x-1)^3 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & -x & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -z & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & z-2 \end{bmatrix}.$$

By construction G has full row rank at the interior points. So we only need to check boundary points.

1. First consider when only one constraint is binding.

   Case $y^2 + (x-1)^3 = 0$. Then G has not rank 4 iff. $x = 1$, $y = 0$, $0 < z < 2$.

When either precisely one of the other constraints is binding one easily verifies G always has rank 4.

2. Next consider the case when precisely two constraints are binding.
   1. Case $y^2 + (x-1)^3 = 0$ and $x = 0$. So, $y = \pm 1$. Then G has rank 4.
   2. Case $y^2 + (x-1)^3 = 0$ and $z = 0$. Then G has not rank 4 iff. $x = 1$, $y = 0$.
   3. Case $y^2 + (x-1)^3 = 0$ and $z = 2$. Then G has not rank 4 iff. $x = 1$, $y = 0$. One easily verifies that in all other cases G has rank 4.
3. Finally consider the case when three constraints are binding. That is: $y^2 + (x-1)^3 = 0$, $x = 0$ and either $z = 0$ or $z = 2$. So, $y = \pm 1$. Then G has rank 4.

So, all points $(1, 0, z)$, with $0 \leq z \leq 2$ are candidate solutions.

Another way to find all points where matrix G has a rank smaller than 4 is using Gaussian elimination. Using this procedure we find that the rank of matrix G coincides with the rank of matrix

$$\tilde{G} = \begin{bmatrix} -1 & 0 & 0 & 0 & -x & 0 & 0 \\ 0 & 2y & 0 & y^2 + (x-1)^3 & -3x(x-1)^2 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -z & 0 \\ 0 & 0 & 0 & 0 & 0 & -z & z-2 \end{bmatrix}.$$

One easily verifies that matrix $\tilde{G}$ has a rank smaller than 4 iff. $y = y^2 + (x-1)^3 = 0$. So, at all points $(1, 0, z)$, with $0 \leq z \leq 2$.

Step 3. Consider $\mathcal{L} := x - (z-1)^2 - \lambda_1(y^2 + (x-1)^3) + \lambda_2 x + \lambda_3 z - \lambda_4(z-2)$.

Step 4. Kuhn-Tucker Conditions:

$$\begin{cases} [1 - 3\lambda_1(x-1)^2 + \lambda_2,\ -2\lambda_1 y,\ -2(z-1) + \lambda_3 - \lambda_4] = [0,\ 0,\ 0]; \\ \lambda_1(y^2 + (x-1)^3) = 0; \\ \lambda_2 x = 0; \\ \lambda_3 z = 0; \\ \lambda_4(z-2) = 0; \\ \lambda_i \geq 0,\ i = 1, \cdots, 4; \\ y^2 + (x-1)^3 \leq 0;\ x \geq 0;\ 0 \leq z \leq 2. \end{cases}$$

From $\lambda_1 y = 0$ it follows that either $\lambda_1 = 0$ or $y = 0$. If $\lambda_1 = 0$, it follows from $1 - 3\lambda_1(x-1)^2 + \lambda_2 = 0$ that $\lambda_2 < 0$, which violates the fact that $\lambda_2 \geq 0$. Similarly, if $\lambda_1 \neq 0$ it follows that $y^2 + (x-1)^3 = 0$. Together with $y = 0$, this implies $x = 1$. But then $\lambda_2 = 0$. Which gives a contradiction in $1 - 3\lambda_1(x-1)^2 + \lambda_2 = 0$.

So, we don't find any additional candidates for a maximum location in this step.

So we find the candidate solutions points $(1, 0, z),\ 0 \leq z \leq 2$.

Weierstrass' Theorem tells us that a global maximum exists. Evaluation of f at these points shows that at $(1, 0, 1)$ a global maximum occurs.

4.5 Mixed constraints: a shortcut

EXERCISE 4.6 Solve the next constrained optimization problems.

a. Maximize $f(x,y) = x - y$, subject to the constraints $y \geq -x$, $y \leq x$ and $x^2 + y^2 = 2$.

b. Maximize $U(x,y,z) = x^\alpha y^\beta z^\gamma$ subject to the constraints $x \geq 15$, $y \geq 10$, $z \geq 5$, $p_1 x + p_2 y + p_3 z = 100$ and $x + y \leq 50$. Here, all parameters are positive, $\alpha + \beta + \gamma = 1$, $p_1 = 1$, $p_2 = 2$ and $p_3 = 3$. In particular evaluate the cases $\alpha = \beta = \frac{1}{4}$, $\gamma = \frac{1}{2}$, and $\alpha = \beta = 0.1, \gamma = 0.8$.

ANSWER 4.6

a. Consider the set $\Gamma = \{(x,y) \mid g_1(x,y) := -x - y \leq 0, g_2(x,y) := y - x \leq 0, h(x,y) := x^2 + y^2 - 2 = 0\}$. Note that Γ describes a part of a circle. So, Γ is a compact set. Since f is continuous we know by Weierstrass' theorem that the problem has a solution. We next use Algorithm 4.20 to solve the problem.

Step 1. $Df = [1 \ -1]$, $Dg_1 = [-1 \ -1]$, $Dg_2 = [-1 \ 1]$ and $Dh = [2x \ 2y]$. All partial derivatives are continuous. So, $f, g, h \in C^1$.

Step 2. Matrix $\tilde{G}$ is given by

$$\tilde{G} = \begin{bmatrix} -1 & -1 & -x-y & 0 \\ -1 & 1 & 0 & y-x \\ 2x & 2y & 0 & 0 \end{bmatrix}.$$

To see whether this matrix has always full row rank on Γ consider next cases:

1. both g_i are non-binding. Then G has not full row rank iff. $x = y = 0$. Which violates the non-binding conditions.
2. g_1 is binding and g_2 not, it follows $(x,y) = (1,-1)$ or $(x,y) = (-1,1)$. In both cases $\tilde{G}$ has full row rank.
3. g_2 is binding and g_1 not, it follows $(x,y) = (1,1)$ or $(x,y) = (-1,-1)$. In both cases $\tilde{G}$ has full row rank.
4. $g_i, \ i = 1,2$ is binding. Then it follows $(x,y) = (0,0)$, which violates $h = 0$.

So we don't find any potential candidate solutions here.

This conclusion can also, again, be obtained by using Gaussian elimination. Using this procedure we find that the rank of matrix $\tilde{G}$ coincides with the rank of matrix

$$\begin{bmatrix} -1 & -1 & -x-y & 0 \\ 0 & 2 & x+y & y-x \\ 0 & 0 & -(x+y)^2 & -(x-y)^2 \end{bmatrix}.$$

So, the rank of matrix $\tilde{G}$ is smaller than 3 iff $-(x+y)^2 = -(x-y)^2 = 0$. That is, if $(x,y) = (0,0)$. But, as already noted above, $(0,0)$ does not belong to the constraint set. Which confirms the above conclusion that we don't find any candidate solutions in this step.

Step 3. Consider $L(x, y, \lambda_1, \lambda_2, \mu_1) := x - y + \lambda_1(x + y) - \lambda_2(y - x) - \mu_1(x^2 + y^2 - 2)$.

Step 4. The necessary conditions are:

1. $[1 + \lambda_1 + \lambda_2 - 2\mu_1 x, \ -1 + \lambda_1 - \lambda_2 - 2\mu_1 y] = [0 \ 0]$;
2. i) $\lambda_1(x + y) = 0$; ii) $\lambda_2(y - x) = 0$;
3. i) $x + y \geq 0$; ii) $x \geq y$; and iii) $\lambda_i \geq 0, \ i = 1, 2$;
4. $x^2 + y^2 = 2$.

1. No constraint is binding. So $\lambda_i = 0, \ i = 1, 2$. Then $2\mu_1 x = 1$ and $2\mu_1 y = -1$. Therefore, $\mu_1(x + y) = 0$. Since g_1 is not binding it follows that $\mu_1 = 0$. Which violates $2\mu_1 x = 1$.
2. g_1 is not binding and $y = x$. So, $\lambda_1 = 0$.
   Substitution into h yields $(x, y) = (1, 1)$ or $(x, y) = (-1, -1)$. Note that $(-1, -1)$ violates g_1.
   Substitution of $(1, 1)$ into item 1. yields $\mu_1 = 0$ and $\lambda_2 = -1$, which violates item 3.
3. g_2 is not binding and $y = -x$. So, $\lambda_2 = 0$.
   Substitution into h yields $(x, y) = (-1, 1)$ or $(x, y) = (1, -1)$. Note that $(-1, 1)$ violates g_2.
   Substitution of $(1, -1)$ into item 1. yields $\mu_1 = \frac{1}{2}$ and $\lambda_2 = 0$.
4. $y = x$ and $y = -x$. This violates $h = 0$.

So, we find one candidate solution $(x, y, \lambda_1, \lambda_2, \mu_1) = (1, -1, 0, 0, \frac{1}{2})$.

By Weierstrass' theorem, this yields the global maximum location of the problem.

b. Consider the set $\Gamma = \{(x, y) \mid g_1(x, y) := -x + 15 \leq 0, \ g_2(x, y) := -y + 10 \leq 0, \ g_3(x, y) := -z + 5 \leq 0, \ g_4(x, y) := x + y - 50 \leq 0, \ h(x, y) := x + 2y + 3z - 100 = 0\}$. Note that Γ is a compact set. Since f is continuous we know by Weierstrass' theorem that the problem has a solution. We next use Algorithm 4.20 to solve the problem.

Step 1. $Df = [\frac{\alpha f}{x} \ \frac{\beta f}{y} \ \frac{\gamma f}{z}]$, $Dg_1 = [-1 \ 0 \ 0]$, $Dg_2 = [0 \ -1 \ 0]$, $Dg_3 = [0 \ 0 \ -1]$, $Dg_4 = [1 \ 1 \ 0]$ and $Dh = [1 \ 2 \ 3]$. All partial derivatives are continuous. So, $f, g, h \in C^1$.

Step 2. Matrix $\tilde{G}$ is given by

$$\tilde{G} = \begin{bmatrix} -1 & 0 & 0 & -x + 15 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -y + 10 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -z + 5 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & x + y - 50 \\ 1 & 2 & 3 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Using Gaussian elimination one easily verifies that the rank of matrix $\tilde{G}$ coincides with that of matrix

$$\tilde{G} = \begin{bmatrix} -1 & 0 & 0 & 15 - x & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 10 - y & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 5 - z & 0 \\ 0 & 0 & 0 & 15 - x & 10 - y & 0 & x + y - 50 \\ 0 & 0 & 0 & 0 & 10 - y & 3(5 - z) & 50 - x - y \end{bmatrix}.$$

From this matrix we see that the only point where $\tilde{G}$ has not full row rank is $(x, y, z) = (15, 10, 5)$. However this point violates $h = 0$. So we don't find any potential candidate solutions here.

Step 3. Consider $L(x, y, z\lambda_1, \lambda_2, \lambda_3, \lambda_4, \mu) := x^\alpha y^\beta z^\gamma + \lambda_1(x-15) + \lambda_2(y-10) + \lambda_3(z-5) - \lambda_4(x+y-50) - \mu(x+2y+3z-100)$.

Step 4. The necessary conditions are:

1. $[\frac{\alpha f}{x} + \lambda_1 - \lambda_4 - \mu,\ \frac{\beta f}{y} + \lambda_2 - \lambda_4 - 2\mu,\ \frac{\gamma f}{z} + \lambda_3 - 3\mu] = [0\ 0\ 0]$;
2. i) $\lambda_1(x-15) = 0$; ii) $\lambda_2(y-10) = 0$, $\lambda_3(z-5) = 0$, $\lambda_4(x+y-50) = 0$;
3. i) $x \geq 15$; ii) $y \geq 10$; iii) $z \geq 5$; iv) $x + y \leq 50$ and iii) $\lambda_i \geq 0,\ i = 1, \cdots, 4$;
4. $x + 2y + 3z = 100$.

1. No constraint is binding.

   So $\lambda_i = 0,\ i = 1, \cdots, 4$. Then $\mu = \frac{\alpha}{x} f = \frac{\beta}{2y} f = \frac{\gamma}{3z} f$. So, $x = \frac{3\alpha}{\gamma} z$ and $y = \frac{3\beta}{2\gamma} z$. Substitution of this into item 4. yields $(x, y, z) = (100\alpha, \frac{100}{2}\beta, \frac{100}{3}\gamma)$ with $\lambda_i = 0$ and $\mu = \frac{f}{100}$.

   So, provided $x > 15$, $y > 10$, $z > 5$ and $x+y < 50$ this yields a candidate solution. Elementary calculations shows that these conditions apply iff $\alpha > \frac{15}{100}$, $\beta > \frac{20}{100}$, $\gamma > \frac{15}{100}$ and $\alpha < \gamma$.

2. Exact one constraint is binding.

   1. No constraint but g_1 is binding. So $\lambda_i = 0,\ i = 2, \cdots, 4$, and $x = 15$. Then $\mu = \frac{\beta}{2y} f = \frac{\gamma}{3z} f$. So, $x = 15$ and $y = \frac{3\beta}{2\gamma} z$. Substitution of this into item 4. yields $(x, y, z) = (15, \frac{85\beta}{2(\beta+\gamma)}, \frac{85\gamma}{3(\beta+\gamma)})$ with $\lambda_i = 0, i = 2, 3, 4$, $\mu = \frac{(\beta+\gamma)f}{85}$ and $\lambda_1 = \mu - \frac{\alpha}{x} f = (\frac{\beta+\gamma}{85} - \frac{\alpha}{15}) f$.

      So, provided $y > 10$, $z > 5$, $x + y < 50$ and $\lambda_1 \geq 0$, this yields a candidate solution. Elaboration of these inequalities shows they are satisfied iff $\alpha,\ \beta$ and γ satisfy $\frac{15}{70}\beta < \gamma < \frac{65}{20}\beta$, $\alpha \leq \frac{15}{100}$ and $\alpha + \beta + \gamma = 1$.

   2. No constraint but g_2 is binding. So $\lambda_i = 0,\ i = 1, 3, 4$, and $y = 10$. Then $\mu = \frac{\alpha}{x} f = \frac{\gamma}{3z} f$. So, $x = \frac{3\alpha}{\gamma} z$ and $y = 10$. Substitution of this into item 4. yields $(x, y, z) = (\frac{80\alpha}{\alpha+\gamma}, 10, \frac{80\gamma}{3(\alpha+\gamma)})$ with $\lambda_i = 0, i = 1, 3, 4$, $\mu = \frac{(\alpha+\gamma)f}{80}$ and $\lambda_2 = 2\mu - \frac{\beta}{x} f = (\frac{2(\alpha+\gamma)}{80} - \frac{\beta}{10}) f$.

      So, provided $x > 15$, $z > 5$, $x + y < 50$ and $\lambda_2 \geq 0$, this yields a candidate solution. Elaboration of these inequalities shows they are satisfied iff $\alpha,\ \beta$ and γ satisfy $\alpha < \gamma < \frac{65}{15}\alpha$, $\beta \leq \frac{20}{100}$ and $\alpha + \beta + \gamma = 1$.

   3. No constraint but g_3 is binding. So $\lambda_i = 0,\ i = 1, 2, 4$, and $z = 5$. Then $\mu = \frac{\alpha}{x} f = \frac{\beta}{2y} f$. So, $x = \frac{2\alpha}{\beta} y$ and $z = 5$. Substitution of this into item 4. yields $(x, y, z) = (\frac{85\alpha}{\alpha+\beta}, \frac{85\beta}{2(\alpha+\beta)}, 5)$ with $\lambda_i = 0, i = 1, 2, 4$, $\mu = \frac{(\alpha+\beta)f}{85}$ and $\lambda_3 = 3\mu - \frac{\gamma}{z} f = (\frac{3(\alpha+\beta)}{85} - \frac{\gamma}{5}) f$.

      So, provided $x > 15$, $y > 10$, $x+y < 50$ and $\lambda_3 \geq 0$, this yields a candidate solution. Elaboration of these inequalities shows this does not occur for any combination of the parameters $\alpha,\ \beta$ and γ.

   4. No constraint but g_4 is binding. So $\lambda_i = 0,\ i = 1, 2, 3$, and $x + y = 50$. Then elementary calculations show $(x, y, z) = (50(1-\beta), 50\beta, \frac{50}{3}(1-\beta))$ with $\lambda_i = 0, i = 1, 2, 3$, $\mu = \frac{\gamma f}{50(1-\beta)}$ and $\lambda_4 = \frac{\alpha-\gamma}{50(1-\beta)}) f$ solve items 1., 2. and 4.

      So, provided $x > 15$, $y > 10$, $z > 5$ and $\lambda_4 \geq 0$, this yields a candidate solution. Elaboration of these inequalities shows they are satisfied iff $\alpha,\ \beta$ and γ satisfy $\alpha \geq \gamma$, $\frac{10}{50} < \beta < \frac{35}{50}$. and $\alpha + \beta + \gamma = 1$.

3. Exact two constraints are binding.
   1. g_1 and g_2 are binding. So $\lambda_i = 0,\ i = 3, 4,\ x = 15,\ y = 10$. Then, by item 4., $z = \frac{65}{3}$. Therefore, by item 1., $\mu = \frac{\gamma}{65}f$, $\lambda_1 = \mu - \frac{\alpha}{x}f = (\frac{\gamma}{65} - \frac{\alpha}{15})f$ and $\lambda_2 = 2\mu - \frac{\beta}{y}f = (\frac{2\gamma}{65} - \frac{\beta}{10})f$. So, provided $\lambda_i \geq 0,\ i = 1, 2$, this yields a candidate solution. Elaboration of these inequalities shows they are satisfied iff $\alpha,\ \beta$ and γ satisfy $\alpha \leq \frac{15}{65}\gamma$, $\beta \leq \frac{20}{65}\gamma$. and $\alpha + \beta + \gamma = 1$.
   2. g_1 and g_3 are binding. So $\lambda_i = 0,\ i = 2, 4,\ x = 15,\ z = 5$. Then, by item 4., $y = 35$. Therefore, also g_4 is binding. This violates our assumption that precisely 2 constraints are binding. So this does not provide an additional candidate solution.
   3. g_1 and g_4 are binding. Then it follows that $z = 5$ and, therefore, also g_3 is binding. So, this does not yield a candidate solution.
   4. g_2 and g_3 are binding. So $\lambda_i = 0,\ i = 1, 4,\ y = 10,\ z = 5$. Then, by item 4., $x = 65$, which violates item 3.
   5. g_2 and g_4 are binding. So $\lambda_i = 0,\ i = 1, 3,\ y = 10,\ x = 40$. Then, by item 4., $z = \frac{40}{3}$. Therefore, by item 1., $\mu = \frac{\gamma}{40}f$, $\lambda_4 = \frac{\alpha}{x}f - \mu = (\frac{\alpha}{40} - \frac{\gamma}{40})f$ and $\lambda_2 = 2\mu + \lambda_4 - \frac{\beta}{y}f = (\frac{2\gamma}{40} + \frac{\alpha}{40} - \frac{\gamma}{40} - \frac{\beta}{10})f$. So, provided $\lambda_i \geq 0,\ i = 1, 3$ this yields a candidate solution. Elaboration of these inequalities shows they are satisfied iff $\alpha,\ \beta$ and γ satisfy $\alpha \geq \gamma$, $\alpha - 4\beta + \gamma \geq 0$ (or, equivalently, $\beta \leq \frac{1}{5}$) and $\alpha + \beta + \gamma = 1$.
   6. g_3 and g_4 are binding. So $\lambda_i = 0,\ i = 1, 2,\ z = 5,\ x + y = 50$. Then, by item 4., $z = \frac{1}{3}x$. Therefore, g_1 is binding too. So this case does not provide a candidate solution.
4. Exact three constraints are binding.
   1. $g_1,\ g_2$ and g_3 are binding. So $\lambda_4 = 0$, $x = 15$, $y = 10$ and $z = 5$. This violates item 4.
   2. $g_1,\ g_2$ and g_4 are binding. This is not possible.
   3. $g_2,\ g_3$ and g_4 are binding. So $\lambda_1 = 0$, $y = 10$, $z = 5$, $x = 40$. This violates item 4.
5. Case 5: All constraints are binding. This is not possible.

So, depending on the parameter choices, we have to evaluate f at different locations to decide at which point the maximum is obtained.

For instance, if $\alpha = \beta = \frac{1}{4}$, $\gamma = \frac{1}{2}$, it follows the solution is given by the solution mentioned in Case (a).

If $\alpha = \beta = 0.1$ and $\gamma = 0.8$, it follows the solution is given by the solution mentioned in Case (c).i.

4.6 Solving the problem without explicit constraint qualifications

EXERCISE 4.7 Solve the next constrained optimization problems.

a. Maximize $U(x, y) = x + \ln(1 + y)$, subject to the constraints $x + 2y \leq 4$, $x \geq 0$ and $y \geq 0$.

b. Minimize $C(K, L) = rK + \omega L$ subject to the constraints $P(K, L) = 10$ where $P(K, L) = \sqrt{K} + \sqrt{L}$; $L \geq 16$, $K \leq 81$ and $\frac{K}{L} \geq 2$.

c. Maximize $U(\mu,\sigma)=\mu-\frac{1}{2}\sigma^2$ w.r.t. ω_1 and ω_2, where $\mu=2\omega_1+3\omega_2$, $\sigma=\omega_1+\omega_2$; and $\omega_i\geq 0$ and $\omega_1+\omega_2\leq 1$.

d. Maximize $U(x,y)=x^{\frac{1}{2}}y^{\frac{1}{4}}$, subject to the constraints $4x+3y=I$, where $I>0$ is some fixed number, $x\geq\frac{1}{8}I$ and $y\geq\frac{1}{10}I$.

ANSWER 4.7

a. As $x\geq 0$ and $y\geq 0$ it follows that $U(x,y)\in C^1$ on the constraint set. Also $g_1(x,y):=x+2y-4$, $g_2(x,y):=-x$ and $g_3(x,y):=-y$ are continuous differentiable. Consider $\Gamma:=\{(x,y)\mid g_i(x,y)\leq 0\}$. Then one easily verifies that Γ is compact. So, by Weierstrass' theorem the problem has a solution.

To find the candidate solutions we proceed by considering the Lagrange function:

$$L(x,y,\lambda_0,\lambda_1,\lambda_2,\lambda_3):=\lambda_0(x+\ln(1+y))-\lambda_1(x+2y-4)+\lambda_2x+\lambda_3y.$$

By Theorem 4.31 the candidate optimal solutions satisfy:

1. $[D_1L(x,y,\lambda)\ D_2L(x,y,\lambda)]=[\lambda_0x-\lambda_1+\lambda_2,\ \frac{\lambda_0}{1+y}-2\lambda_1+\lambda_3]=[0,\ 0]$;
2. $i)\ \lambda_1(x+2y-4)=0$; $ii)\ \lambda_2x=0$; $iii)\ \lambda_3y=0$;
3. $i)\ x+2y\leq 4$; $ii)\ x\geq 0$; $iii)\ y\geq 0$ and $\lambda_i\geq 0,\ i=1,2,3$;
4. $(\lambda_0,\lambda_1,\lambda_2,\lambda_3)\neq(0,0,0,0)$, where $\lambda_0\in\{0,1\}$.

To solve this set of (in)equalities we distinguish two cases.

Case 1: $\lambda_0=1$:

If $\lambda_0=1$ we have from item 1 that $\lambda_1=1+\lambda_2>0$ (as $\lambda_2\geq 0$) and $\lambda_3=2(1+\lambda_2)-\frac{1}{1+y}>0$ (as $y\geq 0$ and $\lambda_2\geq 0$). So, by item 2.i and iii), $x+2y=4$ and $y=0$. So, $(x,y)=(4,0)$ and $\lambda_2=0$. Consequently $(x,y,\lambda_0,\lambda_1,\lambda_2,\lambda_3)=(4,0,1,1,0,1)$ results as a candidate solution.

Case 2: $\lambda_0=0$:

Then, by item 1, $\lambda_2=\lambda_1$ and $\lambda_3=2\lambda_1$. Since not all λ_i are simultaneously zero it follows that they all differ from zero. So, $x=0$; $y=0$ and $x+2y=4$ should apply. Which is clearly not possible. So we don't get any additional candidate solutions here.

So, by Weierstrass' theorem, the maximum is attained at $(4,0)$.

b. Clearly $C(K,L)\in C^1$. Also $g_1(K,L):=16-L$, $g_2(K,L):=K-81$; $g_3(K,L):=2L-K$ and $h(K,L):=\sqrt{K}+\sqrt{L}-10$ are continuous differentiable. Consider $\Gamma:=\{(K,L)\mid g_i(x,y)\leq 0$ and $h(K,L)=0\}$. Then one easily verifies that Γ is compact. So, by Weierstrass' theorem the problem has a solution. To find the candidate solutions we proceed by considering the Lagrange function $\tilde{L}(K,L,\lambda_0,\lambda_1,\lambda_2,\lambda_3,\mu)$ given by

$$\tilde{L}:=-\lambda_0(rK+\omega L)-\lambda_1(16-L)-\lambda_2(K-81)-\lambda_3(2L-K)-\mu(\sqrt{K}+\sqrt{L}-10).$$

By Theorem 4.31 the candidate optimal solutions satisfy:

1. $[D_1\tilde{L}(K,L,\lambda)\ D_2\tilde{L}(K,L,\lambda)]=[-r\lambda_0-\lambda_2+\lambda_3-\frac{\mu}{2\sqrt{K}},\ -\omega\lambda_0+\lambda_1-2\lambda_3-\frac{\mu}{2\sqrt{L}}]=[0,\ 0]$;
2. $i)\ \lambda_1(16-L)=0$; $ii)\ \lambda_2(K-81)=0$; $iii)\ \lambda_3(2L-K)=0$;

3. $i)$ $L \geq 16$; $ii)$ $K \leq 81$; $iii)$ $K \geq 2L$; $iv)$ $\sqrt{K} + \sqrt{L} = 10$ and $\lambda_i \geq 0,\ i = 1, 2, 3$;
4. $(\lambda_0, \lambda_1, \lambda_2, \lambda_3) \neq (0, 0, 0, 0, 0)$, where $\lambda_0 \in \{0, 1\}$.

To solve this set of (in)equalities we distinguish two cases.

Case 1: $\lambda_0 = 1$:

1. no constraint is binding.
   Then, $\lambda_i = 0,\ i = 1, 2, 3$. From item 1. we get: $\mu = 2r\sqrt{K} = 2\omega\sqrt{L}$. Therefore $r\sqrt{K} = \omega\sqrt{L}$. Substitution into item 3.iv) yields: $(K, L) = (\frac{100\omega^2}{(\omega+r)^2}, \frac{100r^2}{(\omega+r)^2})$. So, provided item 3.i)-iii) are met, this provides a candidate solution. That is, provided i) $\omega \leq \frac{3}{2}r$; ii) $\omega \geq \frac{1}{9}r$ and iii) $\omega \geq \sqrt{2}r$ this yields a candidate solution.
2. precisely one constraint is binding:
   1. $g_1 = 0$. That is: $L = 16$ and $\lambda_2 = \lambda_3 = 0$. Then, from $h = 0$, $K = 36$ and $(K, L, \lambda_0, \lambda_1, \lambda_2, \lambda_3, \mu) = (36, 16, 1, \omega - \frac{3}{2}r, 0, 0, -12r)$. So, in case $\omega \geq \frac{3}{2}r$, this yields a candidate solution.
   2. $g_2 = 0$. That is: $K = 81$ and $\lambda_1 = \lambda_3 = 0$. Then, from $h = 0$, $L = 1$ which violates $L \geq 16$.
   3. $g_3 = 0$. That is: $K = 2L$ and $\lambda_1 = \lambda_2 = 0$. Then, from $h = 0$, $K = 2L$ with $L = \frac{100}{3+2\sqrt{2}}$. Elementary calculations show that $\lambda_3 \geq 0$ in this case iff $r\sqrt{K} + \omega\sqrt{L} \geq 0$. That is if $\omega \leq \sqrt{2}r$. So under this condition, this yields a candidate solution.
3. One easily verifies that two or more constraints simultaneously binding does not provide candidate solutions.

Case 2: $\lambda_0 = 0$:

1. no constraint is binding. Then, $\lambda_i = 0,\ i = 1, 2, 3$. This violates item 4.
2. precisely one constraint is binding:
   1. $g_1 = 0$. That is: $L = 16$ and $\lambda_2 = \lambda_3 = 0$. Then, from $h = 0$, $K = 36$. However, by item 1, $\mu = 0$ which implies $\lambda_1 = 0$ too. Which violates item 4.
   2. $g_2 = 0$. That is: $K = 81$ and $\lambda_1 = \lambda_3 = 0$. Similarly as above it follows then that $\lambda_2 = 0$ too, which violates item 4 again.
   3. $g_3 = 0$. That is: $K = 2L$ and $\lambda_1 = \lambda_2 = 0$. Also in this case we obtain similarly as above that item 4 is violated.
3. One easily verifies that two or more constraints simultaneously binding does not provide candidate solutions.

So we don't get any additional candidates here.

So, by Weierstrass' theorem, the minimum is attained at $(K, L) = (36, 16)$ if $\omega \geq \frac{3}{2}r$; at $(K, L) = (\frac{100\omega^2}{(\omega+r)^2}, \frac{100r^2}{(\omega+r)^2})$ if $\sqrt{2}r \leq \omega \leq \frac{3}{2}r$ and at $(K, L) = (\frac{200}{3+2\sqrt{2}}, \frac{100}{3+2\sqrt{2}})$ if $\omega \leq \sqrt{2}r$.

c. Clearly $U(\omega_1, \omega_2)$; $g_1(\omega_1, \omega_2) := -\omega_1$, $g_2(\omega_1, \omega_2) := -\omega_2$ and $g_3(\omega_1, \omega_2) := \omega_1 + \omega_2$ are continuous differentiable. Consider $\Gamma := \{(\omega_1, \omega_2) \mid g_i(\omega_1, \omega_2) \leq 0\}$. Then one easily verifies that Γ is compact. So, by Weierstrass' theorem the problem has a solution. To find the candidate solutions we proceed by considering the Lagrange function: $L(\omega_1, \omega_2, \lambda_0, \lambda_1, \lambda_2, \lambda_3)$ given by

$$L := \lambda_0(2\omega_1 + 3\omega_2 - \frac{1}{2}(\omega_1 + \omega_2)^2 + \lambda_1\omega_1 + \lambda_2\omega_2 - \lambda_3(\omega_1 + \omega_2 - 1).$$

By Theorem 4.31 the candidate optimal solutions satisfy:

1. $[D_1L(\omega_1,\omega_2,\lambda)\ D_2L(\omega_1,\omega_2,\lambda)] = [\lambda_0(2-\omega_1-\omega_2)+\lambda_1-\lambda_3,\ \lambda_0(3-\omega_1-\omega_2)+\lambda_2-\lambda_3] = [0,\ 0]$;
2. $i)\ \lambda_1\omega_1 = 0$; $ii)\ \lambda_2\omega_2 = 0$; $iii)\ \lambda_3(\omega_1+\omega_2-1) = 0$;
3. $i)\ \omega_1 \geq 0$; $ii)\ \omega_2 \geq 0$; $iii)\ \omega_1+\omega_2 \leq 1$ and $\lambda_i \geq 0,\ i = 1,2,3$;
4. $(\lambda_0,\lambda_1,\lambda_2,\lambda_3) \neq (0,0,0,0)$, where $\lambda_0 \in \{0,1\}$.

To solve this set of (in)equalities we distinguish two cases.

Case 1: $\lambda_0 = 1$:

1. no constraint is binding. Then, $\lambda_i = 0,\ i = 1,2,3$. From item 1. we get then an inconsistent set of linear equations. So, this does not provide a candidate solution.
2. precisely one constraint is binding:
   1. $g_1 = 0$. That is: $\omega_1 = 0$ and $\lambda_2 = \lambda_3 = 0$. Then by item 1. $\omega_2 = 3$, which violates item 3.iii). So, we don't obtain a feasible solution.
   2. $g_2 = 0$. That is: $\omega_2 = 0$ and $\lambda_1 = \lambda_3 = 0$. Then, by item 1. $\omega_1 = 2$ and $\lambda_2 = -1$. So, we don't obtain a feasible solution.
   3. $g_3 = 0$. That is: $\omega_1+\omega_2 = 1$ and $\lambda_1 = \lambda_2 = 0$. Then, item 1. yields an inconsistent set of equations. So, we don't obtain a feasible solution.
3. precisely two constraints are binding:
   1. $g_1 = 0$ and $g_2 = 0$. That is: $\omega_1 = \omega_2 = 0$ and $\lambda_3 = 0$. Then, item 1. yields an inconsistent set of equations. So, we don't obtain a feasible solution.
   2. $g_1 = 0$ and $g_3 = 0$. That is: $\omega_1 = \omega_1+\omega_2-1 = 0$ and $\lambda_2 = 0$. Then, we obtain the candidate solution: $(\omega_1,\omega_2,\lambda_0,\lambda_1,\lambda_2,\lambda_3) = (0,1,1,0,0,1)$.
   3. $g_2 = 0$ and $g_3 = 0$. That is: $\omega_2 = \omega_1+\omega_2-1 = 0$ and $\lambda_1 = 0$. Then, item 1. yields $\lambda_2 = -1$. So, we don't obtain a feasible solution.

Case 2: $\lambda_0 = 0$:

Then, by item 1., $\lambda_1 = \lambda_2 = \lambda_3$. So, in order to not violate item 4., all constraints must be simultaneously binding. Which is clearly not possible. So, this does not provide a candidate solution.

Thus, we find only one candidate solution $(\omega_1,\omega_2) = (0,1)$.

Therefore, by Weierstrass' theorem, the maximum is attained at $(\omega_1,\omega_2) = (0,1)$.

d. As $x > 0$ and $y > 0$ it follows that $U(x,y) \in C^1$ on the constraint set. Also $g_1(x,y) := \frac{1}{8}I - x$, $g_2(x,y) := \frac{1}{10} - y$ and $h(x,y) := 4x+3y-I$ are continuous differentiable. Consider $\Gamma := \{(x,y) \mid g_i(x,y) \leq 0 \text{ and } h(x,y) = 0\}$. Then one easily verifies that Γ is compact. So, by Weierstrass' theorem the problem has a solution. To find the candidate solutions we proceed by considering the Lagrange function $L(x,y,\lambda_0,\lambda_1,\lambda_2,\mu)$ given by:

$$L := \lambda_0 x^{\frac{1}{2}}y^{\frac{1}{4}} - \lambda_1(\frac{1}{8}I - x) - \lambda_2(\frac{1}{10} - y) - \mu(4x+3y-I).$$

By Theorem 4.31 the candidate optimal solutions satisfy:

1. $[D_1L(x,y,\lambda)\ D_2L(x,y,\lambda)] = [\frac{1}{2}\lambda_0 x^{-\frac{1}{2}}y^{\frac{1}{4}} + \lambda_1 - 4\mu,\ \frac{1}{4}\lambda_0 x^{\frac{1}{2}}y^{-\frac{3}{4}} + 2\lambda_2 - 3\mu] = [0,\ 0]$;

2. $i)$ $\lambda_1(\frac{1}{8}I - x) = 0$ and $ii)$ $\lambda_2(\frac{1}{10}I - y) = 0$;
3. $i)$ $x \geq \frac{1}{8}I$; $ii)$ $y \geq \frac{1}{10}I$; $iii)$ $4x + 3y = I$ and $\lambda_i \geq 0,\ i = 1, 2$;
4. $(\lambda_0, \lambda_1, \lambda_2) \neq (0, 0, 0)$, where $\lambda_0 \in \{0, 1\}$.

To solve this set of (in)equalities we distinguish two cases.

Case 1: $\lambda_0 = 1$:

1. no constraint is binding. Then, $\lambda_i = 0,\ i = 1, 2$. From item 1. we get: $\mu = \frac{1}{8}x^{-\frac{1}{2}}y^{\frac{1}{4}} = \frac{1}{12}x^{\frac{1}{2}}y^{-\frac{3}{4}}$. So, $y = \frac{2}{3}x$. Substitution into item 3.iii) yields $(x, y) = (\frac{I}{6}, \frac{I}{9})$. This solution satisfies all constraints, so it is a candidate solution.
2. precisely one constraint is binding:
   1. $g_1 = 0$. That is: $x = \frac{1}{8}I$ and $\lambda_2 = 0$. Then by item 3.iii $(x, y) = (\frac{I}{8}, \frac{I}{6})$. But this gives rise to $\lambda_1 < 0$. So, we don't obtain a feasible solution.
   2. $g_2 = 0$. That is: $y = \frac{1}{10}I$ and $\lambda_1 = 0$. Then, by item 3.iii $(x, y) = (\frac{7I}{40}, \frac{I}{10})$. This gives rise to $\lambda_2 = x^{-\frac{1}{2}}y^{-\frac{3}{4}}(\frac{3}{8}y - \frac{1}{4}x) < 0$. So, we don't obtain a feasible solution.
3. precisely two constraints are binding:

   This yields a set of inconsistent equations. So, we don't obtain a feasible solution.

Case 2: $\lambda_0 = 0$:

Then, by item 1., $\lambda_1 = 4\mu$ and $\lambda_2 = 3\mu$. By item 4, it is then not possible that $\mu = 0$. So all constraints must be binding, which yields an inconsistent set of equations. So, we don't obtain a feasible solution.

Thus, we find only one candidate solution: $(x, y) = (\frac{I}{6}, \frac{I}{9})$.

Therefore, by Weierstrass' theorem, the maximum is attained at $(x, y) = (\frac{I}{6}, \frac{I}{9})$.

4.7 A geometric approach to solve the problem

EXERCISE 4.8

Use Theorem 4.43 to determine the tangent cone for next sets at the indicated point a.

a. $S = \{(x, y, z) \mid |x + y| \leq 1;\ x^2 + y^2 + z^2 \leq 1; x + z \leq 1 \text{ and } |y + z| \leq 1\}$ at $(0, 1, 0)$.

b. $S = \{(x, y) \mid x^2 + (y - 1)^2 \leq 2;\ x^2 - y^2 \leq 1 \text{ and } x \leq 1\}$ at $(1, 0)$.

c. $S = \{(x, y) \mid x^2 + y^2 \leq 1; (x - 1)^2 + y^2 \leq 2;\ \text{and } x - y \leq -1\}$ at $(0, 1)$.

ANSWER 4.8

a. Note that at $(0, 1, 0)$ all constraints, except $x + z \leq 1$ are binding. Let $g_1(x, y, z) := x + y = 1$; $g_2(x, y, z) := x^2 + y^2 + z^2 = 1$ and $g_3(x, y, z) := y + z = 1$.

Then $Dg(x, y, z) = \begin{bmatrix} 1 & 1 & 0 \\ 2x & 2y & 2z \\ 0 & 1 & 1 \end{bmatrix}$. So, $Dg(0, 1, 0) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$, which is invertible and thus full row rank. So, by Theorem 4.43, item 1, $T_S(0, 1, 0) = TL_S(0, 1, 0) = \{u \mid [1\ 1\ 0]u \leq$

$0;\ 2[0\ 1\ 0]u \leq 0$ and $[0\ 1\ 1]u \leq 0\}$. That is: $TL_S(0,1,0) = \{[u_1\ u_2\ u_3]^T \mid u_1 + u_2 \leq 0;\ u_2 \leq 0 \text{ and } u_2 + u_3 \leq 0\}$.

b. Note that all constraints are binding at $(1,0)$. Let $g_1(x,y) := x^2+(y-1)^2 = 2$; $g_2(x,y) := x^2-y^2 = 1$ and $g_3(x,y) := x = 1$.
Then $Dg(x,y) = \begin{bmatrix} 2x & 2(y-1) \\ 2x & -2y \\ 1 & 0 \end{bmatrix}$. So, $Dg(1,0) = \begin{bmatrix} 2 & -2 \\ 2 & 0 \\ 1 & 0 \end{bmatrix}$. Note that $Dg(1,0)$ is a 3×2 matrix. So, it can never have full row rank. Consequently, Theorem 4.43, item 1, is not applicable. Also, item 2 of this theorem does not apply since g_2 is not convex.
Note that $Dg(1,0)[-1\ 0]^T < 0$. So, item 3 of this theorem does apply. So, the tangent cone of this set is found as the linearized cone. That is, $T_S(1,0) = TL_S(1,0) = \{[u_1\ u_2]^T \mid u_1 \leq 0 \text{ and } u_2 \geq u_1\}$.

c. Note that all constraints are binding at $(0,1)$. Let $g_1(x,y) := x^2 + y^2 = 1$; $g_2 := (x-1)^2 + y^2 = 2$ and $g_3(x,y) := x - y = -1$.
Then $Dg(x,y) = \begin{bmatrix} 2x & 2y \\ 2(x-1) & 2y \\ 1 & -1 \end{bmatrix}$. So, $Dg(0,1) = \begin{bmatrix} 0 & 2 \\ -2 & 2 \\ 1 & -1 \end{bmatrix}$. Note that $Dg(0,1)$ is a 3×2 matrix. So, it can never have full row rank. Consequently, Theorem 4.43, item 1, is not applicable. Also, item 3 of this theorem does not apply since $Dg_2(0,1)u = -Dg_3(0,1)u$.
Note that all constraints $g_i(x,y)$ are convex functions. So, item 2 of this theorem does apply. So, the tangent cone of this set is found as the linearized cone. That is, $T_S(0,1) = TL(0,1) = \{[u_1\ u_2]^T \mid u_1 = u_2 \text{ and } u_2 \leq 0\}$.

Mixed Exercises

EXERCISE 4.9 Use the theory presented in this chapter to solve next constrained optimization problems.

a. Maximize $f(x,y) = x^3 + y^3$ subject to $|x| \leq 1$ and $|y| \leq 1$.
b. Maximize $f(x,y) = x + y$ subject to $|xy| \leq 1$.
c. Maximize $f(x_1,x_2) = x_1x_2$ subject to $g(x_1,x_2) = x_1 + 8x_2 - 4 \geq 0$ and $h(x_1,x_2) = x_1^2 + x_2^2 - 1 = 0$.
d. Minimize $f(x) = \sqrt{1-x^2},\ x \in [-1,1]$ subject to $g(x) = x \leq 1$.
e. Minimize $f(x,y) = (x+2)^2 + y^2$ subject to $y \geq |x|$. Hint: first rewrite the constrained set as the intersection of two sets $\Gamma_i := \{(x,y) \mid g_i(x,y) \leq 0\}$, where the functions g_i are differentiable.
f. Minimize $f(x,y) = e^{x+y}$ subject to $x^2 \geq 1$ and $y^2 \geq 1$.
g. Minimize $f(x,y) = e^{x+y^2}$ subject to $x^2 \leq 1$ and $y^2 \geq 1$.

ANSWER 4.9

a. Note that f is a continuous function. Furthermore, the constraint set is compact. So, by Weierstrass' theorem, the problem has a global maximum. We next use Algoritm 4.20 to solve the problem.

Step 1. Let $g_1(x,y) := x-1$; $g_2(x,y) := -x-1$; $g_3(x,y) := y-1$ and $g_4(x,y) := -y-1$. Then f and $g_i \in C^1,\ i = 1, \cdots, 4$.

Step 2. $Dg_1 = [1\ 0]$; $Dg_2 = [-1\ 0]$; $Dg_3 = [0\ 1]$ and $Dg_4 = [0\ -1]$. Note that neither g_1 and g_2 nor g_3 and g_4 can be binding at the same time. Using this observation one easily verifies then that if either one or two constraints are simulatenously binding, the corresponding matrix of derivatives has full row rank. So we don't get any candidates for solutions.

Step 3. Consider $L(x,y,\lambda_1,\lambda_2,\lambda_3,\lambda_4)$ given by

$$L := x^3 + y^3 - \lambda_1(x-1) - \lambda_2(-x-1) - \lambda_3(y-1) - \lambda_4(-y-1).$$

Step 4. Then next conditions must be satisfied:

$$\text{i) } 3x^2 - \lambda_1 + \lambda_2 = 0;\ \text{ii) } 3y^2 - \lambda_3 + \lambda_4 = 0;\ \text{iii) } \lambda_i g_i = 0,\ i = 1, \cdots, 4;$$
$$\text{iv) } g_i \leq 0 \text{ and v) } \lambda_i \geq 0,\ i = 1, \cdots, 4.$$

1. Assume no constraint is binding. Consequently, $\lambda_i = 0,\ i = 1, \cdots, 4$. $(x,y,\lambda_1,\lambda_2,\lambda_3,\lambda_4) = (0,0,0,0,0,0)$ is a candidate solution.
2. Assume precisely one constraint is binding:
   1. $x = 1$ and, consequently, $\lambda_i = 0,\ i = 2,3,4$. Then $(x,y,\lambda_1,\lambda_2,\lambda_3,\lambda_4) = (1,0,3,0,0,0)$ is a candidate solution.
   2. $x = -1$ and, consequently, $\lambda_i = 0,\ i = 1,3,4$. Then there is no solution to (i-v).
   3. $y = 1$ and, consequently, $\lambda_i = 0,\ i = 1,2,4$. Then $(x,y,\lambda_1,\lambda_2,\lambda_3,\lambda_4) = (0,1,0,0,3,0)$ is a candidate solution.
   4. $y = -1$ and, consequently, $\lambda_i = 0,\ i = 1,2,3$. Then there is no solution to (i-v).
3. Assume precisely two constraints are binding:
   1. $x = 1,\ y = 1$ and, consequently, $\lambda_i = 0,\ i = 2,4$. Then $(x,y,\lambda_1,\lambda_2,\lambda_3,\lambda_4) = (1,1,3,0,3,0)$ is a candidate solution.
   2. $x = 1,\ y = -1$ and, consequently, $\lambda_i = 0,\ i = 2,3$. Then there is no solution to (i-v).
   3. $x = -1,\ y = 1$ and, consequently, $\lambda_i = 0,\ i = 1,4$. Then there is no solution to (i-v).
   4. $x = -1,\ y = -1$ and, consequently, $\lambda_i = 0,\ i = 1,3$. Then there is no solution to (i-v).

So, we obtain 4 candidate solutions.

Evaluation of f at these solutions shows that the maximum is attained at $(x,y) = (1,1)$.

b. We use Algoritm 4.20 to solve the problem.

Step 1. Let $g_1(x,y) := xy - 1$ and $g_2(x,y) := -xy - 1$. Then f and $g_i \in C^1,\ i = 1,2$.

Step 2. $Dg_1 = [y\ x]$ and $Dg_2 = [-y\ -x]$. Note that g_1 and g_2 cannot be simultaneously binding. Using this observation one easily verifies that if either one of the constraints is binding, the

corresponding matrix of derivatives has not full row rank iff $x = y = 0$. But $(0, 0)$ is not binding. So we don't get any candidates for solutions.

Step 3. Let $L(x, y, \lambda_1, \lambda_2) = x + y - \lambda_1(xy - 1) - \lambda_2(-xy - 1)$.

Step 4. Then next conditions must be satisfied:

$$\text{i) } 1 - \lambda_1 y + \lambda_2 y = 0;\ \text{ii) } 1 - \lambda_1 x + \lambda_2 x = 0;\ \text{iii) } \lambda_i g_i = 0,\ i = 1, 2;$$
$$\text{iv) } g_i \leq 0 \text{ and v) } \lambda_i \geq 0,\ i = 1, 2.$$

1. Assume no constraint is binding. Consequently, $\lambda_1 = \lambda_2 = 0$. Then there is no solution to (i-v).
2. Assume precisely one constraint is binding:
   1. $xy = 1$ and, consequently, $\lambda_2 = 0$. Then $(x, y, \lambda_1, \lambda_2) = (1, 1, 1, 0)$ is a candidate solution.
   2. $xy = -1$ and, consequently, $\lambda_1 = 0$. Then there is no solution to (i-v).
3. Note that not more than one constraint can be binding.

So, we obtain 1 candidate solution $(x, y, \lambda_1, \lambda_2) = (1, 1, 1, 0)$.

Note that $HL(1, 1, 1, 0) = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$. Consequently,

$$[1\ -\frac{D_x g_1(1,1)}{D_y g_1(1,1)}]HL(1,1,1,0)[1\ -\frac{D_x g_1(1,1)}{D_y g_1(1,1)}]^T > 0.$$

So, by Theorem 4.19, f has not a local maximum at $(1, 1)$. Note that with $y = 0$ and x arbitrarily large f attains an arbitrarily large value. So, the problem has no solution.

c. Note, f and g are twice differentiable and $\Gamma := \{(x_1, x_2) \mid g(x_1, x_2) \geq 0 \text{ and } h(x_1, x_2) = 0\}$ is a compact set. Therefore, the problem has by Weierstrass' theorem a solution.

Consider $L(x_1, x_2, \lambda_0, \lambda_1, \mu) := \lambda_0 x_1 x_2 - \lambda_1(4 - x_1 - 8x_2) - \mu(x_1^2 + x_2^2 - 1)$.

By Theorem 4.31 the candidate optimal solutions satisfy:

1. $[D_1 L(x_1, x_2, \lambda_0, \lambda_1, \mu)\ D_2 L(x, y, \lambda_0, \lambda_1, \mu)] = [\lambda_0 x_2 + \lambda_1 - 2\mu x_1,\ \lambda_0 x_1 + 8\lambda_1 - 2\mu x_2] = [0\ 0]$;
2. $\lambda_1(4 - x_1 - 8x_2) = 0$;
3. *i)* $x_1 + 8x_2 \geq 4$; *ii)* $x_1^2 + x_2^2 = 1$ and *iii)* $\lambda_1 \geq 0$;
4. $(\lambda_0, \lambda_1) \neq (0, 0)$, where $\lambda_0 \in \{0, 1\}$.

To solve this set of (in)equalities we distinguish two cases.

Case 1: $\lambda_0 = 1$:

1. g is not binding. Then, $\lambda_1 = 0$. Simple calculations show we obtain next two candidate solutions: $(x_1, x_2, \lambda_0, \lambda_1, \mu) = (\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}, 1, 0, \frac{1}{2})$, yielding a value of f of $\frac{1}{2}$, and $(x_1, x_2, \lambda_0, \lambda_1, \mu) = (-\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}, 1, 0, -\frac{1}{2})$, yielding a value of f of $-\frac{1}{2}$.
2. g is binding. That is: $x_1 + 8x_2 = 4$. Substitution into $h = 0$ yields $(x_1, x_2) = (\frac{-104}{130}, \frac{78}{130})$ and $(x_1, x_2) = (\frac{120}{130}, \frac{50}{130})$. Without verifying whether $\lambda_1 \geq 0$, clearly, both these points will not solve the problem since the candidate solution we obtained in item 1.i) yields a higher value for $f(x_1, x_2)$.

Case 2: $\lambda_0 = 0$. Then, by item 4., $\lambda_1 \neq 0$. So, g must be binding. Consequently, we reobtain the solutions already advertized in Case 1(b). As we already noticed there, these points don't provide the solution for the problem.

So, the solution for the problem is provided by $(x_1, x_2, \lambda_0, \lambda_1, \mu) = (\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}, 1, 0, \frac{1}{2})$.

d. Note, f and g are twice differentiable on $(-1, 1)$. Furthermore, f is continuous on the compact set $[-1, 1]$. Therefore, the problem has by Weierstrass' theorem a solution.

We follow Algorithm 4.20 to find the solution.

Step 1. f is not differentiable at $x = 1$ and $x = -1$. Therefore, this are candidate solutions for a minimum.

Step 2. $g^{'}(x) = 1$. So, no additional candidate solutions occur.

Step 3. Consider $L(x, \lambda) := -\sqrt{1 - x^2} - \lambda(x - 1)$.

Step 4. Then on the interval $(-1, 1)$ additional candidate solutions must satisfy:

$$\text{i) } \frac{x}{\sqrt{1 - x^2}} - \lambda = 0; \text{ ii) } \lambda(1 - x) = 0 \text{ and iii) } \lambda \geq 0.$$

As $x < 1$, by ii), $\lambda = 0$. Which implies by i) that $x = 0$. So, $(x, \lambda) = (0, 0)$ is a candidate solution.

So, there are three candidate solutions: $x = 1$, $x = -1$ and $x = 0$.

Evaluation of f at the candidate solutions $x = 1$, $x = -1$ and $x = 0$ shows that the problem has a global minimum at $x = 1$ and $x = -1$.

e. Step 1: Note, $g(x, y) := |x| - y$ is not differentiable at $x = 0$. However, the constraint set can be rewritten as the intersection of the following constraint sets:

$$\Gamma_1 := \{(x, y) \mid g_1(x, y) := x - y \leq 0\} \text{ and } \Gamma_2 := \{(x, y) \mid g_2(x, y) := -x - y \leq 0\}.$$

For, let $(x, y) \in \Gamma := \{(x, y) \mid |x| \leq y\}$. Assume, $x \geq 0$. Then it follows that $x - y \leq 0$. So $(x, y) \in \Gamma_1$. However, since $x \geq 0$ it follows that $-x - y \leq x - y \leq 0$. That is, $(x, y) \in \Gamma_2$ too. Similarly it follows that if $x \leq 0$, (x, y) belongs to both sets too. Which shows that $\Gamma \subset \Gamma_1 \cap \Gamma_2$.

That also $\Gamma_1 \cap \Gamma_2 \subset \Gamma$ follows straightforwardly.

Thus, the problem is equivalent to minimizing $-(x + 2)^2 - y^2$ subject to $g_i(x, y) \leq 0$.

We follow Algorithm 4.20 to find its solution.

Step 1. From the partial derivatives it is clear that f and g_i are continuous differentiable.

Step 2. $g_1^{'} = [1 \ -1]$ and $g_2^{'} = [-1 \ -1]$. So, $[Dg_1 \ Dg_2]^T$ has always full row rank.

Step 3. Consider $L(x, y, \lambda_1, \lambda_2) := -(x + 2)^2 - y^2 - \lambda_1(x - y) + \lambda_2(x + y)$.

Step 4. Then the candidate optimal solutions satisfy:

1. $[D_1 L(x, y, \lambda_1, \lambda_2) \ D_2 L(x, y, \lambda_1, \lambda_2)] = [-2(x + 2) + \lambda_1 + \lambda_2, \ -2y - \lambda_1 + \lambda_2] = [0 \ 0]$;
2. $i)$ $\lambda_1(x - y) = 0$ and $ii)$ $\lambda_2(x + y) = 0$;
3. $i)$ $y \geq x$; $ii)$ $y \geq -x$ and $iii)$ $\lambda_i \geq 0, \ i = 1, 2$.

Case 1: no constraint is binding: so, $\lambda_i = 0$. Then we have from item 1. that $(x, y) = (-2, 0)$, which violates item 3.ii).

Case 2: precisely one constraint is binding.

1. $g_1 = 0$ and $\lambda_2 = 0$. Then we obtain $\lambda_1 < 0$, which violates item 3.iii).
2. $g_2 = 0$ and $\lambda_1 = 0$. Then we get $(x, y) = (-1, 1)$ with $\lambda_2 = 2$.

Case 3: both constraints are binding. Then we get $(x, y) = (0, 0)$. But this yields $\lambda_1 = -2$ in item 1., which violates item 3.iii).

So, the only candidate solution is $(x, y) = (-1, 1)$ with $(\lambda_1, \lambda_2) = (0, 2)$.

Note that $L(x, y, 0, 2) = -(x + 2)^2 - y^2 + 2(x + y)$. Simple calculations show that $HL(x, y, 0, 2)$ is negative semi-definite. So $L(x, y, 0, 2)$ is concave on $\mathbb{R}^2$. Therefore, the problem has a global minimum at $(x, y) = (-1, 1)$.

f. We follow Algorithm 4.20 to find the solution.

Step 1. Consider $g_1(x, y) := 1 - x^2$ and $g_2(x, y) := 1 - y^2$. From the partial derivatives it is clear that f and g_i are continuous differentiable.

Step 2. $g_1^{'} = [-2x\ 0]$ and $g_2^{'} = [0\ -2y]$. So, $[Dg_1\ Dg_2]^T$ has always full row rank provided either $x \neq 0$ or $y \neq 0$. However, $(0, y)$, with $y^2 \geq 1$, and $(x, 0)$, with $x^2 \geq 0$ are not feasible.

Step 3. Consider $L(x, y, \lambda_1, \lambda_2) := -e^{x+y} - \lambda_1(1 - x^2) - \lambda_2(1 - y^2)$.

Step 4. Then the candidate optimal solutions satisfy:

1. $[D_1L(x, y, \lambda_1, \lambda_2)\ D_2L(x, y, \lambda_1, \lambda_2)] = [-e^{x+y} + 2\lambda_1 x,\ -e^{x+y} + 2\lambda_2 y] = [0\ 0]$;
2. $i)\ \lambda_1(1 - x^2) = 0$ and $ii)\ \lambda_2(1 - y^2) = 0$;
3. $i)\ x^2 \geq 1$; $ii)\ y^2 \geq 1$ and $iii)\ \lambda_i \geq 0,\ i = 1, 2$.

Case 1: no constraint is binding: so, $\lambda_i = 0$. Then item 1. has no solution.

Case 2: precisely one constraint is binding.

1. $g_1 = 0$ and $\lambda_2 = 0$. Then item 1. has no solution.
2. $g_2 = 0$ and $\lambda_1 = 0$. Then item 1. has no solution.

Case 3: both constraints are binding. Then we get $(x, y) = (\pm 1, \pm 1)$. If either $x = -1$ or $y = -1$ item 1. has no solution. If $(x, y) = (1, 1)$, it follows that $(\lambda_1, \lambda_2) = (\frac{1}{2}e^2, \frac{1}{2}e^2)$. So, this is a candidate solution.

So, the only candidate solution is: $(x, y, \lambda_1, \lambda_2) = (1, 1, \frac{1}{2}e^2, \frac{1}{2}e^2)$ with $f(1, 1) = e^2$.

Note that $f(x, y) \geq f(1, 1)$ for all $(x, y) \in \{(x, y) \mid x \geq 1 \text{ and } y \geq 1\}$. So f has a global minimum at $(1, 1)$ if one considers f on this domain. However, note that by choosing an arbitrarily small number a, $f(a, a) = e^{2a}$ can be made arbitrarily close to zero (although it never attains this value), and is therefore smaller than e^2. Therefore, the problem does not have a global minimum.

g. We follow Algorithm 4.20 to find the solution of this problem.

Step 1. Consider $g_1(x, y) := x^2 - 1$ and $g_2(x, y) := 1 - y^2$. From the partial derivatives it is clear that f and g_i are continuous differentiable.

Step 2. $g_1^{'} = [2x\ 0]$ and $g_2^{'} = [0\ -2y]$. So, $[Dg_1\ Dg_2]^T$ has always full row rank provided either $x \neq 0$ or $y \neq 0$. So, $(0, y)$, with $y^2 \geq 1$, are candidate solutions. Note that $(x, 0)$, with $x^2 \geq 0$ are not feasible.

Step 3. Consider $L(x, y, \lambda_1, \lambda_2) := -e^{x+y^2} - \lambda_1(x^2 - 1) - \lambda_2(1 - y^2)$.

Step 4. Then the candidate optimal solutions satisfy:

1. $[D_1L(x,y,\lambda_1,\lambda_2)\ D_2L(x,y,\lambda_1,\lambda_2)] = [-e^{x+y^2} - 2\lambda_1 x,\ -2ye^{x+y^2} + 2\lambda_2 y] = [0\ 0]$;
2. $i)\ \lambda_1(1-x^2) = 0$ and $ii)\ \lambda_2(1-y^2) = 0$;
3. $i)\ x^2 \leq 1$; $ii)\ y^2 \geq 1$ and $iii)\ \lambda_i \geq 0,\ i = 1,2$.

From $D_2L = 0$ it follows, since $y^2 \geq 1$, that $\lambda_2 = e^{x+y^2} > 0$. So, by item 2.ii), $y^2 = 1$.

From $D_1L = 0$ it follows that $x^2 < 1$ (and therefore $\lambda_1 = 0$) is not feasible. So, $x^2 = 1$. By item 3.iii) it follows then from $D_1L = 0$, moreover, that $x = -1$ and $\lambda_1 = \frac{1}{2}e^{x+y^2}$.

So, we obtain the candidate solutions: $(0,y)$, with $y^2 \geq 1$; $(x,y,\lambda_1,\lambda_2) = (-1,1,\frac{1}{2},1)$ and $(x,y,\lambda_1,\lambda_2) = (-1,-1,\frac{1}{2},1)$.

The value f attains at these candidate solutions is: $f(0,y) = e^{y^2}$; $f(-1,1) = f(-1,-1) = 1$. Note that $x + y^2 \geq 0$. Therefore $f(x,y) \geq 1$ for all (x,y) on its domain. So f has a global minimum at $(-1,1)$ and $(-1,-1)$.

One can also show this by considering $HL(x,y,\frac{1}{2},1)$. Using the fact that $x + y^2 \geq 0$ it follows straightforwardly that this matrix is negative definite on the set $D_1 := \{(x,y) \mid x \in [-1,1],\ y \geq 1\}$. One easily verifies that then also an open, convex set $\tilde{D}_1$ around this set exists where $HL(x,y,\frac{1}{2},1)$ is positive definite. This implies $L(x,y,\frac{1}{2},1)$ is concave on this convex, open set. So, by Proposition 4.11, the constraint optimization problem restricted to D_1 has a global minimum at $(-1,1)$. Similarly, it follows that also the constraint optimization problem restricted to $D_2 := \{(x,y) \mid x \in [-1,1],\ y \leq -1\}$ has a global minimum at $(-1,-1)$. As the values at both minimum locations coincide it follows that both candidate solutions provide a global minimum for the original problem.

EXERCISE 4.10 Solve next constrained optimization problems.

a. Maximize $f(x,y,z) = x\ln(xyz) - x - y - z$ s.t. $xyz \geq \frac{1}{2}$.
b. Maximize $f(x,y,z) = x^2 - y^2 + z^2$ s.t. $x^2 + y^2 + z^2 \leq 3$ and $xyz = 1$.
c. Find the point $(x,y) \in \mathbb{R}^2$ that has a minimal distance to the disc, with center $(0,0)$ and radius $\sqrt{2}$, and the hyperbola $cd = 9$. That is, minimize $(x-a)^2 + (y-b)^2 + (x-c)^2 + (y-d)^2$ w.r.t. $x,\ y,\ a,\ b,\ c$ and d s.t. $a^2 + b^2 \leq 2$ and $cd = 9$.
d. Maximize $f(x,y,z) = 14(x+y+z) - xy - xz - yz$ s.t. $x \leq 0$, $x + y = 2$ and $y + z^2 = 5$.
e. Minimize $f(x,y,z) = xy + xz + yz - 3x - 3y - 3z$ s.t. $x \leq 0$, $x - y = -1$ and $y + z^2 = 2$.
f. Minimize $f(x,y,z) = x^2 - xy - yz + z^2$ s.t. $x \geq 0$, $y \geq 0$, $z \geq 0$ and $y \leq x^2 + z^2$.
g. Maximize $f(x,y,z) = x^3 + y^3 + z^3$ s.t. $x^2 + y^2 \leq 1$ and $y^2 + z^2 \leq 1$.
h. Minimize $f(x,y,z) = x^2 + y^2 - z^2$ s.t. $x + y \geq 1$ and $y^2 + z^2 \leq 1$.

ANSWER 4.10

a. We follow Algorithm 4.20 to find the solution of the problem.

Step 1. Let $g(x,y,z) := \frac{1}{2} - xyz$. Then f and $g \in C^1$.

Step 2. $Dg = [yz\ xz\ xy]$. So Dg has not full row rank iff. $yz = xz = xy = 0$. But then at least one of the entries in (x,y,z) must be zero, which implies that $xyz = 0$. Which violates the constraint $g \leq 0$. So we don't get any candidates for solutions.

Step 3. Let $L(x,y,z,\lambda) = x\ln(xyz) - x - y - z\lambda(\frac{1}{2} - xyz)$.

Step 4. Then next conditions must be satisfied:

$$\text{i) } \ln(xyz) + \lambda yz = 0;\ \text{ii) } \frac{x}{y} - 1 + \lambda xz = 0;\ \text{iii) } \frac{x}{z} - 1 + \lambda xy = 0;$$
$$\text{iv) } \lambda(xyz - \frac{1}{2}) = 0;\ \text{v) } g \leq 0 \text{ and vi) } \lambda \geq 0.$$

1. Assume g is not binding. Consequently, $\lambda = 0$. Then by i) $xyz = 1$, ii) $x = y$ and iii) $x = z$. This yields the candidate solution $(1,1,1)$.
2. Next assume g is binding:
   Then $xyz = \frac{1}{2}$ and, consequently, i) $-\ln(2) + \frac{\lambda}{2x} = 0$; ii) $\frac{x}{y} - 1 + \frac{\lambda}{2y} = 0$ and iii) $\frac{x}{z} - 1 + \frac{\lambda}{2z} = 0$. Some elementary calculations yield then next candidate solution: $(x,y,z,\lambda) = (x,(1+\ln 2)x,(1+\ln 2)x,2\ln 2x)$, where $x = \left(\frac{1}{2(1+ln2)^2}\right)^{\frac{1}{3}}$.

So, we obtain 2 candidate solutions: $(x,y,z,\lambda) = (1,1,1,0)$ and $(x,y,z,\lambda) = (x,(1+\ln 2)x,(1+\ln 2)x,2\ln 2x)$, where $x = \left(\frac{1}{2(1+ln2)^2}\right)^{\frac{1}{3}}$

Note that

$$HL(x,y,z,\lambda) = \begin{bmatrix} \frac{1}{x} & \frac{1}{y} + \lambda z & \frac{1}{z} + \lambda y \\ \frac{1}{y} + \lambda z & -\frac{x}{y^2} & \lambda x \\ \frac{1}{z} + \lambda y & \lambda x & -\frac{x}{z^2} \end{bmatrix}.$$

From this we observe that

1. $HL(1,1,1,0)$ is indefinite (notice entry (1,1) is positive, whereas entry (2,2) is negative). So, $(1,1,1)$ is not a maximum location.
2. at $(x,(1+\ln 2)x,(1+\ln 2)x,2\ln 2x)$, with $x = \left(\frac{1}{2(1+ln2)^2}\right)^{\frac{1}{3}}$, the constraint is binding. So, with $Dg = [-yz\ -xz\ -xy] = (1+\ln 2)x^2[(1+\ln 2)\ 1\ 1]$, we have that $Dgv = 0$ iff. $v = [v_1\ -(1+\ln 2)v_1 - v3\ v_3]^T$, with $v_1, v_3 \in \mathbb{R}$. Therefore,

$$\begin{aligned} d(v) &= v^T HL(x,(1+\ln 2)x,(1+\ln 2)x,2\ln 2x)v \\ &= [v_1\ v_3] \begin{bmatrix} 1 & -(1+\ln 2) & 0 \\ 0 & -1 & 1 \end{bmatrix} HL \begin{bmatrix} 1 & 0 \\ -(1+\ln 2) & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_3 \end{bmatrix} \\ &= [v_1\ v_3] \begin{bmatrix} -6.61 & -1.952 \\ -1.952 & -2.3058 \end{bmatrix} \begin{bmatrix} v_1 \\ v_3 \end{bmatrix} < 0, \text{ for all } v_1, v_3 \in \mathbb{R}. \end{aligned}$$

   So, by Theorem 4.19, f has a (local) maximum at this location.

   Note that $f(1,-1,-1) = 1$. So, f has not a global maximum at this location.

b. We follow Algorithm 4.20 to find the solution of the problem.

   Step 1. Let $g(x,y,z) := x^2+y^2+z^2-3$ and $h(x,y,z) := xyz-1$. Then f and $g \in C^1$. Furthermore, the constraint set is compact. So, by Weierstrass' Theorem a global maximum exists.

Step 2. $Dg = [2x\ 2y\ 2z]$ and $Dh = [yz\ xz\ xy]$. Therefore, see Remark 4.28,

$$\tilde{G} = \begin{bmatrix} 2x & 2y & 2z & x^2+y^2+z^2-3 \\ yz & xz & xy & 0 \end{bmatrix}.$$

$\tilde{G}$ has not full row rank iff. $x^2+y^2+z^2-3=0$ and there exists a μ such that $x=\mu yz$, $y=\mu xz$ and $z=\mu xy$. Since none of the entries of (x,y,z) can be zero, it follows that $\tilde{G}$ has not full row rank iff. $x^2+y^2+z^2-3=0$ and $y^2=z^2=x^2$. Together with the restriction imposed by $h(x,y,z)$ we obtain then next candidate solutions: $(1,1,1)$; $(1,-1,-1)$; $(-1,-1,1)$ and $(-1,1,-1)$.

Step 3. Let $L(x,y,z,\lambda) = x^2-y^2+z^2-\lambda(x^2+y^2+z^2-3)-\mu(xyz-1)$.

Step 4. Then next conditions must be satisfied:

$$\text{i) } 2x(1-\lambda)-\mu yz=0;\ \text{ii) } 2y(1+\lambda)+\mu xz=0;\ \text{iii) } 2z(1-\lambda)-\mu xy=0;$$
$$\text{iv) } \lambda(x^2+y^2+z^2-3)=0;\ \text{v) } g\leq 0;\ h=0;\ \text{and vi) } \lambda\geq 0.$$

1. Assume g is not binding. Consequently, $\lambda=0$. Then by i) $2x-\mu yz=0$; ii) $2y+\mu xz=0$; iii) $2z-\mu xy=0$. From $h=0$, it follows that none of the entries of (x,y,z) can be zero. So, $\mu\neq 0$. From i) and ii) we have then $\mu=\frac{2x}{yz}=\frac{-2y}{xz}$. That is, $y^2=-x^2$. Which is not possible ($x\neq 0$ and $y\neq 0$).
2. Next assume g is binding and $\lambda\neq 0$ (the case $\lambda=0$ was in fact already considered above).
   1. First consider the case $\lambda=1$. Then, by i) and v), $\mu=0$. But then ii) is violated.
   2. Next consider the case $\lambda\neq 1$. Then from i) and iii) we obtain: $1-\lambda=\frac{\mu yz}{2x}=\frac{\mu xy}{2z}$. From which we deduce that: $x^2=z^2$. So, $x=z$ or $x=-z$.
      1. Case $x=z$. Then i) and ii) reduce to: i) $2x(1-\lambda-\frac{1}{2}\mu y)=0$; ii) $2y(1+\lambda)+\mu x^2=0$. Or: i) $1-\lambda-\frac{1}{2}\mu y=0$ and ii) $2y(1+\lambda)+\mu x^2=0$. Substitution of $\lambda=1-\frac{1}{2}\mu y$ from i) into ii) yields: $\mu(x^2-y^2)+4y=0$. From v) we obtain: $2x^2+y^2=3$ and $x^2y=1$. Which yields, with $y=\frac{1}{x^2}$, that x must satisfy: $2x^6-3x^4+1=0$. Some elementary analysis shows that $x=1$ and $x=-1$ are the only solutions to this equation. Which shows we do not get additional candidate solutions in this case.
      2. Case $x=-z$. Then i) and ii) reduce to: i) $2x(1-\lambda+\frac{1}{2}\mu y)=0$; ii) $2y(1+\lambda)-\mu x^2=0$. Or: i) $1-\lambda+\frac{1}{2}\mu y=0$ and ii) $2y(1+\lambda)-\mu x^2=0$. Substitution of $\lambda=1+\frac{1}{2}\mu y$ from i) into ii) yields: $\mu(-x^2+y^2)+4y=0$. From v) we obtain: $2x^2+y^2=3$ and $x^2y=-1$. Which yields, with $y=-\frac{1}{x^2}$, that x must satisfy: $2x^6-3x^4+1=0$. Similar to the case above, we see we don't get any additional candidate solutions here.

So, we find four candidate solutions (x,y,z): $(1,1,1)$; $(1,-1,-1)$; $(-1,-1,1)$ and $(-1,1,-1)$. Calculation of f at any of the candidate solutions found above yields a value of 1. So, at all of these locations f has a global maximum.

c. We follow Algorithm 4.20 to find the solution of the problem.

Step 1. Let $f(x,y,a,b,c,d) := -(x-a)^2-(y-b)^2-(x-c)^2-(y-d)^2$; $g(x,y,a,b,c,d) := a^2+b^2-2$ and $h(x,y,a,b,c,d) := cd-9$. Then $f,\ g$ and $h\in C^1$.

Step 2. $Dg = [0\ 0\ 2a\ 2b\ 0\ 0]$ and $Dh = [0\ 0\ 0\ 0\ d\ c]$.

So, see Remark 4.28,

$$\tilde{G} = \begin{bmatrix} 0 & 0 & 2a & 2b & 0 & 0 & a^2+b^2-2 & 0 \\ 0 & 0 & 0 & 0 & d & c & 0 & cd-9 \end{bmatrix}.$$

Given the constraints, one easily verifies that $\tilde{G}$ has always full row rank.

Step 3. Let $L(x, y, a, b, c, d, \lambda, \mu) = f - \lambda g - \mu h$. Then next conditions must be satisfied:

$$\text{i) } 2x - a - c = 0;\ \text{ii) } 2y - b - d = 0;\ \text{iii) } x - (1 + 2\lambda)a = 0;\ \text{iv) } y - (1 + 2\lambda)b = 0;$$
$$\text{v) } 2(x - c) - \mu d = 0;\ \text{vi) } 2(y - d) - \mu c = 0;\ \text{vii) } g \leq 0;\ h = 0;\ \text{and viii) } \lambda \geq 0.$$

1. Assume g is not binding. Consequently, $\lambda = 0$. Then by i)-iv), $a = c$ and $b = d$. But then $a,\ b$ should simultaneously satisfy $a^2 + b^2 < 2$ and $ab = 9$. Some elementary analysis shows this is not possible.
2. Next assume g is binding (and $\lambda > 0$, which case was basically already considered above).
   Then by i) and iii): $c = 2x - a = (1 + 4\lambda)a$.
   So, by vi): $\mu c = \mu(1 + 4\lambda)a = 2y - 2d = b - d = -4\lambda b$. (1)
   Similarly ii) and iv) yield: $d = 2y - b = (1 + 4\lambda)b$.
   So, by v) $\mu d = \mu(1 + 4\lambda)b = 2x - 2c = a - c = -4\lambda a$. (2)
   After subtracting (1) from (2) we obtain: $(\mu(1+4\lambda)-4\lambda)(b-a) = 0$. So either $\mu(1+4\lambda) = 4\lambda$, which implies by (2) that $a = -b$, or $a = b$. Since $a^2 + b^2 = 2$ it follows that $a = \pm b = \pm 1$. Furthermore $cd = (1 + 4\lambda)^2 ab = 9$. So, it follows that $a = b$, and $\lambda = \frac{1}{2}$. From which we obtain, using the other equations, that $c = d = 3$ with $x = y = 2$ if $a = 1$ and $c = d = -3$ with $x = y = -2$ if $a = -1$. Furthermore, $\mu = -\frac{2}{3}$.
   Therefore, we obtain the candidate solutions $(2, 2, 1, 1, 3, 3)$ and $(-2, -2, -1, -1, -3, -3)$ with both $(\lambda, \mu) = (\frac{1}{2}, -\frac{2}{3})$.

So, we find two candidate solutions.

Next consider $L(x, y, a, b, c, d, \frac{1}{2}, -\frac{2}{3})$. Simple differentiation shows that

$$HL = \begin{bmatrix} -4 & 0 & 2 & 2 & 0 & 0 \\ 0 & -4 & 0 & 2 & 0 & 2 \\ 2 & 0 & -3 & 0 & 0 & 0 \\ 0 & 2 & 0 & -3 & 0 & 0 \\ 2 & 0 & 0 & 0 & -2 & \frac{2}{3} \\ 0 & 2 & 0 & 0 & \frac{2}{3} & -2 \end{bmatrix}.$$

An inspection of the eigenvalues (using Matlab) from HL shows that HL is negative definite on $\mathbb{R}^6$. So, by Proposition 4.29, we conclude that at both locations f has a global maximum. Or, in

other words, at both $(x, y) = (2, 2)$ and $(x, y) = (-2, -2)$, the distance of (x, y) to both sets is minimal (in fact the minimal Euclidean distance is $\sqrt{2}$).

d. We follow Algorithm 4.20 to find the solution of the problem.

Step 1. Let $g(x, y, z) := x$, $h_1(x, y, z) := x + y - 2$ and $h_2(x, y, z) := y + z^2 - 5$. Then $f,\ g$ and $h_i,\ i = 1, 2, \in C^1$.

Furthermore, since $y = 5 - z^2$, it follows that $y \leq 5$. Moreover, since $x \leq 0$, it follows that $y = 2 - x \geq 2$. So, $y \in [2, 5]$. Which implies, as $z^2 = 5 - y$ and $x = 2 - y$, that both x and z are bounded too. Consequently, the constraint set is compact. Therefore, by Weierstrass' theorem, the problem has a solution.

Step 2. $Dg = [1\ 0\ 0]$, $Dh_1 = [1\ 1\ 0]$ and $Dh_2 = [0\ 1\ 2z]$.

So, see Remark 4.28,

$$\tilde{G} = \begin{bmatrix} 1 & 0 & 0 & x \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 2z & 0 \end{bmatrix}.$$

Using Gaussian elimination one easily verifies that the rank of $\tilde{G}$ coincides with the rank of matrix

$$\begin{bmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 0 & -x \\ 0 & 0 & 2z & x \end{bmatrix}.$$

So matrix $\tilde{G}$ has not full row rank iff $x = z = 0$. But an easy inspection of the constraints shows no point where $x = z = 0$ is located in the constraint set. Therefore, we don't get any candidate solutions in this step.

Step 3. Let $L(x, y, z, \lambda, \mu_1, \mu_2) := f - \lambda g - \mu_1 h_1 - \mu_2 h_2$.

Step 4. Then next conditions must be satisfied:

$$\text{i) } 14 - y - z - \lambda - \mu_1 = 0;\ \text{ii) } 14 - x - z - \mu_1 - \mu_2 = 0;\ \text{iii) } 14 - x - y - 2\mu_2 z = 0;$$
$$\text{iv) } \lambda x = 0;\ \text{v) } x \leq 0;\ \text{vi) } x + y = 2;\ \text{vii) } y + z^2 = 5;\ \text{and viii) } \lambda \geq 0.$$

1. Assume g is not binding, i.e., $x < 0$. Consequently, $\lambda = 0$. Then by i)-ii), $\mu_1 = 14 - y - z$ and $\mu_2 = y - x$. So, by iii), $14 - x - y - 2z(y - x) = 0$ (1). By vii), $y = 5 - z^2$ and by vi), $x = 2 - y = -3 + z^2$. Substitution of this into (1) shows z is the solution of $4z^3 - 16z + 12 = 0$. Some elementary inspection shows that the roots of this equation are: $z = 1$ and $z = \frac{-1 \pm \sqrt{13}}{2}$. From these 3 roots, $z = \frac{-1 - \sqrt{13}}{2}$ is not located in the constraint set. So, we obtain 2 candidate solutions: $(-2, 4, 1)$ and $(z^2 - 3, 5 - z^2, z)$, with $z = \frac{-1 + \sqrt{13}}{2}$.

2. Next assume g is binding, i.e., $x = 0$. Then, $y = 2$ and $z = \pm\sqrt{3}$. Using this in i)-iii) shows $\lambda = \frac{6}{z} - 2$. So, $z = -\sqrt{3}$ is not feasible.

Therefore we obtain 3 candidate solutions: $(-2,4,1)$, $(z^2-3,5-z^2,z)$, with $z=\frac{-1+\sqrt{13}}{2}$, and $(0,2,\sqrt{3})$.

Evaluation of f at these points shows that the maximum is attained at $(0,2,\sqrt{3})$.

e. With $\tilde{f}:=-f$ consider the maximization of $\tilde{f}$ subject to the constraints.

First, notice that the problem can be easily solved using substitution. One easily verifies that the maximization problem is, after substitution of both equality constraints, equivalent with the maximization of $\tilde{f}(z)=-z^4+2z^3-3z^2+7$, with $z^2\geq 1$.

Below, we elaborate this exercise using the Algorithm 4.20. This, to illustrate the advantage of using Theorem 4.47 compared to Theorem 4.30.

Step 1. Let $g(x,y,z):=x$, $h_1(x,y,z):=x-y+1$ and $h_2(x,y,z):=y+z^2-2$. Then $\tilde{f},\ g$ and $h_i,\ i=1,2,\in C^1$.

Step 2. $Dg=[1\ 0\ 0]$, $Dh_1=[1\ -1\ 0]$ and $Dh_2=[0\ 1\ 2z]$.

So, see Remark 4.28,

$$\tilde{G}=\begin{bmatrix}1&0&0&x\\1&-1&0&0\\0&1&2z&0\end{bmatrix}.$$

Using Gaussian elimination one easily verifies that the rank of $\tilde{G}$ coincides with the rank of matrix

$$\begin{bmatrix}1&0&0&x\\0&-1&0&-x\\0&0&2z&-x\end{bmatrix}.$$

So matrix $\tilde{G}$ has not full row rank iff $x=z=0$. But an easy inspection of the constraints shows no point where $x=z=0$ is located in the constraint set. Therefore, we don't get any candidate solutions in this step.

Step 3. Let $L(x,y,z,\lambda,\mu_1,\mu_2):=\tilde{f}-\lambda g-\mu_1 h_1-\mu_2 h_2$.

Step 4. Then next conditions must be satisfied:

$$\text{i) } 3-y-z-\lambda-\mu_1=0;\ \text{ii) } 3-x-z+\mu_1-\mu_2=0;\ \text{iii) } 3-x-y-2\mu_2 z=0;$$
$$\text{iv) } \lambda x=0;\ \text{v) } x\leq 0;\ \text{vi) } x-y=-1;\ \text{vii) } y+z^2=2;\ \text{and viii) } \lambda\geq 0.$$

1. Assume g is not binding, i.e., $x<0$. Consequently, $\lambda=0$. Then by i)-ii), $\mu_1=3-y-z$ and $\mu_2=6-x-y-2z$. So, by iii), $3-x-y-2z(6-x-y-2z)=0$ (1). By vii), $y=2-z^2$ and by vi), $x=-1+y=1-z^2$. Substitution of this into (1) shows z is the solution of $-4z^3+6z^2-6z=0$. Some elementary inspection shows that the this equation has only one real root: $z=0$. This gives rise to next point: $(1,2,0)$. But this point violates the constraint $x<0$. So we don't find any candidate solution here.
2. Next assume g is binding, i.e., $x=0$. Then, $y=1$ and $z=\pm 1$. Using this in i)-iii) shows $\lambda=2$ in case $z=1$ (with $\mu_2=1$), whereas $\lambda=9$ results in case $z=-1$ (with $\mu_2=-2$).

Therefore we obtain 2 candidate solutions: $a_1 := (0, 1, 1)$ and $a_2 := (0, 1, -1)$.
Evaluation of $\tilde{f}$ at these points shows: $\tilde{f}(a_1) = 5$ and $\tilde{f}(a_2) = 1$.
Next, consider HL. Simple calculations show that

$$HL = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & -2\mu_2 \end{bmatrix}.$$

Note that, at both locations, g is binding. So, Theorem 4.30 does not provide additional information about the optimality of both locations.
To demonstrate the use of Theorem 4.47, consider the nullspace of matrix

$$D[h_1\ h2]^T = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 2z \end{bmatrix}.$$

One easily verifies that $D[h_1\ h_2]^T v = 0$ iff. $v = \tau[1\ 1\ -\frac{1}{2z}]^T$. Note that the linear independence constraint qualification is met at both locations. So, by Theorem 4.47, we have to consider all v for which $Dg(a_i)v \leq 0$, for $i = 1, 2$, respectively. As $Dg(x) = [1\ 0\ 0]$, we see that in both cases v should be restricted to those directions where $\tau \leq 0$.
Evaluation of $d(v) = v^T HLv$ at a_1, with $\mu_2 = 1$, and at a_2, with $\mu_2 = -2$, shows that in both cases $d(v) < 0$ if $v \neq 0$. So, by Theorem 4.47, $\tilde{f}$ has a local maximum at both these points. Clearly, since $\tilde{f}(a_1) = 5 > 1 = \tilde{f}(a_2)$, $\tilde{f}$ has not a global maximum at a_2.
To decide whether a global maximum occurs at a_1, reconsider the maximization problem under the additional constraint that $g_2 := x^2 + y^2 + z^2 \leq r^2$, for a large value of r. Clearly, if this constraint is not binding, we reobtain our points a_1 and a_2 as candidate locations for the maximization problem. In case this additional constraint g_2 is binding, we obtain the intersection points with h_1 and h_2 as additional potential global maximum locations. Substitution of both constraints into g_2, shows that y satisfies $2y^2 - 3y + 1 - r^2$. So, for large values of r this yields $y \approx \frac{3 \pm |r|}{4}$. Since $z^2 = 2 - y \geq 0$ (or, $y \leq 2$), this gives rise to the additional locations $(y - 1, y, \pm\sqrt{2 - y})$, where $y \approx \frac{3-|r|}{4}$. Evaluation of $\tilde{f}$ at these locations yields $\tilde{f} \approx -xy = -(y-1)y < 5$. So, this constrained maximization problem has for arbitrarily large r its global maximum at a_1. So, the original constraint optimization problem has a global maximum at a_1.

f. With $\tilde{f} := -f$ consider the maximization of $\tilde{f}$ subject to the constraints. We follow Algorithm 4.20 to find the solution of this problem.
Step 1. Let $g_1(x, y, z) := -x$, $g_2(x, y, z) := -y$, $g_3(x, y, z) := -z$ and $g_4(x, y, z) := y - x^2 - z^2$. Then $\tilde{f}$ and $g_i,\ i = 1, \cdots, 4, \in C^1$.

Step 2. $Dg_1 = [-1\ 0\ 0]$, $Dg_2 = [0\ -1\ 0]$, $Dg_3 = [0\ 0\ -1]$ and $Dg_4 = [-2x\ 1\ -2z]$.

So, see Remark 4.28,

$$\tilde{G} = \begin{bmatrix} -1 & 0 & 0 & -x & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -y & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -z & 0 \\ -2x & 1 & -2z & 0 & 0 & 0 & y - x^2 - z^2 \end{bmatrix}.$$

Using Gaussian elimination one easily verifies that the rank of $\tilde{G}$ coincides with the rank of matrix

$$\begin{bmatrix} -1 & 0 & 0 & -x & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -y & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -z & 0 \\ 0 & 0 & 0 & 2x^2 & y & 2z^2 & y - x^2 - z^2 \end{bmatrix}.$$

So matrix $\tilde{G}$ has not full row rank iff $x = y = z = 0$. Therefore, $(0, 0, 0)$ is a candidate solution.

Step 3. Let $L(x, y, z, \lambda_1, \lambda_2, \lambda_3, \lambda_4) := \tilde{f} + \lambda_1 x + \lambda_2 y + \lambda_3 z - \lambda_4(y - x^2 - z^2)$.

Step 4. Then next conditions must be satisfied:

$$\text{i) } 2(\lambda_4 - 1)x + y + \lambda_1 = 0;\ \text{ii) } x + z + \lambda_2 - \lambda_4 = 0;\ \text{iii) } y + 2(\lambda_4 - 1)z + \lambda_3 = 0;$$
$$\text{iv) } \lambda_i g_i = 0,\ i = 1, \cdots, 4;\ \text{v) } g_i(x, y, z) \leq 0,\ i = 1, \cdots, 4;\ \text{and viii) } \lambda_i \geq 0,\ i = 1, \cdots, 4.$$

1. Assume g_4 is not binding, i.e., $y < x^2 + z^2$. Consequently, $\lambda_4 = 0$. Then by ii), $x = z = \lambda_2 = 0$ and thus, by iii), $y = \lambda_3 = 0$. So, $(0, 0, 0)$ results again with $\lambda_i = 0,\ i = 1, \cdots, 4$.
2. If g_4 is binding, $y = x^2 + z^2$.
   1. Consider the case $x = 0$. Then, by i), $y = \lambda_1 = 0$. So, we reobtain $(0, 0, 0)$.
   2. If $x > 0$, by v), $\lambda_1 = 0$. Furthermore, as $y = x^2 + z^2 > 0$, also $\lambda_2 = 0$.

      Now, if $z = 0$, by iii), $y = \lambda_3 = 0$. But this violates $y > 0$. So, $z > 0$ and thus, by v), $\lambda_3 = 0$. From i) and iii) we next conclude that $2(\lambda_4 - 1)(z - x) = 0$. So, either $\lambda_4 = 1$ or $z = x$ applies. If $\lambda_4 = 1$, by iii), $y = 0$, which violates our previous assumption $y > 0$. So, $z = x$. Therefore, by ii), $\lambda_4 = 2x$. Substitution of this into iii) yields then as a candidate solution $(x, y, z) = \frac{1}{9}(3, 2, 3)$, with $\lambda = (0, 0, 0, \frac{2}{3})$.

So we find two candidate solutions: $(x, y, z) = (0, 0, 0)$ and $(x, y, z) = \frac{1}{9}(3, 2, 3)$.

Next, consider HL. Simple calculations show that

$$HL = \begin{bmatrix} 2\lambda_4 - 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 2\lambda_4 - 2 \end{bmatrix}.$$

Now, $Dg_4 = [-2x\ 1\ -2z]$. Elementary algebraic calculations show that the nulspace of this matrix is given by the range of matrix $B = \begin{bmatrix} 1 & 0 \\ 2x & 2z \\ 0 & 1 \end{bmatrix}$.
Consequently at $(x, y, z) = \frac{1}{9}(3, 2, 3)$, with $\lambda_4 = \frac{2}{3}$,

$$B^T HLB = \begin{bmatrix} 2\lambda_4 - 2 + 4x & 2x + 2z \\ 2x + 2z & 2\lambda_4 - 2 + 4z \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix}.$$

Clearly, this matrix is not negative semi-definite. So, $\tilde{f}$ has not a local maximum here.
Finally, consider $a = (0, 0, 0)$ with $\lambda_i = 0,\ i = 1, \cdots, 4$. As already noticed in Step 2, the linear independence constraint qualification is not met at this point. However, one easily verifies that all constraints are convex. So, we conclude from Theorem 4.44 that $T_a(D) = TL_a(D)$. Next, notice that

$$TL_a(D) = \{v \mid \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} v \leq 0\}.$$

From this one easily establishes that $v = [v_1\ 0\ v_3]^T$, where $v_i \geq 0,\ i = 1, 3$.
Since (a, λ) satisfy the (KKT) conditions, by Theorem 4.48, we have that a will be a local maximum if $v^T HL(a, \lambda)v < 0$, for all $v \in TL_a(D),\ v \neq 0$.
One easily verifies that $v^T HL(a, \lambda)v = -2v_1^2 - 2v_2^2$. So, indeed, $\tilde{f}$ has a local maximum at a.
Moreover, as $\tilde{f}(1, 2, 1) = 3 > 0 = \tilde{f}(0, 0, 0)$, a does not provide a global maximum.

g. Note that the constraint set is compact. As f is continuous we have from Weierstrass' theorem that a solution exists. We follow Algorithm 4.20 to find the solution of the problem.
Step 1. Let $g_1(x, y, z) := x^2 + y^2 - 1$ and $g_2(x, y, z) := y^2 + z^2 - 1$. Then f and $g_i \in C^1,\ i = 1, 2$.
Step 2. $Dg_1 = [2x\ 2y\ 0]$ and $Dg_2 = [0\ 2y\ 2z]$. Note that if g_1 is binding $Dg_1 \neq 0$. Similarly, if g_2 is binding, $Dg_2 \neq 0$. If both constraints are simultaneously binding, Dg has not full row rank iff. $(x, y, z) = (0, 1, 0)$. So, this is a candidate solution.
Step 3. Let $L(x, y, \lambda_1, \lambda_2) = x^3 + y^3 + z^3 - \lambda_1(x^2 + y^2 - 1) - \lambda_2(y^2 + z^2 - 1)$.
Step 4. Then next conditions must be satisfied:

$$\text{i) } 3x^2 - 2\lambda_1 x = 0;\ \text{ii) } 3y^2 - 2\lambda_1 y - 2\lambda_2 y = 0;\ \text{iii) } 3z^2 - 2\lambda_2 z = 0;$$
$$\text{iv) } \lambda_i g_i = 0,\ i = 1, 2;\ \text{v) } g_i \leq 0 \text{ and vi) } \lambda_i \geq 0,\ i = 1, 2.$$

1. Assume no constraint is binding. Consequently, $\lambda_i = 0,\ i = 1, 2$. Then $(0, 0, 0)$ results as a candidate solution.

2. Assume precisely one constraint is binding.
   1. Case $x^2+y^2=1$. Then $\lambda_2=0$. Then by iii), $z=0$. From i) and ii) (together with $x^2+y^2=1$) we obtain then next candidate solutions $(x,y,z,\lambda_1,\lambda_2)=(0,1,0,\frac{3}{2},0)$; $(x,y,z,\lambda_1,\lambda_2)=(1,0,0,\frac{3}{2},0)$ and $(x,y,z,\lambda_1,\lambda_2)=(\frac{1}{2}\sqrt{2},\frac{1}{2}\sqrt{2},0,\frac{3}{4}\sqrt{2},0)$.
   2. Case $y^2+z^2=1$. Then $\lambda_1=0$. Then by i), $x=0$. From ii) and iii) (together with $y^2+z^2=1$) we obtain then next candidate solutions $(x,y,z,\lambda_1,\lambda_2)=(0,1,0,0,\frac{3}{2})$; $(x,y,z,\lambda_1,\lambda_2)=(0,0,1,0,\frac{3}{2})$ and $(x,y,z,\lambda_1,\lambda_2)=(0,\frac{1}{2}\sqrt{2},\frac{1}{2}\sqrt{2},0,\frac{3}{4}\sqrt{2})$.
3. Assume both constraints are binding.

   If either $x=0$ or $z=0$, we again find the candidate solution $(x,y,z)=(0,1,0)$. If $y=0$, we find the candidate solution $(x,y,z,\lambda_1,\lambda_2)=(1,0,1,\frac{3}{2},\frac{3}{2})$.

   Otherwise, from $y^2=1-x^2$, substitution into $g_2=0$ gives $x=\pm z$. Whereas conditions i)-iii) yield: $x=z=\frac{2}{3}\lambda_1$ and $y=x+z$. Which yields the candidate solution: $(x,y,z,\lambda_1,\lambda_2)=(\frac{1}{5}\sqrt{5},\frac{2}{5}\sqrt{5},\frac{1}{5}\sqrt{5},\frac{3}{10}\sqrt{5},\frac{3}{10}\sqrt{5})$.

So, we find the eight candidate solutions (x,y,z): $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, $(1,0,1)$, $(\frac{1}{2}\sqrt{2},\frac{1}{2}\sqrt{2},0)$, $(0,\frac{1}{2}\sqrt{2},\frac{1}{2}\sqrt{2})$ and $(\frac{1}{5}\sqrt{5},\frac{2}{5}\sqrt{5},\frac{1}{5}\sqrt{5})$.

Evaluation of the candidate solutions shows that the maximum is attained at $(x,y,z)=(1,0,1)$.

h. We use Algorithm 4.20 again to find the solution of this problem.

Step 1. Let $g_1(x,y,z):=1-x-y$ and $g_2(x,y,z):=y^2+z^2-1$. Then f and $g_i\in C^1,\ i=1,2$.

Step 2. $Dg_1=[-1\ -1\ 0]$ and $Dg_2=[0\ 2y\ 2z]$. Note that if g_1 is binding $Dg_1\neq 0$. Similarly, if g_2 is binding, $Dg_2\neq 0$. If both constraints are simultaneously binding, one easily verifies Dg has always full row rank. So, there are no candidate solutions here.

Step 3. Let $L(x,y,\lambda_1,\lambda_2)=-x^2-y^+z^2+\lambda_1(x+y-1)-\lambda_2(y^2+z^2-1)$.

Step 4. Then next conditions must be satisfied:

$$\text{i)}\ -2x+\lambda_1=0;\ \text{ii)}\ -2y+\lambda_1 y-2\lambda_2 y=0;\ \text{iii)}\ 2z-2\lambda_2 z=0;$$
$$\text{iv)}\ \lambda_i g_i=0,\ i=1,2;\ \text{v)}\ g_i\leq 0 \text{ and } \text{vi)}\ \lambda_i\geq 0,\ i=1,2.$$

1. Assume no constraint is binding. Consequently, $\lambda_i=0,\ i=1,2$. Then $(0,0,0)$ satisfies i)-iii), but violates the first constraint.
2. Assume precisely one constraint is binding.
   1. Case $x+y=1$. Then $\lambda_2=0$. Then one easily verifies that the candidate solution $(x,y,z,\lambda_1,\lambda_2)=(\frac{1}{2},\frac{1}{2},0,1,0)$ results.
   2. Case $y^2+z^2=1$. Then $\lambda_1=0$. From i), ii) and vi) we then obtain $x=y=0$, which violates the first constraint.
3. Assume both constraints are binding. $x+y=1$ and $y^2+z^2=1$.

   Inspection of the conditions shows that $(x,y,z,\lambda_1,\lambda_2)=(\frac{2}{3},\frac{1}{3},\frac{\sqrt{8}}{3},\frac{4}{3},1)$ and $(x,y,z,\lambda_1,\lambda_2)=(\frac{2}{3},\frac{1}{3},-\frac{\sqrt{8}}{3},\frac{4}{3},1)$ result as candidate solutions.

So we have three candidate solutions: $(x,y,z,\lambda_1,\lambda_2)=(\frac{1}{2},\frac{1}{2},0,1,0)$, $(x,y,z,\lambda_1,\lambda_2)=(\frac{2}{3},\frac{1}{3},\frac{\sqrt{8}}{3},\frac{4}{3},1)$ and $(x,y,z,\lambda_1,\lambda_2)=(\frac{2}{3},\frac{1}{3},-\frac{\sqrt{8}}{3},\frac{4}{3},1)$.

Note that the smallest value of the objective function of these 3 points is attained at $(x, y, z, \lambda_1, \lambda_2) = (\frac{2}{3}, \frac{1}{3}, \frac{\sqrt{8}}{3}, \frac{4}{3}, 1)$ and $(x, y, z, \lambda_1, \lambda_2) = (\frac{2}{3}, \frac{1}{3}, -\frac{\sqrt{8}}{3}, \frac{4}{3}, 1)$. So, both these points are candidates for a solution of the problem. Unfortunately the constraint set is not bounded. Therefore, we consider the Lagrangian at $(\lambda_1, \lambda_2) = (\frac{4}{3}, 1)$. This function is given by $\tilde{L}(x, y, z) = -x^2 - 2y^2 + \frac{4}{3}x + \frac{4}{3}y - \frac{1}{3}$. Therefore

$$H\tilde{L} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Clearly, $H\tilde{L}$ is negative semi-definite on $\mathbb{R}^3$. So, $\tilde{L}$ is concave on $\mathbb{R}^3$. Therefore we conclude that the constraint optimization problem attains a global minimum at both these locations.

EXERCISE 4.11 Find all local maxima of next constrained optimization problems.

a. Optimize $f(x, y, z) = -6x + y + z$ s.t. $y \leq x^3$ and $z \leq x^3$.

b. Optimize $f(x, y, z) = xyz$ s.t. $y \leq x^3$ and $z \leq x^3$.

ANSWER 4.11

a. We follow Algorithm 4.20 to find the solution of the problem.

Step 1. Let $g_1(x, y, z) := y - x^3$ and $g_2(x, y, z) := z - x^3$. Then f and $g_i \in C^1,\ i = 1, 2$.

Step 2. $Dg_1 = [-3x^2\ 1\ 0]$ and $Dg_2 = [-3x^2\ 0\ 1]$. Note that if g_1 is binding $Dg_1 \neq 0$. Similarly, if g_2 is binding, $Dg_2 \neq 0$. If both constraints are simultaneously binding, one easily verifies Dg has always full row rank. So, there are no candidate solutions here.

Step 3. Let $L(x, y, \lambda_1, \lambda_2) = -6x + y + z - \lambda_1(y - x^3) - \lambda_2(z - x^3)$.

Step 4. Then next conditions must be satisfied:

$$\text{i)}\ -6 + 3\lambda_1 x^2 + 3\lambda_2 x^2 = 0;\ \text{ii)}\ 1 - \lambda_1 = 0;\ \text{iii)}\ 1 - \lambda_2 = 0;$$
$$\text{iv)}\ \lambda_i g_i = 0,\ i = 1, 2;\ \text{v)}\ g_i \leq 0 \text{ and vi)}\ \lambda_i \geq 0,\ i = 1, 2.$$

From ii) and iii) we have $\lambda_i = 1,\ i = 1, 2$. Substitution of this into i) and next evaluation of the other conditions yields next candidate solutions: $(x, y, z, \lambda_1, \lambda_2) = (1, 1, 1, 1, 1)$ and $(x, y, z, \lambda_1, \lambda_2) = (-1, -1, -1, 1, 1)$.

Clearly, neither of these candidates yields a global maxima of the problem, since the objective function attains a larger value at, for example $(x, y, z) = (10, 1000, 1000)$. To determine whether a local maximum occurs at the candidate solutions, we consider Theorem 4.19.

To that end consider $L(x, y, z, 1, 1) = 2x^3 - 6x$. Note that both constraints are binding. Therefore matrix

$$B = \begin{bmatrix} & & 1 & \\ -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} & \begin{bmatrix} -3x^2 \\ -3x^2 \end{bmatrix} & \end{bmatrix} = \begin{bmatrix} 1 \\ 3x^2 \\ 3x^2 \end{bmatrix}.$$

With matrix B as above one easily verifies that $B^T HL(x,y,z,1,1,1)B = 12x$. Therefore $B^T HL(1,1,1,1,1)B > 0$ and $B^T HL(-1,-1,-1,1,1)B < 0$. So the problem has a local maximum at $(-1-,-1,-1)$ and has not a local maximum at $(1,1,1)$.

b. First, consider the first part of Algorithm 4.20 to find candidate solutions of the problem.

Step 1. Let $g_1(x,y,z) := y - x^3$ and $g_2(x,y,z) := z - x^3$. Then f and $g_i \in C^1,\ i = 1,2$.

Step 2. $Dg_1 = [-3x^2\ 1\ 0]$ and $Dg_2 = [-3x^2\ 0\ 1]$. Note that if g_1 is binding, $Dg_1 \neq 0$. Similarly, if g_2 is binding, $Dg_2 \neq 0$. If both constraints are simultaneously binding, one easily verifies Dg has always full row rank. So, there are no candidate solutions here.

Step 3. Let $L(x,y,\lambda_1,\lambda_2) = xyz - \lambda_1(y - x^3) - \lambda_2(z - x^3)$.

Step 4. Then next conditions must be satisfied:

$$\text{i) } yz + 3\lambda_1 x^2 + 3\lambda_2 x^2 = 0;\ \text{ii) } xz - \lambda_1 = 0;\ \text{iii) } xy - \lambda_2 = 0;$$
$$\text{iv) } \lambda_i g_i = 0,\ i = 1,2;\ \text{v) } g_i \leq 0 \text{ and vi) } \lambda_i \geq 0,\ i = 1,2.$$

To solve this set of equations consider next 3 possibilities:

1. No constraint is binding. Then $\lambda_i = 0$. This yields $(x,y,z,\lambda_1,\lambda_2) = (\bar{x},0,0,0,0)$, where $\bar{x} > 0$, as candidate solutions.
2. precisely one constraint is binding.
   1. g_1 is binding, i.e., $y = x^3$. Then $\lambda_2 = 0$. This yields $(x,y,z,\lambda_1,\lambda_2) = (\bar{x},0,0,0,0)$, where $\bar{x} > 0$, and $(x,y,z,\lambda_1,\lambda_2) = (0,0,\bar{z},0,0,0,0)$, where $\bar{z} < 0$, as candidate solutions.
   2. g_2 is binding, i.e., $z = x^3$. Then $\lambda_1 = 0$. This yields $(x,y,z,\lambda_1,\lambda_2) = (\bar{x},0,0,0,0)$, where $\bar{x} > 0$, and $(x,y,z,\lambda_1,\lambda_2) = (0,\bar{y},0,0,0,0,0)$, where $\bar{y} < 0$, as candidate solutions.
   3. Both constraints are binding. This yields $(x,y,z,\lambda_1,\lambda_2) = (0,0,0,0,0)$.

To determine, next, whether a local maximum occurs at the candidate solutions we consider Theorem 4.19.

To that end consider $L(x,y,z,0,0) = xyz$. Elementary calculations show

$$HL = \begin{bmatrix} 0 & z & y \\ z & 0 & x \\ y & x & 0 \end{bmatrix}.$$

If both constraints are binding, at $(0,0,0)$, $B = \begin{bmatrix} 1 \\ -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -3x^2 \\ -3x^2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

Then $B^T HL(0,0,0,0,0)B = 0$. So the theorem does not provide a conclusive answer. However, if we evaluate the objective function along the curve $(\epsilon,\epsilon^3,\epsilon^3)$ we see that for $\epsilon > 0$, the objective values are larger than zero. So, the problem has not a local maximum at $(0,0,0)$.

Next consider the candidates obtained in case g_1 is binding and g_2 not.

First consider the candidate solutions: $(x,y,z) = (\bar{x},0,0)$, where $\bar{x} > 0$. As $Dg_1 = [-3x^2\ 1\ 0]$, we

have that $Dg_1(\bar{x},0,0) = [-3\bar{x}^2\ 1\ 0]$ with $\bar{x} > 0$. Note that unfortunately $Dg_{1z} = 0$. However, for example, $Dg_{1x} \neq 0$. Therefore, we apply the Implicit Function Theorem with respect to the first variable rather than the third, which yields

$$B = \begin{bmatrix} \frac{-1}{3\bar{x}^2} & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Evaluated at $(\bar{x},0,0)$, $B^T HLB$ equals $\begin{bmatrix} 0 & \bar{x} \\ \bar{x} & 0 \end{bmatrix}$.
Since this matrix is indefinite (and thus not negative semi-definite), it follows that none of these points provides a local maximum.
Next consider the candidate solutions: $(x,y,z) = (0,0,\bar{z})$, where $\bar{z} < 0$. As $Dg_1 = [-3x^2\ 1\ 0]$, we have that $Dg_1(0,0,\bar{z},0,0) = [0\ 1\ 0]$. Again, unfortunately, $Dg_{1z} = 0$. However, $Dg_{1y} \neq 0$. Therefore, we apply the Implicit Function Theorem with respect to the second variable rather than the third, which yields

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Evaluated at $(0,0,\bar{z})$, $B^T HLB$ equals $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.
Since this matrix is negative semi-definite, the theorem does not allow us to conclude whether these candidate solutions are local maxima. A more detailed inspection is needed. To that end consider the parameterized curve $(x,y,z) = (\epsilon,-\epsilon,\bar{z})$, with $\epsilon > 0$, in the constraint set. The objective function takes strictly positive values along this curve. Hence, the optimization problem has not a local maximum at $(0,0,\bar{z})$.
Next consider the case g_2 is binding and g_1 not.
We have already seen that the candidate solutions $(x,y,z) = (\bar{x},0,0)$, where $\bar{x} > 0$, do not provide local maxima. We therefore consider the candidate solutions $(x,y,z) = (0,\bar{y},0)$, where $\bar{y} < 0$.
Since $Dg_2 = [-3x^2\ 0\ 1]$, it follows that $Dg_2(0,\bar{y},0) = [0\ 0\ 1]$. Hence

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Evaluated at $(0,\bar{y},0)$, $B^T HLB$ equals $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.
Since this matrix is negative semi-definite, the theorem does not allow us to conclude whether these candidate solutions yield local maxima. A more detailed analysis is therefore required. To

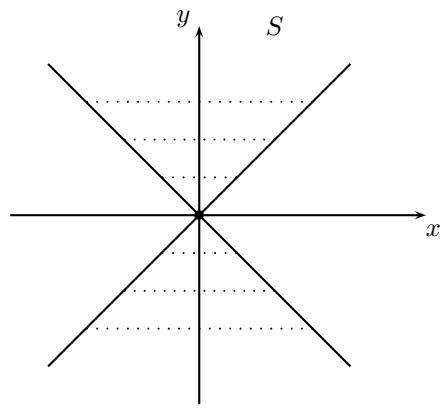


FIGURE 4.1 – $S = \{(x, y) \mid x^2 - y^2 \leq 0.\}$

this end, consider the parameterized curve $(x, y, z) = (\epsilon, \bar{y}, -\epsilon)$, with $\epsilon > 0$, in the constraint set. The objective function takes strictly positive values along this curve. Therefore $(0, \bar{y}, 0)$ do not provide a local maximum.

Finally, if no constraint is binding we found the candidate solutions: $(x, y, z) = (\bar{x}, 0, 0)$, where $\bar{x} > 0$, from which we already showed they are not local maxima.

So the problem has no local maxima.

EXERCISE 4.12

Determine the tangent and linearized cone for next sets at the indicated point a, respectively.

a. $S = \{(x, y) \, ||x| \leq 1 \text{ and } |y| \leq 1\}$ at $(1, 1)$.

b. $S = \{(x, y) \, |x^2 + y^2 \leq 1\}$ at $(1, 0)$.

c. $S = \{(x, y) \, |x^2 - y^2 \leq 0\}$ at $(0, 0)$.

d. $S = \{(x, y, z) \, | \log(1 + x^2 + y^2 + z^2) \leq \log(4);\ \frac{1}{2+x} + \frac{1}{2+x} + \frac{1}{2+x} + \frac{1}{2+x} \geq 1;\ x + y + z \leq 3;\ x \geq 1;\ |z| \leq 2\}$ at $(0, 0)$.

ANSWER 4.12

a. Note that $Dg(1, 1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, which is invertible and therefore has full row rank. Hence, by Theorem 4.43, it follows that $T_S(1, 1) = TL_S(1, 1)$.

Next, $TL_S(1, 1) = \{u \mid D(x)u \leq 0 \text{ and } D(y)u \leq 0\}$ evaluated at $(x, y) = (1, 1)$. Consequently, $TL_S(1, 1) = \{(u_1, u_2) \mid [1\ 0]u \leq 0 \text{ and } [0\ 1]u \leq 0\} = \{(u_1, u_2)^T \, |u_1 \leq 0 \text{ and } u_2 \leq 0\}$.

b. Observe that $Dg(1, 0) = 2 \begin{bmatrix} 1 & 0 \end{bmatrix}$, which has full row rank. Therefore, by Theorem 4.43, it follows that $T_S(1, 1) = TL_S(1, 1)$.

Next, $TL_S(1, 0) = \{u \mid D(x^2 + y^2)u \leq 0\}$ evaluated at $(x, y) = (1, 0)$. Hence, $TL_S(1, 0) = \{(u_1, u_2) \mid 2[1\ 0]u \leq 0\} = \{(u_1, u_2)^T \, |u_1 \leq 0\}$.

c. We illustrated S in Figure 4.1.

From this one easily verifies that $T_S(0,0) = S$.

Next, $TL_S(0,0) = \{u \mid D(x^2 - y^2)u \leq 0\}$ evaluated at $(x,y) = (0,0)$. Hence, $TL_S(0,0) = \{(u_1,u_2) \mid 2[0\ 0]u \leq 0\} = \mathbb{R}^2$.

Observe that $Dg(0,0) = 2[0\ 0]$, which does not have full row rank.

d. Note that all constraints except $|z| \leq 2$ are binding. Define

$$\begin{aligned} g_1(x,y,z) &:= \log(1 + x^2 + y^2 + z^2);\ g_2(x,y,z) := -\frac{1}{2+x} - \frac{1}{2+y} - \frac{1}{2+z}; \\ g_3(x,y,z) &:= x + y + z \text{ and } g_4(x,y,z) := -x. \end{aligned}$$

Then, $Dg(x,y,z) = \begin{bmatrix} \frac{2x}{1+x^2+y^2+z^2} & \frac{2y}{1+x^2+y^2+z^2} & \frac{2z}{1+x^2+y^2+z^2} \\ \frac{1}{(2+x)^2} & \frac{1}{(2+y)^2} & \frac{1}{(2+z)^2} \\ 1 & 1 & 1 \\ -1 & 0 & 0 \end{bmatrix}$. So, $Dg(1,1,1) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ 1 & 1 & 1 \\ -1 & 0 & 0 \end{bmatrix}$.

Observe that for $u := [1\ \ -2\ 0]^T$ we have $Dg(1,1,1)u < 0$. Hence, by Theorem 4.43 (item 3), it follows that $T_S(1,1,1) = TL_S(1,1,1)$.

Next, $TL_S(1,1,1) = \{u \mid Dg(1,1,1)u \leq 0\} = \{[u_1\ u_2\ u_3]^T \mid [1\ 1\ 1]u \leq 0 \text{ and } u_1 \geq 0\}$.

Advanced, Deepening and Broadening Exercises

EXERCISE 4.13 *(Reformulating optimization problems into the standard framework)*

In this exercise a number of optimization problems are considered that can be reformulated into the standard framework.

Reformulate next problems as a (mixed) constrained maximization problem (4.11).

a. Maximize $f(x)$ subject to $\max_x\{g_1(x), g_2(x)\} \leq a$.

b. Maximize $f(x)$ subject to $|g_1(x)| + |g_2(x)| \leq a$.

c. Minimize $f(x)$ where $f(x) = \max_x\{g_1(x), g_2(x)\}$.

d. Minimize the Euclidean distance between two sets given by $S_i = \{x \mid g_i(x) \leq b_i\},\ i = 1,2$.

ANSWER 4.13

a. Maximize $f(x)$ subject to $g_1(x) \leq a$ and $g_2(x) \leq a$.

b. Maximize $f(x)$ subject to $g_1(x)+g_2(x) \leq a$, $-g_1(x)+g_2(x) \leq a$, $g_1(x)-g_2(x) \leq a$, $-g_1(x)-g_2(x) \leq a$.

c. Find t that maximizes $-t$ subject to $g_1(x) \leq t$ and $g_2(x) \leq t$.

d. Minimize $f(x) = \sqrt{\sum_{i-1}^{n}(x_i - y_i)^2}$ subject to $g_1(x) \leq b_1$ and $g_2(y) \leq b_2$. Or, determine $-t$ that maximizes $t := -\sqrt{\sum_{i-1}^{n}(x_i - y_i)^2}$ subject to $g_1(x) \leq b_1$ and $g_2(y) \leq b_2$.

EXERCISE 4.14 *(Solving an inequality constrained optimization problem as an equality constrained optimization problem)*

In this exercise you are asked to show that an optimization problem subject to inequality constraints can be solved also as an equality constrained optimization problem by the introduction of so-called *slack variables*. For simplicity you are asked to prove this result if there is only one inequality restriction.

Assume $f(x),\ g(x) : \mathbb{R}^n \to \mathbb{R}$ are continuous on an open set $S \subset \mathbb{R}^n$. Consider the set $C := \{x \in \mathbb{R}^n \mid g(x) \leq b,\ x \in S\}$ and $C_1 := \{(x, s) \in \mathbb{R}^{n+1} \mid g(x) + s^2 = b,\ x \in S\}$. Assume $a \in C$. Let $f_1(x, s) := f(x)$ for all s.

a. Show that $s(x) := \sqrt{b - g(x)}$ is a continuous function on C.

b. Assume $f(x)$ has a local maximum at $x = a$ on C. Show that $f_1(x, s)$ has a local maximum at $(x, s) = (a, \sqrt{b - g(a)})$ on C_1.

c. Assume $f_1(x, s)$ has a local maximum at $(x, s) = (a, \sqrt{b - g(a)})$ on C_1. Show that $f(x)$ has a local maximum at $x = a$ on C.

d. Is $(x, s) = (a, \sqrt{b - g(a)})$ on C_1 the only location where $f_1(x, s) = f(a)$? Comment on this.

ANSWER 4.14

a. Note that $b - g(x) \geq 0$ on $S \cap C$. As $g(x)$ is continuous, $b - g(x)$ is also continuous, and therefore its square root, $s(x)$, is continuous as well.

b. Sine $f(x)$ has a local maximum at $x = a$ on C, by definition $f(x) \leq f(a)$ for all x in the intersection of C with some ball B centered at a with radius ϵ.

   Now, consider an arbitrary element (x, s) in the intersection of C_1 with a ball B_1 centered at $(a, \sqrt{b - g(a)})$ with radius ϵ. Then (x, s) satisfies

$$\sqrt{(x - a)^2 + (s - \sqrt{b - g(a)})^2} \leq \epsilon.$$

   Which implies that $x \in B$.

   Furthermore, since $(x, s) \in C_1$, it follows that $g(x) \leq b$. Hence, $x \in C$. Therefore, x lies in the intersection $C \cap B$, where by assumption $f(x) \leq f(a)$.

c. Since $f_1(x, s)$ has a local maximum at $(a, \sqrt{b - g(a)})$ on C_1, by definition $f(x) \leq f(a)$ for all (x, s) in the intersection of C_1 and some ball B_1 centered at $(a, \sqrt{b - g(a)})$ with radius δ_1.

   Recall from item a that $s(x) = \sqrt{b - g(x)}$ is continuous. In particular, this implies that for every $\epsilon_2 > 0$ there exists a ball B_2, centered at a with radius δ_2, such that for all $x \in B_2$,

$$|s(a) - s(x)| < \epsilon_2.$$

   Now consider an arbitrary x in the intersection of C with a ball B, centered at a with radius δ_2, where ϵ_2 and δ_2 satisfy $\epsilon_2 + \delta_2 \leq \delta_1$. Then, x satisfies $\|x - a\| \leq \delta_2$ and $g(x) \leq b$. From the latter

inequality it follows that there exists $s(x) := \sqrt{b - g(x)}$ such that

$$g(x) + s^2(x) = b.$$

As noted above, $s(x)$ is continuous and, since $\|x - a\| \leq \delta_2$, it follows that $|s(x) - s(a)| < \epsilon_2$. Consequently,

$$\sqrt{(x-a)^2 + (s(x) - s(a))^2} \leq \delta_2 + \epsilon_2 \leq \delta_1.$$

Hence $(x, s(x)) \in B_1 \cap C_1$, which implies that $f(x) \leq f(a)$.

d. No, also $f_1(a, -\sqrt{b - g(a)}) = f(a)$. In fact one can show, similarly as above, that also $f_1(x, s)$ has a local maximum at $(a, -\sqrt{b - g(a)})$. So, f_1 has at least two locations where it attains the same maximum value.

EXERCISE 4.15 *(Proof Lemma 4.7)*

Assume $g(x) : \mathbb{R}^n \to \mathbb{R}$ is continuous differentiable on an open set $S \subset \mathbb{R}^n$ containing a. Consider the set $C = \{x \in \mathbb{R}^n \mid g(x) \leq b\}$ and assume $g(a) = b$. Prove that

1. $Dg_u(a)u \geq 0$ for any direction u pointing into C. Hint: Consider $\tilde{g}(t) = g(a + tu)$ for appropriate $t \geq 0$.
2. $Dg_u(a)u \leq 0$ for any direction u pointing outside C.
3. $Dg(a)$ is perpendicular to the tangent plane of the graph of $g(x)$ at $x = a$.

ANSWER 4.15

a. Let u be any direction such that, for t_0 sufficiently small and all $0 \leq t \leq t_0$, we have $x(t) = a + tu \in C$. Then define $\tilde{g}(t) := (g(x(t))$, which satisfies $\tilde{g}(t) \leq b$ for all $0 \leq t \leq t_0$ and $\tilde{g}(0) = b$.
Note that $\tilde{g}(t)$ is continuously differentiable on $(0, t_0)$. Moreover, $\tilde{g}(t)$ attains a maximum at $t = 0$. By the chain rule,

$$\tilde{g}^{'}(t) = Dg(x(t))x^{'}(t) = Dg(x(t))u.$$

Since $\tilde{g}(t)$ has a maximum at $t = 0$, there exists $0 < t_1 \leq t_0$ such that $\tilde{g}^{'}(t) \geq 0$ for all $t \in (0, t_1)$. Because $\tilde{g}(t)$ is continuous, it follows that $\lim_{t \downarrow 0} \tilde{g}^{'}(t) \geq 0$. Hence, $Dg(a)u \geq 0$.

b. Similar to item a. Consider $\tilde{g}(t) := g(x(t)$ where $x(t) = a + tu \in C = \{x \mid g(x) \geq b\}$. Note that $\tilde{g}(t)$ has a minimum at $t = 0$.

c. The tangent plane to the graph of $g(x)$ at $x = a$ is given by

$$\{x \mid Dg(a)(x - a) = 0\}.$$

Equivalently, the condition $Dg(a)(x - a) = 0$ can be interpreted as stating that every vector $x - a$, with x lying on the tangent plane, has innerproduct zero with the vector $Dg^T(a)$. In other words, vector $Dg^T(a)$ is orthogonal to every direction vector in the tangent plane of g at $x = a$. Hence, $Dg^T(a)$ is perpendicular to the tangent plane of g at $x = a$.

EXERCISE 4.16 *(Proof Theorem 4.19)*

Consider the assumptions and notation in Theorem 4.19.

a. Show that $f(a)$ will be a local maximum of the constrained optimization problem to maximize $f(x)$ subject to the constraint set defined by $S_c = \{x \in \mathbb{R}^n \mid g_1(x) = \tilde{b}_1\}$, where $\tilde{b}_1 = [b_1 \cdots b_{m_1}]^T$, if the advertized constrained Hessian matrix is negative definite.

b. Use item a. and Corollary 3.39 to show that a necessary condition for existence of a local maximum at a for the constraint optimization problem (4.8) is that $d(v) = v^T H\tilde{L}(a)v \leq 0$ for all $v \in \{\ v \mid Dg_1(a)v = 0\}$.

c. Show that the constraint optimization problem (4.8) has a maximum at $x = a$ under the stated assumptions. Hint: Introduce slack variables into the problem (see Exercise 4.14) to rewrite the problem as an equality constraint problem. Next, use Theorem 3.38 to conclude.

ANSWER 4.16

a. By assumption, $(a, \lambda_1, \cdots, \lambda_{m_1}, 0, \cdots, 0)$ satisfies the (KKT) conditions of the original constraint optimization problem (4.8), where g_1 contains the binding constraints and g_2 the non-binding constraints.

   Now, (with a slight abuse of notation) consider the Lagrange function

$$L_c := f(x) - \sum_{i=1}^{m_1} \mu_i(g_i(x) - b_i)$$

   associated with the maximization of $f(x)$ over the constrained set S_c. One can easily verify that $(x, \mu) := (a, \lambda_1, \cdots, \lambda_{m_1})$ satisfies the (KKT) conditions for this problem and that a is a regular point.

   Note that the Hessian associated with this problem, $HL_c(x, \lambda_1, \cdots, \lambda_{m_1})$, coincides with $H\tilde{L}(x)$. Since

$$d(v) = v^T H\tilde{L}(a)v < 0,$$

   for all $v \neq 0$ satisfying $Dg_1(a)v = 0$, it follows that

$$v^T HL_c(a, \lambda_1, \cdots, \lambda_{m_1})v < 0$$

   for this same set of vectors. By Theorem 3.38, this implies that $f(x)$ has a strict local maximum at a on S_c.

b. Consider the constrained optimization problem of maximizing $f(x)$ over the set S_c, where $g_1(x) = \tilde{b}_1$. Recall from the proof of item a. that $(a, \lambda_1, \cdots, \lambda_{m_1})$ satisfy the (KKT) conditions associated with this problem, and that a is a regular point for this constrained optimization problem.

   Furthermore, we observed that $H\tilde{L}(x)$ coincides with the Hessian $HL_c(x, \lambda_1, \cdots, \lambda_{m_1})$. By Corollary 3.39, a necessary condition for f to have a maximum on the constrained set S_c is that

$$d(v) = v^T HL_c v \leq 0 \text{ for all } v \in \{\ v \mid Dg_1(a)v = 0\}.$$

Consequently, it also holds that $d(v) = v^T H\tilde{L}(a)v \leq 0$ for all $v \in \{\ v \mid Dg_1(a)v = 0\}$.

c. Consider the constrained optimization problem of maximizing $f(x)$ on the set S_c, where the first m_1 constraints are binding, $g_1(x) = \tilde{b}_1$, and the remaining constraints are non-binding. By assumption the original problem (4.8) has a solution at $x = a$ if and only if the problem of maximizing $f(x)$ over the set

$$S_{ei} := \{x \mid g_1(x) = \tilde{b}_1,\ g_2(x) < \tilde{b}_2\}$$

has a maximum at $x = a$.

From Exercise 4.14, $f(x)$ will have a local maximum on this set S_{ei} if the function $f_1(x, s) := f(x)$ has a local maximum for the corresponding equality constrained problem where, for every non-binding constraint i, a slack variable s_i is introduced. That is, we consider maximizing $f_1(x, s)$ over the set

$$\{(x, s) \mid g_i(x) = b_i,\ i = 1, \cdots, m_1, g_i(x) + s_i^2 = b_i,\ i = m_1 + 1, \cdots, m\}.$$

The associated Lagrange function for this equality-constrained problem is

$$L(x, s, \lambda) := f(x) - \sum_{i=1}^{m_1} \lambda_i(g_i(x) - b_i) - \sum_{i=m_1+1}^{m} \lambda_i(g_i(x) + s_i^2 - b_i).$$

Let

$$\tilde{g}_i(x, s) := g_i(x),\ i = 1, \cdots, m_1 \text{ and } \tilde{g}_i(x, s) := g_i(x) + s_i^2,\ i = m_1 + 1, \cdots, m.$$

Then

$$D\tilde{g}(a, \bar{s}) = \begin{bmatrix} Dg_1(a) & 0 \\ Dg_2(a) & \text{diag}(2s_i(a)) \end{bmatrix}.$$

Since $b_i - g_i(a) > 0,\ i = m_1 + 1, \cdots, m$, we have $s_i(a) \neq 0$ for these i. By assumption, $Dg_1(a)$ has full row rank, which immediately implies that $D\tilde{g}(a, \bar{s})$ also has full row rank, where $\bar{s} := (s_{m_1+1}(a), \cdots, s_m(a))$. Therefore, $(a, \bar{s})$ is a regular point for this equality-constraint optimization problem.

By assumption the (KKT) conditions for the original problem (4.8) are satisfied. That is, with

$$L_{ce}(x, \lambda) := f(x) - \sum_{i=1}^{m} \lambda_i(g_i(x) - b_i),$$

we have

$$Df(a) = \sum_{i=1}^{m} \bar{\lambda}_i Dg_i(a),\ \bar{\lambda}_i = 0,\ i = m_1 + 1, \cdots, m, \text{ and } \bar{\lambda}_i \geq 0,\ i = 1, \cdots, m_1.$$

Now, consider the Lagrange conditions for the equality-constrained problem:

$$\begin{aligned} 1.a\ L_x &= Df(x) - \sum_{i=1}^{m_1} \lambda_i Dg_i(x) - \sum_{i=m_1+1}^{m} \lambda_i Dg_i(x) = 0; \\ 1.b\ L_{s_i} &= -2\lambda_i s_i = 0,\ i = m_1+1, \cdots, m. \end{aligned}$$

It is straightforward to verify that $x := a$; $s_i := g_i(a) - b_i,\ i = m_1+1, \cdots, m,$; and $\lambda_i := \bar{\lambda}_i,\ i = 1, \cdots, m$, satisfy conditions 1.a and 1.b.
By Theorem 3.38 $(a, \bar{s})$ will be a local maximum if $B^T HL(x, s, \bar{\lambda})B$ is negative definite at $(a, \bar{s})$, where $HL(x, s, \bar{\lambda}) = \begin{bmatrix} Hf(x) - \sum_{i=1}^{m_1} \bar{\lambda}_i Hg_i(x) & 0 \\ 0 & 0 \end{bmatrix}$ and B represents the nulspace of $D\tilde{g}(a, \bar{s})$, specifically

$$B = \begin{bmatrix} B_1 \\ -[\text{diag}(2\bar{s}_i)]^{-1}(Dg_2(a)B_1) \end{bmatrix},\ \text{where } B_1 = \begin{bmatrix} I \\ -(D_y g_1(a))^{-1} D_x g_1(a) \end{bmatrix}.$$

A direct calculation shows that

$$B^T HL(x, s, \bar{\lambda}) = B_1^T (Hf(x) - \sum_{i=1}^{m_1} \bar{\lambda}_i Hg_i(x))B_1.$$

Hence, $B^T HL(x, s, \bar{\lambda})B$ is negative definite at $(a, \bar{s})$ if and only if

$$B_1^T (Hf(x) - \sum_{i=1}^{m_1} \bar{\lambda}_i Hg_i(x))B_1$$

is negative definite at $x = a$.
In conclusion, the constraint maximization problem (4.8) has a local maximum at a regular point $x = a$ if, the first-order (KKT) conditions are satisfied and

$$B_1^T (Hf(x) - \sum_{i=1}^{m_1} \bar{\lambda}_i Hg_i(x))B_1$$

is negative definite at $x = a$.

EXERCISE 4.17 *(Reformulation Chebyshev distance problem as a Linear Programming Problem)*
The Chebyshev distance between 2 vectors x and y is defined as the maximum number of the absolute difference between entries of both vectors, i.e. $\max_i |x_i - y_i|$. In this exercise you are asked to show that the problem to find a vector in a convex polyhedron S that has a minimal Chebyshev distance to a given vector $c \in \mathbb{R}^n$ (not in S) can be reformulated as a linear programming problem.
a. Let $t \geq 0$. Show that $|x_i - y_i| \leq t$ iff. $x_i - y_i \leq t$ and $x_i - y_i \geq -t$.
b. Show that $\max_i |x_i - y_i| \leq t$ iff $x_i - t \leq y_i$ and $-x_i - t \leq -y_i,\ i = 1, \cdots, n$.

c. Assume the convex polyhedron S is described by the intersection of the half spaces $a_i^T x \leq b_i$. Show that the problem to find $\min_{s\in S} \max_i |s_i - c_i|$ (with $c \notin S$) is equivalent with the linear program to find $\max_{x,t} -t$ subject to the inequality constraints $a_i^T x - t \leq b_i + c_i$ and $-a_i^T x - t \leq -b_i - c_i,\ i = 1, \cdots, n$.

d. Consider 3 data points $(x_1, y_1) = (0, -1)$, $(x_2, y_2) = (1, 2)$ and $(x_3, y_3) = (2, 3)$. Determine parameters α and β that make the line $y = \alpha + \beta x$ as "close" to the points as possible in the sense of Chebyshev's distance. Verify your result.

ANSWER 4.17

a. Assume $|x_i - y_i| \leq t$. Then, if $x_i - y_i \geq 0$ clearly $x_i - y_i \leq t$. In case $x_i - y_i \leq 0$ it follows that $-x_i + y_i \leq t$. Hence, both inequalities apply.
Conversely, if both inequalities hold, it immediately follows that $|x_i - y_i| \leq t$.

b. From item a., it follows that $\max_i |x_i - b_i| \leq t$ iff. $x_i - b_i \leq t$ and $x_i - b_i \geq -t$ for all i. Rearranging these inequalities then gives the stated result.

c. Since $c \notin S$, the minimum distance from c to S is achieved on the boundary of S; that is, at a point $s \in S$ where $a_i^T x = b_i$, for $i = 1, \cdots, m$.
By item b., letting $x_i := a_i^T x$ and $y_i := b_i + c_i$, for a fixed $t > 0$ we have $\max_i |x_i - y_i| \leq t$ iff

$$a_i^T x - t \leq b_i + c_i \text{ and } -a_i^T x - t \leq -b_i - c_i,\ i = 1, \cdots, n.$$

Therefore, to find a point $s \in S$ that minimizes the distance to c, we must solve

$$\max_{x,t\geq 0} -t \text{ subject to } a_i^T x - t \leq b_i + c_i \text{ and } -a_i^T x - t \leq -b_i - c_i,\ i = 1, \cdots, n.$$

d. If the data points lay exactly on a line, the parameters α and β would satisfy the equations $\alpha = -1$, $\alpha + \beta = 2$ and $\alpha + 2\beta = 3$. Equivalently, in matrix form, this can be written as

$$A[\alpha\ \beta]^T = [-1\ 2\ 3]^T, \text{ where } A = [1\ 1\ 1; 0\ 1\ 2]^T.$$

Thus, our goal is to determine $x = [\alpha\ \beta]^T$ such that the Chebyshev distance between Ax and $b = [-1\ 2\ 3]^T$ is minimized.
By item c., this can be reformulated as the problem of maximizing $f(x, t) = -t$ subject to the constraints a) $x_1 - t \leq -1$; b) $-x_1 - t \leq 1$; c) $x_1 + x_2 - t \leq 2$; d) $-x_1 - x_2 - t \leq -2$; e) $x_1 + 2x_2 - t \leq 3$ and f) $-x_1 - 2x_2 - t \leq -3$.
Clearly, by Weierstrass' theorem, this problem has a solution. The with this problem corresponding Lagrange function is

$$\begin{aligned} L(x, t, \lambda) := &-t - \lambda_1(x_1 - t + 1) - \lambda_2(-x_1 - t - 1) - \lambda_3(x_1 + x_2 - t - 2) - \lambda_4(-x_1 - x_2 - t + 2) \\ &- \lambda_5(x_1 + x_2 - t - 3) - \lambda_6(-x_1 - x_2 - t + 3). \end{aligned}$$

From which we derive the necessary conditions:

$$\text{1.a)}\ \sum_{i=1}^{6}\lambda_i = 1;\ \ \text{b)}\ \lambda_2 - \lambda_1 + \lambda_4 - \lambda_3 + \lambda_6 - \lambda_5 = 0;\ \text{c)}\ \lambda_4 - \lambda_3 + \lambda_6 - \lambda_5 = 0;$$

$$\text{2.a)}\ \lambda_1(x_1 - t + 1) = 0;\ \text{b)}\lambda_2(-x_1 - t - 1) = 0;\ \text{c)}\ \lambda_3(x_1 + x_2 - t - 2) = 0;$$

$$\text{d)}\ \lambda_4(-x_1 - x_2 - t + 2) = 0;\ \text{e)}\ \lambda_5(x_1 + 2x_2 - t - 3) = 0\ \text{f)}\ \lambda_6(-x_1 - 2x_2 - t + 3) = 0.$$

$$\text{3.a)}\ \lambda_i \geq 0\text{and b) all constraints must be satisfied.}$$

First, note that constraints 1.b) and 1.c) can be rewritten equivalently as:

$$\text{1.b}')\ \lambda_1 = \lambda_2 \text{ and } \text{1.c}')\ \lambda_4 - \lambda_3 + \lambda_6 - \lambda_5 = 0.$$

1. Case 1: $x_1 - t + 1 = 0$ and $x_1 + t + 1 = 0$.
   Then it follows that $x_1 = -1$ and $t = 0$. From this we obtain from constraints c) and d) that $x_2 = 3$. But this violates constraint e).
2. Case 2: $x_1 - t + 1 = 0$ and $x_1 + t + 1 > 0$.
   Then we have $t = 1 + x_1 \neq 0$ and $\lambda_1 = \lambda_2 = 0$.
   1. if $x_1 + x_2 - t - 2 = 0$ it follows that $x_2 = 3$. Which again violates constraint e).
   2. if $x_1 + x_2 - t - 2 \neq 0$ we have $\lambda_3 = 0$. So, $\lambda_5 = \lambda_4 + \lambda_6$.
      1. if $\lambda_5 = 0$ it follows $\lambda_i = 0,\ i = 1, \cdots, 6$. Which violates condition 1.a)
      2. if $\lambda_5 \neq 0$, it follows that $x_2 = 2$. If $\lambda_6 \neq 0$ it follows $x_1 = -1$, which violates $t \neq 0$. So, $\lambda_4 > 0$ and $\lambda_6 = 0$. This results in the solution: $\lambda = (0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0)$ with $(t, x_1, x_2) = (\frac{1}{2}, -\frac{1}{2}, 2)$.

   In a similar way, one can examine the other two remaining cases, $x_1 - t + 1 > 0$ with $x_1 + t + 1 = 0$ and $x_1 - t + 1 > 0$ with $x_1 + t + 1 = 0$, respectively, which yield the same solution.

Observe that with $y = -\frac{1}{2} + 2x$ we have at $x_1 = 0,\ y_1 = -\frac{1}{2}$; at $x_2 = 1,\ y_2 = \frac{3}{2}$; and at $x_3 = 2,\ y_3 = \frac{7}{2}$. The Chebyshev distance between $(-1, 2, 3)$ and $(-\frac{1}{2}, \frac{3}{2}, \frac{7}{2})$ is indeed $\frac{1}{2}$.

EXERCISE 4.18 *(Optimization problem equivalence under change of variables)*
Consider the maximization of $f(x)$ subject to the constraints $g_i(x) \leq b_i,\ i = 1, \cdots, k$, and $h_j(x) = c_j,\ j = 1, \cdots, m$, on a set $C \subset \mathbb{R}^n$.
Assume $\phi(x) = [\phi_1(x) \cdots \phi_n(x)]^T$ is a mapping from $D \subset R^n$ to $\mathbb{R}^n$ which maps no two elements of its domain to a single value in its range (*one-to-one mapping*). Moreover assume $\phi(x)$ is such that $C \subset \phi(D)$.

a. Show that above optimization problem has a solution at x^* iff. with $\tilde{f}(z) := f(\phi(z))$, $\tilde{g}_i(z) := g_i(\phi(z))$ and $\tilde{h}_i(z) := h_i(\phi(z))$, the maximization of $\tilde{f}(z)$ subject to the constraints $\tilde{g}_i(z) \leq b_i,\ i = 1, \cdots, k$, and $\tilde{h}_j(z) = c_j,\ j = 1, \cdots, m$, has a solution at some point z^*. What is the relationship between the solution of both problems and between x^* and z^*?
b. Use the approach sketched in item a. to find the maximum of $f(x_1, x_2) = ax_1^{\alpha} + \ln(x_2)$ subject to $x_1^{\alpha}\ln(x_2) \leq 4$. Here a, α and x_i are strictly positive numbers.

ANSWER 4.18

a. By assumption, for every $x \in C$ there exists a $z \in D$ such that $x = \phi(z)$. In particular, for x^* there exists a z^* with $x^* = \phi(z^*)$. If f attains its maximum on the constraint set at x^*, then by definition,

$$f(x) \leq f(x^*) \text{ for all } x \text{ in the constraint set.}$$

That is, there exists a set $D_1 \subset \mathbb{R}^n$ such that for all $z \in D_1$,

$$f(\phi(z)) \leq f(\phi(z^*)),\ g(\phi(z)) \leq b_i, \text{ and } h(\phi(z)) = c_j.$$

Moreover, since ϕ is a one-to-one mapping, there are no other points outside D_1 satisfying the constraints. In other words, the constraint optimization problem of maximizing

$$\tilde{f}(z) \text{ subject to } \tilde{g}_i(z) \leq b_i,\ i = 1, \cdots, k, \text{ and } \tilde{h}_j(z) = c_j,\ j = 1, \cdots, m,$$

has a solution at some point z^*.

The converse follows in a similar manner. That is, if ϕ^{-1} denotes the inverse of ϕ on D_1, then $z^* = \phi^{-1}(x^*)$ and the maximum value of the (equivalent) optimization problem is $\tilde{f}(z^*)$.

b. Consider $\phi(x_1, x_2) = [x_1^{\alpha_1}\ \ln(x_2)]^T$. One directly verifies that for every $(z_1, z_2) \in \mathbb{R}^2$ with $z_i > 0$, $\phi(x_1, x_2) = [z_1, z_2]$ has a unique solution with $x_i > 0$. So, ϕ is an invertible mapping on the positive quadrant in $\mathbb{R}^2$.

The corresponding equivalent optimization problem is to maximize $\tilde{f}(z_1, z_2) = az_1 + z_2$ subject to the constraint $z_1 z_2 \leq 4$ (with $z_i > 0$).

The corresponding Lagrange function for this problem is: $L(z_1, z_2) = az_1 + z_2 - \lambda(z_1 z_2 - 4)$.

From this we obtain the necessary conditions:

$$\text{1.a) } a - \lambda z_2 = 0;\ \text{b) } 1 - \lambda z_1 = 0;\ \text{2) } \lambda(z_1 z_2 - 4) = 0;\ \text{3.a) } \lambda \geq 0 \text{ and b) } z_1 z_2 \leq 4.$$

From 1.b) we obtain $\lambda = \frac{1}{z_1} \neq 0$. So, by 2), $z_1 z_2 = 4$. Together with conditions 1 we obtain $z_2 = az_1$, with $z_1 = \frac{2}{\sqrt{a}}$.

Note that the conditions for applying Weierstrass' theorem are satisfied. Consequently, the maximum value of the problem is $4\sqrt{a}$, which is attained at $(x_1, x_2) = (\left(\frac{2}{\sqrt{a}}\right)^{\frac{1}{\alpha_1}}, e^{2\sqrt{a}})$.

EXERCISE 4.19 *(Geometric programming)*

In this exercise we consider the so-called Geometric programming problem. In this problem objective and constraint functions are described by so-called posynomial functions. Using a transformation this problem can be rewritten as a convex optimization problem. You are asked to complete the details to reformulate this geometric programming problem as a convex optimization problem. Extensions and more information on this problem can be found in, e.g., [5][Chapter 4.5].

Consider the *monomial* $p_i(x) := a_i x_1^{\alpha_{1i}} * \cdots * x_n^{\alpha_{ni}}$, where $a_i > 0$, $x_i > 0$ and $\alpha_{ij} \in \mathbb{R}$. Then

$f(x) := \sum_{i=1}^{k} p_i(x)$ is called a *posynomial* .
Let $f(x)$ and $g_i(x)$ be posynomials and b_i strictly positive numbers. The geometric programming problem is defined as:

$$\min_x f(x), \text{ subject to the constraints } g_i(x) \leq b_i.$$

a. Let $y_i := \ln(x_i)$. Use this transformation (see Exercise 4.18), to formulate the corresponding equivalent optimization problem of maximizing $\tilde{f}(y) := -\sum_{i=1}^{k} \tilde{p}_i(y)$ subject to $\tilde{g}_i(y) \leq b_i$. Here, with $\alpha_i := [\alpha_{1i} \cdots \alpha_{ni}]^T$ and $\beta_i := \ln(a_i)$, $\tilde{p}_i(y) = e^{\alpha_i^T y + \beta_i}$.
b. Show that $D\tilde{p}_i(y) = \alpha_i^T \tilde{p}_i(y)$ and $H\tilde{p}_i(y) = \alpha_i \alpha_i^T \tilde{p}_i(y)$.
c. Show that $\tilde{f}(y)$ is a concave function and $\tilde{g}_i(y)$ are convex functions.
d. Use above approach to find the maximum of $f(x_1, x_2) = \frac{x_2^2 + 2x_1x_2 + 1}{x_1^2 x_2^2}$ subject to $x_i \geq 1,\ i = 1, 2$. You may use a result from next chapter that the necessary KKT conditions are in this case sufficient too.

ANSWER 4.19

a. With $y_i := \ln(x_i)$, $\phi(x) = [e^{x_1}, \cdots, e^{x_n}]$ maps the positive orthant in $\mathbb{R}^n$ one-to-one to itself. Using this transformation, by Exercise 4.18, we obtain the equivalent optimization problem: $p_i(x) := a_i x_1^{\alpha_{1i}} * \cdots * x_n^{\alpha_{ni}} = e^{\ln(a_i)} e^{\alpha_{1i} y_1} * \cdots * e^{\alpha_{ni} y_n} = e^{\alpha_{1i} y_1 + \cdots + \alpha_{ni} y_n + \ln(a_i)} = e^{\alpha_i^T y + \beta_i}$, where $\alpha_i = [\alpha_{1i} \cdots \alpha_{ni}]^T$ and $\beta_i = \ln(a_i)$.
b. Straightforward differentiation shows this result.
c. From item b. it follows that

$$H(-\tilde{f}(y)) = \sum_{i=1}^{k} H\tilde{p}_i(y) = \sum_{i=1}^{k} \alpha_i \alpha_i^T \tilde{p}_i(y).$$

As $\alpha_i \alpha_i^T$ is a positive semi-definite matrix and $\tilde{p}_i(y) > 0$ for all y it follows that $H(-\tilde{f}(y))$ is a positive semi-definite matrix for all y. That is, $-\tilde{f}(y)$ is convex and, therefore, $\tilde{f}(y)$ is concave on its domain. Similarly it follows that $\tilde{g}_i(y)$ are convex functions on the domain.
d. With $y_i = \ln(x_i)$ the equivalent convex optimization problem is

$$\max_{y_i} \tilde{f}(y_1, y_2) = e^{-2y_1} + 2e^{-y_1 - y_2} + e^{-2y_1 - 2y_2} \text{ subject to } g_i(y_1, y_2) = e^{y_i} \geq 1,\ i = 1, 2.$$

The with this problem associated Lagrange function is

$$L(y_1, y_2, \lambda_1, \lambda_2) := e^{-2y_1} + 2e^{-y_1 - y_2} + e^{-2y_1 - 2y_2} + \lambda_1(e^{y_1} - 1) + \lambda_2(e^{y_2} - 1).$$

From which we obtain the necessary conditions

1.a) $-2e^{-2y_1} - 2e^{-y_1-y_2} - 2e^{-2y_1-2y_2} + \lambda_1 e^{y_1} = 0$; 1.b) $-2e^{-y_1-y_2} - 2e^{-2y_1-2y_2} + \lambda_2 e^{y_2} = 0$;
2.a) $\lambda_1(e^{y_1} - 1) = 0$; 2.b) $\lambda_2(e^{y_2} - 1) = 0$; 3.a) $\lambda_i \geq 0,\ i = 1, 2$; and $e^{y_i} \geq 1,\ i = 1, 2$.

From items 1. and 3.a) it is obvious that $\lambda_i > 0,\ i = 1, 2$. So, by item 2, $e^{y_i} = 1,\ i = 1, 2$. That is, $y_i = 0,\ i = 1, 2$. Since this is a convex optimization problem (and Slater's condition is met) it follows (see next chapter) that this point yields the maximum of the problem. Or, at $x_i = 1,\ i = 1, 2$. the problem attains its maximum.

EXERCISE 4.20 *(Linear-fractional programming)*
Consider the problem to maximize the ratio of two affine functions over a constraint set described by linear inequalities. In this exercise you are asked to show that this problem can be reformulated as a linear programming problem.
Let $f(x) = \frac{a_1^T x + d_1}{a_2^T x + d_2}$ and $g_j(x) = c_j^T x,\ j = 1, \cdots, k$, where $a_i,\ c_j \in \mathbb{R}^n$ and $d_i,\ b_j \in \mathbb{R}$, $i = 1, 2;\ j = 1, \cdots, k$.
Then, the linear-fractional optimization problem is formalized as the maximization of $f(x)$ on the nonempty set defined by $C = \{x \in \mathbb{R}^n \mid g_j(x) \leq b_j,\ j = 1, \cdots, k,$ and $a_2^T x + d_2 > 0\}$.

a. Show that if $\hat{x}$ is a solution for the linear-fractional optimization problem, with $\hat{y} := \frac{\hat{x}}{a_2^T \hat{x} + d_2}$ and $\hat{z} := \frac{1}{a_2^T \hat{x} + d_2}$, $(\hat{y}, \hat{z})$ is a solution to the quasi linear programming problem:
max$_{y,z}\ a_1^T y + d_1 z$ subject to $c_j^T y - b_j z \leq 0,\ j = 1, \cdots, k$, $a_2^T y + d_2 z = 1$ and $z > 0$.

b. Show that if $(\hat{y}, \hat{z})$ is a solution to the quasi linear programming problem:
max$_{y,z}\ a_1^T y + d_1 z$ subject to $c_j^T y - b_j z \leq 0,\ j = 1, \cdots, k$, $a_2^T y + d_2 z = 1$ and $z > 0$, the linear-fractional programming problem has the solution $\hat{x} = \frac{\hat{y}}{\hat{z}}$.

c. Let $(\hat{y}, \hat{z})$ be a solution to the quasi linear programming problem:
max$_{y,z}\ a_1^T y + d_1 z$ subject to $c_j^T y - b_j z \leq 0,\ j = 1, \cdots, k$, $a_2^T y + d_2 z = 1$ and $z > 0$.
Show that $(\hat{y}, \hat{z})$ also solves the linear programming problem:
max$_{y,z}\ a_1^T y + d_1 z$ subject to $c_j^T y - b_j z \leq 0,\ j = 1, \cdots, k$, $a_2^T y + d_2 z = 1$ and $z \geq 0$. Hint: show this by assuming this is not the case.

d. Assume the constrained set in the linear-fractional programming problem is bounded.
Show that the linear-fractional programming problem has a solution iff. the linear programming problem formulated in item c. has a solution.
Show that if, moreover, in this case $d_2 \neq 0$ the solution of this linear programming problem can be found by determining the maximum w.r.t. y of the linear program $(a_1^T - \frac{d_1}{d_2} a_2^T)y + \frac{d_1}{d_2}$ subject to the constraints $\frac{1}{d_2}(1 - a_2^T y) \geq 0$ and $(c_j^T + \frac{b_j}{d_2} a_2^T)y \leq \frac{b_j}{d_2},\ j = 1, \cdots, k$.

e. Consider a company producing 2 products. Assume its production of product i is x_i. The profit of product i is $2x_1$ and $3x_2$, respectively. Whereas the cost of producing product i is $x_1 + 2$ and $2x_2 + 3$, respectively. The production constraints are $x_i \geq 0$ and $x_1 + x_2 \leq 10$. Assume the company wants to maximize its efficiency expressed as the ratio of total profit divided by total cost. Formulate

the corresponding linear-fractional programming problem, and solve it using the above introduced equivalent linear programming problem.

ANSWER 4.20

a. Assume that $\hat{x}$ is a solution for the linear-fractional optimization problem. Then

$$f(\hat{x}) \geq f(x) \text{ for all } x \text{ satisfying } g_j(x) \leq b_j \text{ and } a_2^T x + d_2 > 0.$$

Let $y := \frac{x}{a_2^T x + d_2}$ and $z := \frac{1}{a_2^T x + d_2}$. It follows immediately that the constraint set C can be described in terms of (y, z) as

$$\{(y, z) \mid z > 0,\ a_2^T y + d_2 z = 1 \text{ and } c_j^T y \leq b_j^T z\}.$$

Now define

$$\hat{y} := \frac{\hat{x}}{a_2^T x + d_2} \text{ and } \hat{z} := \frac{1}{a_2^T \hat{x} + d_2}.$$

Then

$$f(\hat{y}, \hat{z}) = a_1^T \hat{y} + d_1 \hat{z} \geq f(y, z) = a_1^T y + d_1 z \text{ for all } (y, z) \in C.$$

b. By assumption, for all (y, z) in the constraint set we have

$$a_1^T \hat{y} + d_1 \hat{z} \geq a_1^T y + d_1 z.$$

Since $a_2^T y + d_2 z = 1$ for all (y, z) in the constraint set, it follows that

$$\frac{a_1^T \hat{y} + d_1 \hat{z}}{a_2^T \hat{y} + d_2 \hat{z}} \geq \frac{a_1^T y + d_1 z}{a_2^T y + d_2 z}.$$

Now let $x := \frac{y}{z}$ and $\hat{x} := \frac{\hat{y}}{\hat{z}}$. An elementary rearrangement of the above inequality yields

$$\frac{a_1^T \hat{x} + d_1}{a_2^T \hat{x} + d_2} \geq \frac{a_1^T x + d_1}{a_2^T x + d_2},$$

for all x satisfying $c_j^T x \leq b_j,\ j = 1, \cdots, k$, and $a_2^T x + d_2 > 0$.

c. Assume that $(\hat{y}, \hat{z})$ solves the linear programming problem under the restriction $z > 0$. Suppose there exists a point (y_0, z_0) such that $f(y_0, z_0) > f(\hat{y}, \hat{z})$ when we consider the enlarged constraint set where $z \geq 0$. Now consider the convex combination

$$(y_\lambda, z_\lambda) = \lambda(y_0, z_0) + (1 - \lambda)(\hat{y}, \hat{z}).$$

Then

$$c_j^T y_\lambda = c_j^T(\lambda y_0 + (1 - \lambda)\hat{y}) = \lambda c_j^T y_0 + (1 - \lambda) c_j^T \hat{y} \leq b_j;\ a_2^T y_\lambda + d_2 z_\lambda = 1 \text{ and } z_\lambda > 0 \text{ if } \lambda > 0.$$

If $z_0 > 0$, then (y_0, z_0) would belong to the feasible set of the lp-problem where $z > 0$. But this would contradict the optimality of $(\hat{y}, \hat{z})$.
However, if $z_0 = 0$, then for sufficiently small $\lambda > 0$ we have

$$f(y_\lambda, z_\lambda) > f(\hat{y}, \hat{z}),$$

while (y_λ, z_λ) belongs to the feasible set of the lp-problem with $z > 0$. This again contradicts the optimality of $(\hat{y}, \hat{z})$.
Therefore, our assumption that there exists a point (y_0, z_0) such that $f(y_0, z_0) > f(\hat{y}, \hat{z})$ must be false.

d. " $\Rightarrow$ " This implication follows from the previous item c.
" $\Leftarrow$ " Assume that the maximum is attained at $(\hat{y}, \hat{z}) = (\hat{y}, 0)$. Then $c_j^T \hat{y} \leq 0$ and $a_2^T \hat{y} = 1$. However, it is easy to verify that for any $\lambda > 0$, the point $\lambda \hat{y}$ also belongs to the constraint set C. This contradicts the assumption that C is bounded. Hence, the maximum can not be attained with $z = 0$; that is $\hat{z} > 0$. By item b., this implies that the linear-fractional programming problem has a solution.
The equivalent LP problem is obtained immediately by substituting $z = \frac{1}{d_2}(1 - a_2^T y)$ into the extended LP problem.

e. The linear-fractional programming problem is to maximize, w.r.t. x_1 and x_2

$$\frac{2x_1 + 3x_2}{x_1 + 2x_2 + 5} \text{ subject to } x_i \geq 0,\ i = 1, 2, \text{ and } x_1 + x_2 \leq 10.$$

The corresponding linear programming problem is to maximize w.r.t. $y_1,\ y_2$ and z

$$f(y_1, y_2, z) = 2y_1 + 3y_2 \text{ subject to } y_1 + y_2 - 10z \leq 0,\ y_i \geq 0,\ i = 1, 2, y_1 + 2y_2 + 5z = 1 \text{ and } z \geq 0.$$

Equivalently, this can be written as the maximization of

$$f(y_1, y_2) = 2y_1 + 3y_2 \text{ subject to } y_i \geq 0, i = 1, 2,\ 3y_1 + 5y_2 \leq 2 \text{ and } 1 - y_1 - 2y_2 \geq 0.$$

Note that the constraint $y_1 + 2y_2 \leq 1$ is redundant in this case. Simple calculations show that the maximum is attained at $(y_1, y_2) = (\frac{2}{3}, 0)$.
Thus, $\hat{z} = \frac{1}{15}$, and consequently $(x_1, x_2) = (10, 0)$ solves the linear-fractional programming problem.

EXERCISE 4.21 *(Proof Proposition 4.3)*
In this exercise we provide a proof for Proposition 4.3. Let $f(s)$ be differentiable on an open set $S \subset \mathbb{R}^n$ and $x \in S$. Then, f is differentiable in x in the direction u. Furthermore, the directional

derivative of f in a in the direction u, denoted by $D_u f(x)$, equals

$$D_u f(x) = Df(x) * \frac{u}{\|u\|}.$$

Here, $\|u\|$ denotes the length of the vector $u = [u_1, \cdots, u_n]^T$: $\sqrt{u_1^2 + \cdots u_n^2}$.

a. Consider a fixed vector x in the interior of the domain of f and some fixed direction vector $u \in \mathbb{R}^n$. Show that there exist a $\delta > 0$ such that for all scalars $t \in (-\delta, \delta)$, $x + tu$ belongs to the interior of the domain of f.

b Show that the scalar function $F(t) := f(g(t))$, where $g(t) := x + tu$, is differentiable for $t \in (-\delta, \delta)$.

c. Show that the derivative of the section function at point x equals $\lim_{t \downarrow 0} \frac{F(t)-F(0)}{t\|u\|}$.

d. Show that $D_u f(x) = Df(x) * \frac{u}{\|u\|}$.

ANSWER 4.21

a. As x is in the interior of the domain of f, by definition there exists a μ such that all y satisfying $\|x - y\| < \mu$, belong to the interior of the domain of f. So, with $\delta := \frac{\mu}{\|u\|}$, we have that for all $t \in (-\delta, \delta)$ $y := x + tu$ satisfy $\|x - y\| = \|tu\| = t\|u\| < \mu$.

b. $g(t)$ is differentiable. f is differentiable at every $t \in (-\delta, \delta)$ at $g(t)$. So, by the chain rule of differentiation, $F(t)$ is differentiable at every $t \in (-\delta, \delta)$. Moreover, $F^{'}(t) = Df(g(t))u$.

c. The section function around point x is given by $f(x + tu)$, where $t \in (-\delta, \delta)$. Consequently, the difference quotient of the section function at x relative to the point $x + tu$ for $t > 0$, is

$$\frac{f(x + tu) - f(x)}{t\|u\|},$$

where the denominator represents the distance between the points $x + tu$ and x).
By definition, the directional derivative of f at x in the direction u is obtained by taking the limit of this difference quotient if $t > 0$ approaches zero.

d. By item b, for all t_0 in a neighborhood of $t_0 = 0$, the difference quotient $\frac{F(t_0+t)-F(t_0)}{t}$ has a limit if t approaches zero. In particular, for $t_0 = 0$, it follows that the difference quotient $\frac{f(x+tu)-f(x)}{t}$ has a limit if t approaches zero. By item b, this limit equals $Df(g(0))u = Df(x)u$. It follows that the difference quotient $\frac{f(x+tu)-f(x)}{t\|u\|}$ also has a limit if t approaches zero namely $\frac{Df(x)u}{\|u\|}$.
Therefore, the directional derivative of f at x in the direction u exists and equals $\frac{Df(x)u}{\|u\|}$.

EXERCISE 4.22 *(Exercise where KKT conditions don't reduce the set of potential solutions)*
Consider the problem to minimize $f(x, y) = x^2 + y^2$ subject to the constraints $x \geq 0$; $y \geq 0$ and $xy = 0$.

a. Show that all points where $xy = 0$ are obtained as potential candidates for a solution of the problem in Step 2 of Algorithm 4.20.

b. Show, using Fritz-John's theorem, that all points (x, y) where $xy = 0$ qualify as a candidate solution.

c. Solve the minimization problem.

ANSWER 4.22

a. To find all candidate solutions we proceed along the lines of Algorithm 4.20.

Step 1. Note that with $g_1(x,y) := -x$, $g_2(x,y) := -y$ and $h(x,y) := xy$ all functions $f,\ g_1,\ g_2$ and h are continuous differentiable.

Step 2. Consider, see Remark 4.28, matrix

$$\tilde{G} := \begin{bmatrix} -1 & 0 & -x & 0 \\ 0 & -1 & 0 & -y \\ y & x & 0 & 0 \end{bmatrix}.$$

The rank of this matrix equals the rank of matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ y & x & 1 \end{bmatrix} \tilde{G} = \begin{bmatrix} -1 & 0 & -x & 0 \\ 0 & -1 & 0 & -y \\ 0 & 0 & -xy & -xy \end{bmatrix}.$$

Clearly, this lastmentioned matrix has not full row rank iff. $xy = 0$.

Consequently, all points where $xy = 0$ qualify as a candidate solution for the problem.

b. Consider

$$L(x,y,\lambda_0,\lambda_1,\lambda_2,\mu) := -\lambda_0(x^2+y^2) + \lambda_1 x + \lambda_2 y - \mu xy.$$

By Fritz-John's theorem, if the problem attains a maximum, then the following system of (in)equalities holds:

1.i) $-2\lambda_0 x + \lambda_1 - \mu y = 0$; ii) $-2\lambda_0 y + \lambda_2 - \mu x = 0$;

2.i) $\lambda_1 x = 0$; ii) $\lambda_2 y = 0$;

3.i) $x \geq 0$; ii) $y \geq 0$, iii) $xy = 0$ and iv) $\lambda_0 \in \{0,1\},\ \lambda_i \geq 0,\ i = 1,2$;

4. $(\lambda_0,\lambda_1,\lambda_2) \neq (0,0,0)$.

Assume $\lambda_0 = 0$. Then by 1. $\lambda_1 = \mu y$ and $\lambda_2 = \mu x$. One easily verifies that now all points $(x,y) \neq (0,0)$ with either $x = 0$ or $y = 0$ satisfy 1-4.

Next, consider $(x,y,\lambda_0,\lambda_1,\lambda_2) = (0,0,1,0,0)$. One easily verifies that this point also satisfies all conditions 1-4.

Thus, Fritz-John's theorem also implies that all points satisfying $xy = 0$ are candidate solutions to the problem.

c. So, apparently at any point of the constraint set the minimum may occur.

From $xy = 0$, we have that either $x = 0$ or $y = 0$. If $x = 0$, the problem reduces to the problem to minimize y^2 subject to the constraint $y \geq 0$. This yields $(x,y) = (0,0)$ as the candidate solution. Similarly, if $y = 0$, we obtain $(x,y) = (0,0)$ as the candidate solution. Since $f(x,y) \geq 0$, it is clear that $(0,0)$ solves the problem.

EXERCISE 4.23 *(Optimization over the nonnegative orthant)*
Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous differentiable. Consider the problem to maximize $f(x)$ on the non-negative orthant, i.e., on the constraint set $\{x \in \mathbb{R}^n \mid x_i \geq 0,\ i = 1, \cdots, n\}$.
Show that at a maximum location a next conditions apply:

$$1.\ D_i f(a) \leq 0,\ i = 1, \cdots, n;\ 2.\ a_i D_i f(a) = 0;\ \text{and } 3.\ a_i \geq 0,\ i = 1, \cdots, n.$$

ANSWER 4.23 We proceed along the lines of Algorithm 4.20.
Step 1. Note that with $g_i(x) := -x_i$ all functions f and g_i are continuous differentiable.
Step 2. Consider matrix $G := [Dg(x) \mid \text{diag}(g_i(x))] = [-I \mid \text{diag}(x_i)]$.
Note that matrix G has always full row rank. So, all potential maximum locations are found as solutions from the KKT conditions formalized below in Step 3.
Step 3. Let $L(x, \lambda_1, \cdots, \lambda_m) := f(x) + \sum_{i=1}^m \lambda_i x_i$.
Step 4. Then, by Kuhn-Tucker's theorem, at a maximum location a of the problem next 3 conditions are satisfied:

$$1.\ [D_1 L(a, \lambda) \cdots D_n L(a, \lambda)] = 0;\ 2.\ \lambda_i a_i = 0;\ 3.\ a_i \geq 0 \text{ and } \lambda_i \geq 0.$$

Condition 1 reduces to $D_i f(a) + \lambda_i = 0$. Since $\lambda_i \geq 0$, we conclude from this that $D_i f(a) \leq 0$. Moreover, we obtain from this equality that $\lambda_i = -D_i f(a)$. So, by condition 2. it follows then that $D_i f(a) a_i = 0$.

EXERCISE 4.24 *(Barrier Interior-Point Method to solve numerically a constrained optimization problem)*
In this exercise we discuss a numerical approach, the so-called *Barrier Method*, to find the solution of a constrained optimization problem (under the assumption a solution exists).
Consider the problem to find the maximum of $f : \mathbb{R}^n \to \mathbb{R}$ subject to $g(x) \leq b$.

a. Let $I(x) : \mathbb{R} \to \mathbb{R}$ be the function that attains an arbitrarily small value "$-\infty$" if $x > 0$ and a value 0 if $x \leq 0$. Assume there is just one constraint. (this can be generalized simply to the case that there are m constraints). Show that the constrained optimization problem has a solution iff. $f_I(x) := f(x) + I(g(x) - b)$ has a maximum (this result can be generalized simply to the case that there are m constraints).
b. Let $t > 0$. Consider the function $\tilde{I}_t(x) = \frac{1}{t} \ln(-x)$ for $x < 0$. Plot $\tilde{I}_1(x)$, $\tilde{I}_{10}(x)$ and $\tilde{I}_{100}(x)$ in one graph. What can you say about $\tilde{I}_t(x)$ in relationship with $I(x)$.
c. Consider the necessary conditions for the constrained maximization problem. Furthermore consider for a fixed $t > 0$, the first-order conditions for a maximum of the function $\tilde{f}_t(x) := f(x) + \frac{1}{t} \sum_{i=1}^m \ln(b_i - g_i(x))$ on the constrained set $S := \{x \mid g_i(x) < b_i,\ i = 1, \cdots, m\}$. Assume that for all $t > 0$ this unconstrained optimization problem has a unique solution $x^*(t)$. Show that if $x^*(t)$ converges, for $t \to \infty$, this converged solution solves the KKT conditions of the constrained

maximization problem. Hint: use $\lambda_i(t) := \frac{1}{t}\frac{1}{b_i - g_i(x^*(t))}$ to rewrite the first-order conditions in the unconstrained optimization problem.

d. Consider the constrained maximization problem and the associated Lagrange function $L(x,\lambda) = f(x) - \sum_{i=1}^m \lambda_i(g_i(x) - b_i)$. Show that for all $\lambda \geq 0$, $L(x,\lambda) \geq f(x)$ and $\min_{\lambda \geq 0} L(x,\lambda) = f(x)$. Conclude from this that $f(x^*) = \max_{x \in S} \min_{\lambda \geq 0} L(x,\lambda)$ (i).
Show that with $\lambda_i := \lambda_i(t) := \frac{1}{t}\frac{1}{b_i - g_i(x^*(t))}$, $x^*(t)$ is a stationary point of $L(x,\lambda(t))$.
Assume $L(x,\lambda(t))$ has its unique maximum at $x^*(t)$ (ii). Use (i) and (ii) to conclude that $f(x^*) \leq f(x^*(t)) + \frac{m}{t}$.

e. Consider, for a fixed $t > 0$, the function $\tilde{f}_t(x) := f(x) + \frac{1}{t}\sum_{i=1}^m \ln(b_i - g_i(x))$ on the constrained set $S := \{x \mid g_i(x) < b_i,\ i = 1, \cdots, m\}$.
Assume $\tilde{f}_t(x)$ has a unique maximum at $x^*(t)$ for all $t > 0$ and we want to approximate the maximum of our constrainted optimization problem with an error of at most $\epsilon > 0$. Use item d to show that if we choose $t \geq \frac{m}{\epsilon}$ we achieve this goal.

f. Consider for a fixed $t > 0$ the Newton-Raphson procedure (see Exercise 2.48) to approximate the maximum location $x^*(t)$ of $\tilde{f}_t(x)$.

$$x_{N+1} = x_N - H\tilde{f}_t^{-1}(x_N) D\tilde{f}_t^T(x_N),\ x_0 = \bar{x} \in S,$$

Show that

$$\begin{aligned} D\tilde{f}_t(x) &= Df(x) - \frac{1}{t}\sum_{i=1}^m \frac{1}{b_i - g_i(x)} Dg_i(x) \text{ and} \\ H\tilde{f}_t(x) &= Hf(x) - \frac{1}{t}\sum_{i=1}^m \left(\frac{1}{(b_i - g_i(x))^2} Dg_i^T(x) Dg_i(x) + \frac{1}{b_i - g_i(x)} Hg_i(x)\right). \end{aligned}$$

g. Since it is usually difficult to find for a large value of t the maximum location $x^*(t)$ of $\tilde{f}_t(x)$ one uses often a recursive procedure to find this location. One, often used, approach is to first approximate $x^*(t)$ for a smaller value of $\bar{t}$ (e.g. $\bar{t} = 1$) and next use this approximation $x^*(\bar{t})$ as an initialization point to approximate $x^*(t)$ for a larger value than $\bar{t}$. This process is then repeated until $t > \frac{m}{\epsilon}$, where ϵ is the maximal allowed error towards the sought-after maximum $f(x^*)$. Implement this method starting with $t = 1$ and doubling this value each iteration (using the Newton-Raphson procedure outlined in item f) to find the maximum of $f(x,y,z) = -x^2 - e^y - z^2$ on the constraint set $S = \{(x,y,z) \mid x^2 + y^2 \leq 16 \text{ and } x^4 + z^2 \leq 25\}$. Observe that with $t = 200$ we approximate the maximum of $-e^{-4}$ within a distance of 0.01.

ANSWER 4.24

a. Note that $f_I(x)$ coincides with $f(x)$ for all x satisfying the constraint $g(x) \leq b$ and $f_I(x) = -\infty$ elsewhere.

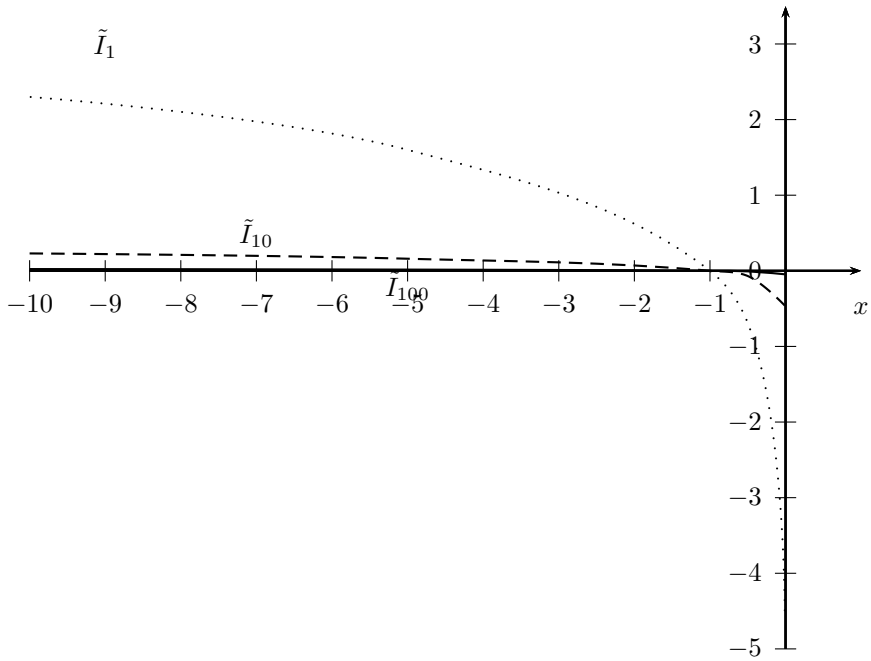


FIGURE 4.2 – $\tilde{I}_j,\ j = 1, 10, 100.$

b. Observe that if t gets larger, $\tilde{I}_t$ approaches the function which is almost zero for $x \in (-a, 0)$, where a is large, and which diverges to "$-\infty$" for values of x very close to zero. Note that for every t, $\tilde{I}_t(x)$ is infinitely many times differentiable for all $x < 0$.

c. If, for a fixed t, $x^*(t)$ is a solution for the unconstrained optimization problem the first-order conditions are

$$D\tilde{f}_t(x^*(t)) = Df(x^*(t)) - \frac{1}{t}\sum_{i=1}^{m} \frac{1}{b_i - g_i(x^*(t))} Dg_i(x^*(t)) = 0, \text{ where } g_i(x^*(t)) < b_i,\ i = 1, \cdots, m.$$

The KKT conditions for the constrained maximization problem are, with

$$L(x, \lambda) := f(x) - \sum_{i=1}^{m} \lambda_i (g_i(x) - b_i)$$

1. $D_x L(x^*, \lambda) = Df(x^*) - \sum_{i=1}^{m} \lambda_i Dg_i(x^*) = 0$; 2. $\lambda_i(g_i(x^*) - b_i) = 0$; 3. $\lambda_i \geq 0$ and $g_i(x^*) \leq b_i$.

Now, let $\lambda_i(t) := \frac{1}{t}\frac{1}{b_i - g_i(x^*(t))}$ in the unconstrained optimization problem.
Then the first-order condition is equivalent with

$$1.\ Df(x^*(t)) - \sum_{i=1}^{m} \lambda_i(t) Dg_i(x^*(t)) = 0;\ 2.\ \lambda_i(t)(b_i - g_i(x^*(t)) = \frac{1}{t};$$
$$3.\ \lambda_i(t) > 0 \text{ and } g_i(x^*(t) - b_i) < 0.$$

From this it follows that, if $x^*(t)$ converges as $t \to \infty$, the limit satisfies the first-order optimality conditions of the constrained optimization problem.

d. Since $g_i(x) \leq b$, for all $\lambda_i \geq 0$ we have $-\lambda_i(g_i(x) - b_i) \geq 0$. Therefore,

$$L(x, \lambda) = f(x) - \sum_{i=1}^{m} \lambda_i(g_i(x) - b_i) \geq f(x) \text{ and } \min_{\lambda \geq 0} L(x, \lambda) = f(x) (\text{ by choosing } \lambda_i = 0).$$

So, with $x^* = \arg\max_{x \in S} f(x)$, we have $f(x^*) = \max_{x \in S} \min_{\lambda \geq 0} L(x, \lambda)$.
Next choose, for a fixed t, $\lambda_i = \lambda_i(t) := \frac{1}{t}\frac{1}{b_i - g_i(x^*(t))}$ in $L(x, \lambda)$. That is, consider $L(x, \lambda(t))$. Then $DL_x(x^*(t), \lambda(t)) = 0$ (see item c.).
Assuming $L(x, \lambda(t))$ has a unique maximum at $x^*(t)$, we have:

$$\begin{aligned} f(x^*) &= \max_{x \in S} \min_{\lambda \geq 0} L(x, \lambda) \leq \max_{x \in S} L(x, \lambda(t)) = f(x^*(t)) - \sum_{i=1}^{m} \lambda_i(t)(g_i(x^*(t)) - b_i) \\ &= f(x^*(t)) + \frac{m}{t}. \end{aligned}$$

e. From item d we have that $f(x^*) - f(x^*(t)) \leq \frac{m}{t}$. So, if we want this error to be smaller than ϵ, we should choose t such that $\frac{m}{t} \leq \epsilon$.

f. Straightforward differentiation of $\tilde{f}_t(x)$ and its derivative yields the advertized result.

g. With $f(x, y, z) = -x^2 - e^y - z^2$; $g_1(x, y, z) := x^2 + y^2$ and $g_2(x, y, z) := x^4 + z^2$ we have

$$Df = [-2x\ -e^y\ -2z];\ Dg_1 = [2x\ 2y\ 0] \text{ and } Dg_2 = [4x^3\ 0\ 2z].$$

Therefore,

$$Hf = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -e^y & 0 \\ 0 & 0 & -2 \end{bmatrix};\ Hg_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix};\ \text{and } Hg_2 = \begin{bmatrix} 12x^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Consequently,

$$\begin{aligned} D\tilde{f}_t(x) &= Df(x) - \frac{1}{t}\sum_{i=1}^{m} \frac{1}{b_i - g_i(x)} Dg_i(x) \\ &= [-2x\ -e^y\ -2z] - \frac{1}{t}\frac{1}{16 - (x^2 + y^2)}[2x\ 2y\ 0] - \frac{1}{t}\frac{1}{25 - (x^4 + z^2)}[4x^3\ 0\ 2z], \end{aligned}$$

and

$$\begin{aligned} H\tilde{f}_t(x) &= Hf(x) - \frac{1}{t}\sum_{i=1}^{m}\left(\frac{1}{(b_i - g_i(x))^2} Dg_i^T(x)Dg_i(x) + \frac{1}{b_i - g_i(x)} Hg_i(x)\right) \\ &= \begin{bmatrix} -2 & 0 & 0 \\ 0 & -e^y & 0 \\ 0 & 0 & -2 \end{bmatrix} - \frac{1}{t}\frac{1}{(16-(x^2+y^2))^2}\begin{bmatrix} 4x^2 & 4xy & 0 \\ 4xy & 4y^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &\quad -\frac{1}{t}\frac{1}{(25-(x^4+z^2))^2}\begin{bmatrix} 16x^6 & 0 & 8x^3z \\ 0 & 0 & 0 \\ 8x^3z & 0 & 4z^2 \end{bmatrix} \\ &\quad -\frac{1}{t}\frac{1}{16-(x^2+y^2)}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \frac{1}{t}\frac{1}{25-(x^4+z^2)}\begin{bmatrix} 12x^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}. \end{aligned}$$

Next, use for $t = 1, 2, 4, 16, .., 256,$ the iteration scheme

$$x_{N+1} = x_N - H\tilde{f}_t^{-1}(x_N)D\tilde{f}_t^T(x_N),\ x_0 = \bar{x} \in S,$$

until the distance between 2 successive iterations is smaller than, e.g., 0.01. For $t = 1$ you may use, e.g., $\bar{x} = [1\ 1\ 1]^T$ to start the procedure.

EXERCISE 4.25 *(Proof Theorem 4.35)*
Consider the constrained maximization problem (4.1), where $f,\ g_i,\ i = 1, \cdots, k, \in C^1$. Assume that the problem has a solution at $s = a$ and all constraints $g_i,\ i = 1, \cdots, k_1$, are binding at this point while the remaining constraints are non-binding. Let

$$L(s, \lambda_0, \lambda_1, \cdots, \lambda_k) := \lambda_0 f(s) - \lambda_1(g_1(s) - b_1) - \cdots - \lambda_k(g_k(s) - b_k)$$

denote the Lagrange function associated with this problem.
Below, you are asked to show - using Minkowski's[2] separation theorem - that, for the Lagrangian function defined above, the necessary conditions stated in Theorem 4.35 apply.
Minkowski's separation theorem states that if $Z_i,\ i = 1, 2,$ are two disjoint nonempty convex subsets of $\mathbb{R}^n$, then there exists a hyperplane $a^T x = b$ that seperates both sets. That is,

$$a^T x \leq b, \text{ for all } x \in Z_1 \text{ and } a^T x \geq b, \text{ for all } x \in Z_2.$$

(see also Lemma 5.7, where you are asked to prove a special case of this result).

2 Minkowski was a Polish mathematician who lived from 1864-1909

a. Show that there does not exist a vector $d \in \mathbb{R}^n$ such that the inequalities $Df(a)d > 0$ and $Dg_i(a)d < 0,\ i = 1, \cdots, k_1$, hold simultaneously.
b. Define the sets

$$Z_1 := \{z \mid [-Df^T(a)\ Dg_1^T(a)\ \cdots\ Dg_{k_1}^T(a)]^T x, x \in \mathbb{R}^n\}$$

and

$$Z_2 := \{z \in \mathbb{R}^{k_1+1} \mid z_i < 0,\ i = 1, \cdots, k_1 + 1\}.$$

Show that both sets are convex.
c. Use Minkowski's separating hyperplane theorem to show that there exists a vector $a \geq 0$, with $a \neq 0$, such that $a^T z = 0$ for all $z \in Z_1$.
d. Show that, for the Lagrangian function defined above, the necessary conditions stated in Theorem 4.35 apply.

ANSWER 4.25

a. Recall that the directional derivative of g_i in the direction d is given by $D_d g_i(a) = Dg_i(a)d$.
Since, by assumption, $Dg_i(a)d < 0$, the vector d points into the interior of the set $g_i(x) \leq 0$.
Because f attains a maximum at $x = a$ on the constraint set, it follows that for any direction d satisfying $Dg_i(a)d < 0,\ i = 1, \cdots, k_1$, we must have $f(x) \leq f(a)$ along that direction. Hence, the directional derivative of f in the direction d satisfies $Df(a)d \leq 0$.
In other words, there exists no direction d such that $Dg_i(a)d < 0,\ i = 1, \cdots, k_1$, and simultaneously $Df(a)d > 0$.
b. Let $\lambda \in [0, 1]$.
Note that Z_1 is the image of a linear subspace. Let $A := [-Df^T(a)\ Dg_1^T(a)\ \cdots\ Dg_{k_1}^T(a)]^T$. Then, with $z_i = Ax_i,\ i = 1, 2$, also $\lambda z_1 + (1 - \lambda)z_2 = A(\lambda x_1 + (1 - \lambda)x_2)$ belongs to Z_1. Therefore, Z_1 is a convex set.
Furthermore, if $z_i \in Z_2,\ i = 1, 2$, then it is clear that every component of $z := \lambda z_1 + (1 - \lambda)z_2$ is also negative. Thus, $z \in Z_2$. So, Z_2 is a convex set too.
c. By item a, all vectors $z \in Z_1$ satisfy $z \geq 0$. Clearly, all $z \in Z_2$ satisfy $z < 0$. As both sets are convex, Minkowski's separating hyperplane theorem implies that there exists a vector $a \neq 0$ such that

$$a^T z \geq 0 \text{ for all } z \in Z_1 \text{ and } a^T z \leq 0 \text{ for all } z \in Z_2.$$

From the condition $a^T z \leq 0$ for all $z \in Z_2$, and in particular for the vectors
$e_i = [0, \cdots, 0, -1, 0, \cdots, 0]^T$, it follows that every entry of the vector a must be nonnegative. That is, $a \geq 0$.
Furthermore, from the definition of Z_1, we observe that if $z \in Z_1$, then also $-z \in Z_1$. Consequently,

for any $z \in Z_1$, we have both $a^T z \geq 0$ and $a^T(-z) \geq 0$, which implies $a^T z = 0$ for all $z \in Z_1$. In other words, there exists a vector $a \geq 0$, with $a \neq 0$, such that $a^T z = 0$ for all $z \in Z_1$.

d. Let $a := [a_0, a_1, \cdots, a_{k_1}] \neq 0$, with $a_i \geq 0$. From item c, we have $a^T Ax = 0$, for all $x \in \mathbb{R}^n$. In particular evaluating this expression at the vectors $e_i = [0, \cdots, 0, 1, 0, \cdots, 0]^T$ yields $a^T A = 0$. That is, there exist nonnegative constants, not all zero, such that

$$a_0 Df(a) - a_1 Dg_1(a) - \cdots - a_{k_1} Dg_{k_1} = 0.$$

Next, choose $\lambda_i = a_i,\ i = 0, \cdots, k_1$, and set $\lambda_i = 0$ for the non-binding constraints. With these choices of λ_i it follows that all four conditions in Theorem 4.35 are satisfied.

EXERCISE 4.26 *(Proof Theorem 4.45)*

In this exercise we provide a proof for Theorem 4.45 using Farkas' lemma.

Farkas' lemma states next result.

Let $A \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^n$. Then exactly one of the next two assertions holds:

1. There exists an $x \in \mathbb{R}^m$ such that $Ax = b$ and $x \geq 0$.
2. There exists a $y \in \mathbb{R}^n$ such that $A^T y \leq 0$ and $b^T y > 0$.

a. Assume $B \in \mathbb{R}^{k \times n}$ and the dimension of the null space $\mathcal{N}(B)$ of matrix B is s. Show that $d \in \mathcal{N}(B)$ iff. there exists a $u \in \mathbb{R}^k$ such that $d = Hu$, where $H \in \mathbb{R}^{n \times s}$ is a full column rank matrix.

b. Assume the conditions of Theorem 4.45 are satisfied. Show that there exist matrices $A_1,\ B_1$ and a vector c such that for all d satisfying $A_1 d \leq 0$ and $B_1 d = 0$ we have $c^T d \leq 0$.

c. Use item a and Farkas' lemma to show that under the conditions of Theorem 4.45 there exist a matrix $A^T := A_1 H$ and vector $b^T := c^T H$ such that $Ax = b$ and $x \geq 0$.

d. Use the fundamental theorem of linear algebra, that the null space $\mathcal{N}(B)$ of matrix B is orthogonal to the row space $\mathcal{R}(B^T)$, to obtain from item c the result that there exist $\lambda_i \geq 0,\ i = 1, \cdots, k_1$ and $\mu_i,\ i = 1, \cdots, m$, such that $Df(a) = \sum_{i=1}^{k_1} \lambda_i Dg_i(a) + \sum_{i=1}^{m} \mu_i Dh_i(a)$.

e. Show that under the conditions of Theorem 4.45 the conditions (KKTM) apply.

ANSWER 4.26

a. Note that the null space of any matrix constitutes a linear subspace. Since the dimension of the null space $\mathcal{N}(B)$ of matrix B is by assumption s, there exist s basis vectors that span this subspace. Consider the matrix H that has as its columns these s basis vectors. Then the advertized result follows. Note that $H \in \mathbb{R}^{n \times s}$.

b. Since the conditions of the theorem are met, we know that the set of all tangent directions is given by

$$S := \{d \mid Dg_i(a)d \leq 0,\ i = 1, \cdots, k_1,\ \text{and}\ Dh_i(a)d = 0,\ i = 1, \cdots, m.\}$$

That is, with $A_1 := [Dg_1^T(a)\ \cdots\ Dg_{k_1}^T(a)]^T$ and $B_1 := [Dh_1^T(a)\ \cdots\ Dh_m^T(a)]^T$,

$$S = \{d \mid A_1 d \leq 0 \text{ and } B_1 d = 0\}.$$

By Theorem 4.39, setting $c := Df^T(a)$, we have $c^T d \leq 0$ for all $d \in S$.

c. From item a, we know that a vector d satisfies $B_1 d = 0$ iff $d = Hx$ for some vector x, where the matrix H spans the null space of B_1. Define

$$S = \{x \mid A_1 Hx =: Ax \leq 0\}.$$

Then, for all x satisfying $Ax \leq 0$, we have $c^T Hx \leq 0$. Or, setting $b^T := c^T H$, it follows that for all x with $Ax \leq 0$, the inequality $b^T x \leq 0$ holds.

By Farkas' lemma, we can therefore conclude that, under the conditions stated in the theorem, there exists a matrix $A^T = A_1 H$, a vector $b^T = c^T H$, and a nonnegative vector x such that $Ax = b$.

d. From item c, we know that there exists a vector $x \geq 0$ such that

$$H^T A_1 x = H^T c, \text{ or equivalently }, H^T(A_1 x - c) = 0.$$

This implies that $A_1 x - c$ lies in the null space of H^T. By the fundamental theorem of linear algebra, it follows that there exists a vector v such that $A_1 x - c = B_1 v$. Defining $\lambda := x$ and $\mu := -v$, we obtain $c = A_1 \lambda + B_1 \mu$.

e. Basically this result follows from item d. Choosing additionally $\lambda_i = 0,\ i = k_{1+1}, \cdots, k$, one easily verifies all (KKTM) conditions are met.

EXERCISE 4.27 *(Linear independence constraint qualification implies fulfilment Mangasarian-Fromovitz constraint qualification)*

Show that if the Linear independence constraint qualification is satisfied in Theorem 4.43, also the Mangasarian-Fromovitz constraint qualification is met. Hint: consider the equation $D\tilde{g}_1(a) = [-1 \ \cdots \ -1 \ 0 \ \cdots \ 0]^T$.

ANSWER 4.27

By assumption, the matrix $D\tilde{g}_1(a)$ has full row rank. This implies (see linear algebra) that, for any vector b, the equation $D\tilde{g}_1(a)u = b$ has a solution, for instance $u = D\tilde{g}_1^T(a)(D\tilde{g}_1(a)D\tilde{g}_1^T(a))^{-1}b$.

Now take $b := [-1 \ \cdots \ -1 \ 0 \ \cdots \ 0]^T$. Then there exists a vector u such that $Dg_1(a)u < 0$ and $Dh(a) = 0$. These are precisely the conditions required by the Mangasarian-Fromovitz constraint qualification.

Problems with an Economic Content

EXERCISE 4.28 *(consumer choice problem)*

Consider a consumer who can buy basic consumption goods and luxury goods. The price of the consumption good is p_1 and that of the luxury good p_2. The minimum amount of consumption goods he needs for living is m_1. His total budget he can spend on the goods is b. Utility he obtains from

buying an amount of x and y of the consumption and luxury goods is $U(x, y)$. Both products have the property that if one gets too much of it, utility from it declines.

a. Formulate the constrained optimization problem for this consumer.

b. Show the problem has no solution if the budget b is smaller than $p_1 m_1$.

c. Assume the consumer is insatiable for either one of both goods. Show the problem attains its maximum either at $(m_1, \frac{b-p_1 m_1}{p_2})$, at $(\frac{b}{p_1}, 0)$ or at the point $(x, \frac{b-p_1 x}{p_2})$ where x solves the equation $\frac{p_1}{p_2} = \frac{U_x^{'}}{U_y^{'}}$.
Explain under which conditions the solution $(m_1, \frac{b-p_1 m_1}{p_2})$ occurs.

d. Next assume the consumer's utility from both products declines after a certain threshold $(\bar{x}, \bar{y})$ has occurred. That is: $U(x, y) \leq U(\bar{x}, \bar{y})$ for all positive values of x and y.
Consider a poor consumer which budget, $\bar{b}$, is insufficient to realize this optimum.
Calculate by how much his income should be raised to maximize his utility.
Another way to realize this maximum is that government lowers the prices of consumer goods. Indicate how one can calculate this new price which makes the consumer can realize this optimum utility.

ANSWER 4.28

a. The maximization problem is: $max_{x,y} U(x, y)$ subject to the constraints $p_1 x + p_2 y \leq b$, $x \geq m_1$, $y \geq 0$.

b. If $b < p_1 m_1$, the constraint set is empty. So there exists no solution.

c. The constraint set is compact, so the maximum is attained either at an interior point or on the boundary of the set. Since the consumer is insatiable in at least one of the two goods - say, the basic consumption good - this implies that the condition $U_x^{'}(x, y) = 0$ has no solution. Hence, there is no interior optimum.
Moreover, because the consumer is insatiable in this good, the optimum is also not attained at the corner point $(m_1, 0)$.
Therefore, maximum utility must be achieved on the budgetline, i.e., where $p_1 x + p_2 y = b$. This reduces the problem to maximizing

$$U(x, \frac{b - p_1 x}{p_2}),\ m_1 \leq x \leq \frac{b}{p_1}.$$

The maximum is thus attained either at $(m_1, \frac{b-p_1 m_1}{p_2})$, at $(\frac{b}{p_1}, 0)$ or at a point where $U(x, \frac{b-p_1 x}{p_2})$ has a stationary point. This occurs when

$$\frac{dU}{dx} U(x, \frac{b - p_1 x}{p_2}) = U_x^{'}(x, y) + U_y^{'}(x, y) \frac{-p_1}{p_2} = 0.$$

If the corner solution $(m_1, \frac{b-p_1 m_1}{p_2})$ arises, it indicates that the consumer has a strong preference for the luxury good.

d. An increase in income of $p_1\bar{x} + p_2\bar{y} - \bar{b}$ would achieve this.
Alternatively, the government could adjust the price p_1 of the consumption good so that the new price p_1^N satisfies $U(\bar{x}, \frac{\bar{b}-p_1^N\bar{x}}{p_2}) = U(\bar{x}, \bar{y})$.

EXERCISE 4.29 *(optimal fertilization)*
Consider a farmer who wants to grow two crops a year on his land of L acres. The growth curve of his crops is approximately given by the *logistic curves* $f_i(s) = \frac{\bar{f}_i}{1+e^{-(s-10)}}$, where $\bar{f}_i$ is the maximum yield that can be realized per acre for crop $i,\ i = 1, 2$. The farmer can use a different fertilizer for each crop to increase its yield. If he uses an amount of x_i per time unit of the fertilizer, the growth of the crop increases by a fraction of $\delta_i x_i,\ i = 1, 2$. Every year he can cultivate his land for an uninterrupted period of at most T time units. All prices the farmer has to pay for his fertilizer, p_i (per unit), and the prices he gets on the market for his crop, r_i (per unit), are fixed. Assume the farmer has a maximum budget b per year he wants to spend on fertilizer and he wants to maximize his income.

a. Show that the optimization problem the farmer is confronted with can be formulated as the maximization w.r.t. the variables s_i and $x_i,\ i = 1, 2$, of the function: $\pi(s_1, s_2, x_1, x_2) := (1 + \delta_1 x_1)\frac{\bar{f}_1}{1+e^{-(s_1-10)}}Lr_1 + (1 + \delta_2 x_2)\frac{\bar{f}_2}{1+e^{-(s_2-s_1-10)}}Lr_2 - s_1p_1x_1L - (s_2 - s_1)p_2x_2L$ subject to $0 \le s_1 \le s_2 \le T$; $s_1p_1x_1L + (s_2 - s_1)p_2x_2L \le b$ and $x_i \ge 0$.

b. Let $\alpha_i := \delta_i\bar{f}_i r_i,\ i = 1, 2$. Show that if (s_1, s_2, x_1, x_2) solve the problem, there must exist (s_1, s_3, x_1, x_2) and $\lambda_i,\ i = 1, \cdots, 6$, satisfying next equations:
1.i) $\alpha_1\frac{e^{-(s_1-10)}}{(1+e^{-(s_1-10)})^2}(\frac{1}{\delta_1} + x_1) + \lambda_1 - \lambda_3 - (1 + \lambda_6)p_1x_1 = 0$;
ii) $\alpha_2\frac{e^{-(s_3-10)}}{(1+e^{-(s_3-10)})^2}(\frac{1}{\delta_2} + x_2) + \lambda_2 - \lambda_3 - (1 + \lambda_6)p_2x_2 = 0$;
iii) $((1 + \lambda_6)s_1 - \lambda_4)(1 + e^{-(s_1-10)}) = \frac{\alpha_1}{p_1}$;
iv) $((1 + \lambda_6)s_3 - \lambda_5)(1 + e^{-(s_3-10)}) = \frac{\alpha_2}{p_2}$;
2. With $g_1 := -s_1$; $g_2 := -s_3$; $g_3 := s_1 + s_3 - T$; $g_4 := -x_1$; $g_5 := -x_2$; and $g_6 := s_1p_1x_1L + s_3p_2x_2L - b$, $\lambda_i g_i = 0,\ i = 1, \cdots, 6$;
3. $\lambda_i \ge 0$ and $g_i \le 0,\ i = 1, \cdots, 6$.
Moreover, in that case $s_2 = s_3 + s_1$.

c. Show from conditions 1.iii) and iv) that the farmer will always grow both crops.

d. Show that if the farmer does not use fertilizer to increase the revenue for either one of his crops, to harvest optimally he should use the whole season T to grow his crops.

e. Show that if a point $(\bar{s}_1, \bar{s}_3, \bar{x}_1, \bar{x}_2)$ in the interior of the constrained set and $\bar{\lambda}_i,\ i = 1, \cdots, 6$, satisfy the necessary conditions in item b., the profit of the farmer is: $L(\frac{\bar{f}_1 r_1}{1+e^{-(\bar{s}_1-10)}} + \frac{\bar{f}_2 r_2}{1+e^{-(\bar{s}_3-10)}})$.

f. Show that for the farmer it is always optimal to use either the whole growth season T to grow his crops and/or to spend his complete budget for fertilization. Hint: show that if an interior solution satisfies the necessary conditions one can always find a point in the constrained set where profit is higher.

g. A numerical case study. Consider $p_1 = 1$, $p_2 = \frac{3}{2}$, $\delta_1 = 0.01$, $\delta_2 = 0.015$, $T = 30$, $b = 2.5 * 10^4$, $L = 10$, $\bar{f}_1 = 1500$, $\bar{f}_2 = 1300$, $r_1 = 1.0067$ and $r_2 = 1.0498$.
Use above items b-f and, e.g. Matlab, to verify next (approximate) results.

i) Show that $(s_1, s_3, x_1, x_2) = (15, 13, 11.16, 107.19)$ is an interior solution that satisfies the necessary conditions.

ii) Show that if the budget restriction is not binding next candidate solutions result:
$(s_1, s_3, x_1, x_2) = (17, 13, 0, 104.8)$ with $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) = (0, 0, 1.37, 1.91, 0, 0)$.
$(s_1, s_3, x_1, x_2) = (15, 15, 1.074, 0)$ with $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) = (0, 0, 9.073, 0, 1.444, 0)$.

iii) Show that if not the whole growth season is used next (approximate) candidate solution results: $(s_1, s_3, x_1, x_2) = (14.9764, 12.9527, 11.4345, 119.8507)$ with $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) = (0, 0, 0, 0, 0, 0.0014)$.

iv) Show that if as well the whole budget for fertilization is used as crops are grown for the whole growth season $(s_1, s_3, x_1, x_2) = (12.9171, 17.0829, 193.5419, 0)$ and $(s_1, s_3, x_1, x_2) = (17.084, 12.916, 0, 129.0389)$ result as candidate solutions.

v) Show that in the optimal situation the farmer should spend all his budget for fertilization on the first crop for approximately 13 time units and next use the rest of the growth season to grow his second crop.

ANSWER 4.29

a. The revenu from crop i, using an amount x_i of fertilizer per acre when the crop has grown for s time periods, is given by

$$\delta_i x_i \frac{\bar{f}_i}{1 + e^{-(s-10)}} r_i, \ i = 1, 2.$$

The cost of fertilizer per acre over this period is: $sp_i x_i$.

Letting s_i denote the time at which crop i is harvested, we then obtain the stated problem formulation.

b. Define

$$L(s_1, s_2, x_1, x_2, \tilde{\lambda}_1, \cdots, \tilde{\lambda}_5, \lambda_6) := \pi(s_1, s_2, x_1, x_2) - \sum_{i=1}^{5} \tilde{\lambda}_i g_i - \lambda_6 g_6.$$

Then, the following are the necessary (KKT) conditions for this problem.

1.i) $\alpha_1 \frac{e^{-(s_1-10)}}{(1+e^{-(s_1-10)})^2}(\frac{1}{\delta_1} + x_1) - \alpha_2 \frac{e^{-(s_2-s_1-10)}}{(1+e^{-(s_2-s_1-10)})^2}(\frac{1}{\delta_2} + x_2) + \frac{\tilde{\lambda}_1}{L} - \frac{\tilde{\lambda}_2}{L} -$
$(1 + \lambda_6)(p_1 x_1 - p_2 x_2) = 0;$

ii) $\alpha_2 \frac{e^{-(s_2-s_1-10)}}{(1+e^{-(s_2-s_1-10)})^2}(\frac{1}{\delta_2} + x_2) + \frac{\tilde{\lambda}_2}{L} - \frac{\tilde{\lambda}_3}{L} - (1 + \lambda_6) p_2 x_2 = 0;$

iii) $((1 + \lambda_6)s_1 - \frac{\tilde{\lambda}_4}{p_4 L})(1 + e^{-(s_1-10)}) = \frac{\alpha_1}{p_1};$

iv) $((1 + \lambda_6)(s_2 - s_1) - \frac{\tilde{\lambda}_5}{p_5 L})(1 + e^{-(s_2-s_1-10)}) = \frac{\alpha_2}{p_2};$

2. With $g_1 := -s_1;\ g_2 := -s_3;\ g_3 := s_2 - T;\ g_4 := -x_1;\ g_5 := -x_2;$ and
$g_6 := s_1 p_1 x_1 L + (s_2 - s_1) p_2 x_2 L - b,\ \lambda_i g_i = 0,\ i = 1, \cdots, 6;$

3. $\lambda_i \geq 0$ and $g_i \leq 0,\ i = 1, \cdots, 6.$

By introducing $s_3 := s_2 - s_1$, defining $\lambda_i := \frac{\tilde{\lambda}_i}{L}$ for $i = 1, 2, 3$, and $\lambda_i := \frac{\tilde{\lambda}_i}{p_i L}$ for $i = 4, 5$, and then adding conditions 1.i) and 1.ii) we obtain the stated conditions.

c. If $s_i = 0$ for $i = 1, 3$, we obtain a contradiction in conditions 1.iii) and 1.iv), since $\lambda_i \geq 0$ for $i = 4, 5$. Therefore, it must be that $s_1 > 0$ and $s_3 = s_2 - s_1 > 0$. From this, we conclude that the farmer will also cultivate the second crop.

d. If, for example, $x_1 = 0$, then condition 1.i) implies that $\lambda_3 > 0$. Hence, the constraint g_3 must be binding.

e. Since, by assumption, $\bar{s}_i > 0$ and $\bar{x}_i > 0$, it follows from the necessary conditions that $\bar{\lambda}_i = 0$ for $i = 1, \cdots, 6$. Hence, by conditions 1.iii) and 1.iv),

$$s_1(1 + e^{-(s_1 - 10)}) = \frac{\alpha_1}{p_1} \text{ and } s_3(1 + e^{-(s_3 - 10)}) = \frac{\alpha_2}{p_2}.$$

Substituting this into the profit $\bar{\pi} := \pi(\bar{s}_1, \bar{s}_2, \bar{x}_1, \bar{x}_2)$ yields:

$$\begin{aligned}
\bar{\pi} &= (1 + \delta_1 \bar{x}_1) \frac{\bar{f}_1}{1 + e^{-(\bar{s}_1 - 10)}} L r_1 + (1 + \delta_2) \bar{x}_2 \frac{\bar{f}_2}{1 + e^{-(\bar{s}_3 - 10)}} L r_2 - \bar{s}_1 p_1 \bar{x}_1 L - \bar{s}_3 p_2 \bar{x}_2 L \\
&= L\left(\frac{\bar{f}_1 r_1}{1 + e^{-(\bar{s}_1 - 10)}} + \frac{\bar{f}_2 r_2}{1 + e^{-(\bar{s}_3 - 10)}} + \bar{x}_1\left(\frac{\alpha_1}{1 + e^{-(\bar{s}_1 - 10)}} - \bar{s}_1 p_1\right) + \bar{x}_2\left(\frac{\alpha_2}{1 + e^{-(\bar{s}_3 - 10)}} - \bar{s}_3 p_2\right)\right) \\
&= L\left(\frac{\bar{f}_1 r_1}{1 + e^{-(\bar{s}_1 - 10)}} + \frac{\bar{f}_2 r_2}{1 + e^{-(\bar{s}_3 - 10)}}\right).
\end{aligned}$$

f. In item c, it was already shown that in the optimal situation the farmer will cultivate both crops, i.e., $\bar{s}_i > 0$.

In item d, it was shown that if no fertilizer is applied to either crop, the farmer will use the full growth period for cultivation.

Thus, the only remaining case to analyze is whether a maximum can occur in the interior of the constrained set. Suppose $(\bar{s}_1, \bar{s}_3, \bar{x}_1, \bar{x}_2)$ lies in the interior of the constrained set and $\bar{\lambda}_i = 0$ for $i = 1, \cdots, 6$, satisfying the necessary conditions from item b. From item e, the profit in that case is

$$\bar{\pi} = L\left(\frac{\bar{f}_1 r_1}{1 + e^{-(\bar{s}_1 - 10)}} + \frac{\bar{f}_2 r_2}{1 + e^{-(\bar{s}_3 - 10)}}\right).$$

Following item e, the profit at an arbitrary point is

$$L\left(\frac{\bar{f}_1 r_1}{1 + e^{-(s_1 - 10)}} + \frac{\bar{f}_2 r_2}{1 + e^{-(s_3 - 10)}} + x_1\left(\frac{\alpha_1}{1 + e^{-(s_1 - 10)}} - s_1 p_1\right) + x_2\left(\frac{\alpha_2}{1 + e^{-(s_3 - 10)}} - s_3 p_2\right)\right).$$

Now, consider $f_i(s) = \frac{\alpha_i}{1 + e^{-(s - 10)}} - s p_i$ for $i = 1, 2$. We know that $f_1(\bar{s}_1) = f_2(\bar{s}_3) = 0$. Next, choose $s_i := \bar{s}_i + \Delta s$ with Δs positive but small (so that the constraint on g_3 is still satisfied) and set $x_i = 0$. Then (s_1, s_2, x_1, x_2) belongs to the constrained set. It can be easily verified that the profit at this point is higher than $\bar{\pi}$.

Since the constrained set is compact and the profit function is continuous, Weierstrass' theorem guarantees that a maximum exists. As we have just shown, this maximum cannot occur at an

interior point of the constrained set satisfying the necessary conditions. Therefore, at the optimum, either g_3 and/or g_6 must be binding.

g. Note from item c, that in the optimal situation $\lambda_i = 0$ for $i = 1, 2$.

i) Verify (numerically) that with $\lambda_i = 0,\ i = 1, \cdots, 6$ the advertized solution satisfies approximately the necessary conditions. Note that $\pi = 27972$ in this case.

ii) If the budget constraint is not binding we have $\lambda_6 = 0$ and, by item f, $s_1 + s_3 = T$. By item d either $x_1 > 0$ or $x_2 > 0$. Elaborating these cases yields the advertized results with $\pi = 18087$ and $\pi = 28555$, respectively.

iii) If the time constraint g_3 is not binding it follows that $\lambda_3 = 0$ and, by item f, the budget constraint is binding. That is: $s_1 p_1 x_1 L + s_3 p_2 x_2 L = b$. Furthermore, it is clear from conditions 1 and 2 that always $x_i > 0,\ i = 1, 2$. So: $\lambda_4 = \lambda_5 = 0$. From condition 3 and 4 we have: $1 + \lambda_6 = \frac{\alpha_1}{s_1 p_1 (1+e^{10-s_1})} = \frac{\alpha_2}{s_3 p_2 (1+e^{10-s_3})}$. Substitution of this into conditions 1 and 2 yields then $x_1 = \frac{s_1 e^{10-s_1}}{\delta_1 (1+(1-s_1)e^{10-s_1})}$ and $x_2 = \frac{s_3 e^{10-s_3}}{\delta_2 (1+(1-s_3)e^{10-s_3})}$.
So, by numerically solving the set of equations:
1. $\frac{\alpha_1}{s_1 p_1 (1+e^{10-s_1})} = \frac{\alpha_2}{s_3 p_2 (1+e^{10-s_3})}$ and 2. $\frac{s_1^2 p_1 L e^{10-s_1}}{\delta_1 (1+(1-s_1)e^{10-s_1}))} + \frac{s_2^2 p_2 L e^{10-s_3}}{\delta_1 (1+(1-s_3)e^{10-s_3}))} = b$, we obtain the advertized solutions with $\pi = 28002$.

iv) With both g_3 and g_6 binding it is clear we cannot have simultaneously $x_i = 0,\ i = 1, 2$.
If $x_2 = 0$ it follows that $x_1 > 0$, so $\lambda_4 = 0$. After substitution of $1 + \lambda_6$ from condition 3 and λ_3 from condition 2 into condition 1, we find s_1 as the solution solving this equation. λ_5 follows then finally from condition 4. We find then the approximate solution $(12.9171, 17.0829, 193.5419, 0)$ with $\lambda_3 = 1.1436$, $\lambda_5 = 5.3097$ and $\lambda_6 = 0.1090$. This results in $\pi = 30688$.
Similarly we find with $x_1 = 0$, the approximate solution $(17.084, 12.916, 0, 129.0389)$ with $\lambda_3 = 1.2639$, $\lambda_4 = 2.0363$ and $\lambda_6 = 0.0024$. This results in $\pi = 28093$.
Finally, with $x_i > 0,\ i = 1, 2,\ \lambda_i = 0,\ i = 4, 5$. By substitution of λ_6 from condition 3 into condition 4 and using g_3, we obtain an equation in s_1. After solving this equation we can then determine next from conditions 1, 2 and the budget constraint x_1, x_2 and λ_3. We find approximately $s_1 = 15.8542$ and $s_1 = 9.734$. However in both cases this results in $\lambda_6 < 0$. So, we don't get any additional candidate solutions.

v) Since the constraint set is compact and the profit function continuous we know from Weierstrass' theorem that the problem has a solution. Evaluation of profit at the candidate solutions shows that maximum profit occurs at $(s_1, s_3, x_1, x_2) = (12.9171, 17.0829, 193.5419, 0)$.

EXERCISE 4.30 *(green farming)*

Consider a farmer who grows an amount x of a product using traditional farming and an amount y using a more environmentally friendly way. He can grow at most a total amount c of the product. The price he can get on the market for his product using traditional farming is $p_1(x, y)$ and using the environmentally friendly way is $p_2(x, y)$, where $p_1(x, y) \leq p_2(x, y)$. The production cost are $C_1(x)$ and $C_2(y)$, respectively, where $C_1(z) \leq C_2(z)$. All functions are assumed to be continuous.

a. Assume the farmer wants to maximize his profit. Formulate the optimization problem and show the problem has a solution.
b. Formulate the KKT conditions.
c. Consider the case $p_i(x,y) = \alpha_i - (x+y)$ and $C_i(z) = \beta_i z + \gamma_i,\ i = 1,2$, where the constants α_i, β_i and γ_i are all nonnegative and satisfy $\alpha_1 \leq \alpha_2$, $\beta_1 \leq \beta_2$, $\gamma_1 \leq \gamma_2$ and $\beta_1 - \alpha_1 \neq \beta_2 - \alpha_2$.
Show that the problem has no interior solution.
d. Show that under the assumption that $(\frac{\alpha_i - \beta_i}{2})^2 \neq \gamma_1 - \gamma_2,\ i = 1,2$, the constraint qualification is met.
e. Show from the KKT conditions that, under the parametric assumptions stated in items c and d, if the farmer decides to produce he will always specialize in the production of either one of both products. Show he will specialize in the environmentally friendly way iff. $\alpha_2 - \alpha_1 > \beta_2 - \beta_1$.
f. Assume the farmer grows the product using traditional farming and the government wants the farmer to grow it in the environmentally way. Indicate how the government might achieve this goal.

ANSWER 4.30

a. The farmer wants to maximize profit $\pi(x,y) := xp_1(x,y) + yp_2(x,y) - C_1(x) - C_2(y)$, subject to the constraints $x \geq 0$, $y \geq 0$, $x + y \leq c$ and $\pi(x,y) \geq 0$.
Note that the profit function $\pi(x,y)$ is continuous. Furthermore, the constraint set is compact. So, by Weierstrass' theorem the problem has a solution.
b. Consider $L(x,y,\lambda_1,\lambda_2,\lambda_3,\lambda_4) := (1+\lambda_4)\pi(x,y) - \lambda_1(x+y-c) + \lambda_2 x + \lambda_3 y$.
Then the KKT conditions are:
i) $(1+\lambda_4)(p_1 + xp_{1x}^{'} + yp_{2x}^{'} - C_1^{'}) - \lambda_1 + \lambda_2 = 0$;
ii) $(1+\lambda_4)(xp_{1y}^{'} + p_2 + yp_{2y}^{'} - C_2^{'}) - \lambda_1 + \lambda_3 = 0$;
iii) $\lambda_1(c - x - y) = 0$; $\lambda_2 x = 0$; $\lambda_3 y = 0$; $\lambda_4 \pi = 0$;
iv) $x + y \leq c$; $x \geq 0$; $y \geq 0$; $\pi \geq 0$ and $\lambda_i \geq 0,\ i = 1, \cdots, 4$.
c. Note that $\pi_x^{'} = -2(x+y) + \alpha_1 - \beta_1$ and $\pi_y^{'} = -2(x+y) + \alpha_2 - \beta_2$. So, $\pi^{'}(x,y) = [0\ 0]$ has no solutions.
d. $Dg = \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \\ -\pi_x^{'} & -\pi_y^{'} \end{bmatrix}$. From this it is directly clear that if one constraint is binding, the constraint qualification is met. In case two constraints are simulataneously binding the constraint qualification is not met if, e.g., $x = 0$ and $\pi = 0$ together with $\pi_y^{'} = 0$ would apply. However, elaborating this case shows that due to the parametric assumption $(\frac{\alpha_2 - \beta_2}{2})^2 \neq \gamma_1 - \gamma_2$, this cannot occur. Similar one shows that the case $y = 0$, $\pi = 0$ together $\pi_x^{'} = 0$ has no solution either.
e. By items c and d we have that the solution is attained at the boundary of the feasible solution set. The KKT conditions are:
i) $(1+\lambda_4)(\alpha_1 - \beta_1 - 2x - 2y) - \lambda_1 + \lambda_2 = 0$;

ii) $(1+\lambda_4)(\alpha_2-\beta_2-2x-2y)-\lambda_1+\lambda_3=0$;
iii) $\lambda_1(c-x-y)=0$; $\lambda_2 x=0$; $\lambda_3 y=0$; $\lambda_4\pi=0$;
iv) $x+y\leq c$; $x\geq 0$; $y\geq 0$; $\pi\geq 0$ and $\lambda_i\geq 0,\ i=1,\cdots,4$.
From items i) and ii), we have

$$2(1+\lambda_4)(x+y)=(\alpha_1-\beta_1)(1+\lambda_4)-\lambda_1+\lambda_2=(\alpha_2-\beta_2)(1+\lambda_4)-\lambda_1+\lambda_3.$$

From this, it follows that:

$$(\alpha_1-\beta_1)(1+\lambda_4)+\lambda_2=(\alpha_2-\beta_2)(1+\lambda_4)+\lambda_3.$$

Or, equivalently,

$$\lambda_2=\lambda_3+(\alpha_2-\alpha_1+\beta_1-\beta_2)(1+\lambda_4).$$

Since $\alpha_2-\alpha_1+\beta_1-\beta_2\neq 0$ and $\lambda_4\geq 0$, we conclude that $\lambda_2\neq\lambda_3$.
Therefore, either $x=0$ and $y\neq 0$, $x\neq 0$ and $y=0$ or $x=y=0$. Note that the last case corresponds to the farmer producing nothing.
Furthermore, if production occurs, $\lambda_2>0$ (which implies $x=0$) iff. $\alpha_2-\alpha_1+\beta_1-\beta_2>0$, which confirms the claim.

f. For example, the government could increase the difference between α_2 and α_1 by providing a subsidy on the price of the environmentally produced product and/or by taxing the traditionally produced product.
Additionally, the government could tax the inputs required for traditional farming and/or subsidize those needed for environmentally friendly production, thereby reducing the gap between β_2 and β_1.

CHAPTER 5

Constrained Optimization: The Convex and Linear Case

5.1 Convex sets and their relationship with convex functions

EXERCISE 5.1 *(Proof Lemma 5.2)*
Let $S, T \subset \mathbb{R}^n$ be convex. Show that $S \cap T$ is either empty or convex.

ANSWER 5.1
If the intersection is not empty, we may assume that it contains at least two points, s and t (otherwise the intersection is trivially a convex set). Since s and t belong to S we know that the line connecting them belongs also to S. Similarly, since s and t belong to T we know that the line connecting them belongs to T too. That is, the line connecting s and t belongs to both S and T. So to the intersection of both sets.

EXERCISE 5.2 *(Proof Lemma)*
Let $S \subset \mathbb{R}^n$ be convex and $g : S \to \mathbb{R}$ be a convex function and $b \in \mathbb{R}$. Consider $G := S \cap \{x \in \mathbb{R}^n \mid g(x) \leq b\}$. Show that G is either empty or convex.

ANSWER 5.2
Assume $G \neq \emptyset$. Let $x_1, x_2 \in G$. Then $g(x_i) \leq b,\ i = 1, 2$. But then also

$$g(\alpha x_1 + (1-\alpha)x_2) \leq \alpha g(x_1) + (1-\alpha)g(x_2) \leq \alpha b + (1-\alpha)b = b.$$

That is, for arbitrary $\alpha \in [0, 1]$, $\alpha x_1 + (1-\alpha)x_2 \in G$. That is, G is convex.

5.2 Convex Programming and Saddle-Point Equilibria

EXERCISE 5.3 Consider the problem to minimize $f(x,y) = (x+1)^2+(y+1)^2$, subject to the constraints $(x-1)^2 \leq 4$, and $y \geq 0$.

a. Is the constraint set compact?

b. Solve the problem using Corollary 5.11.

c. Show that the Lagrangian $L(x,y,\lambda_1,\lambda_2) := -(x+1)^2 - (y+1)^2 - \lambda_1((x-1)^2-4) + \lambda_2 y$ has a saddle-point equilibrium at $((x,y),(\lambda_1,\lambda_2)) = ((-1,0),(0,2))$.

ANSWER 5.3

a. Note that $(0,a)$, with a arbitrarily large, belongs to the constraint set. So this set is not compact.

b. Let $f(x,y) := -(x+1)^2 - (y+1)^2$; $g_1(x,y) := (x-1)^2 - 4$ and $g_2(x,y) := -y$. Then

$$Df = [-2(x+1),\ -2(y+1)] \text{ and } Hf = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Hf is negative semi-definite on its domain. So, f is a concave function. Furthermore,

$$Hg_1 = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } Hg_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

which both are positive semi-definite. So, both g_1 and g_2 are convex functions.

Finally, for example at $(0,1)$, we have $g_1(0,1) < 0$ and $g_2(0,1) < 0$. Hence, Slater's condition is satisfied. Therefore, by Corollary 5.11, all solutions to the problem are given by those that satisfy the Kuhn-Tucker conditions (1-3).

So, consider the with this problem associated Lagrange function:

$$L(x,y,\lambda_1,\lambda_2) := -(x+1)^2 - (y+1)^2 - \lambda_1((x-1)^2-4) + \lambda_2 y.$$

The with this problem associated Kuhn-Tucker conditions (1-3) are:

1. $[D_1L(x,y,\lambda)\ D_2L(x,y,\lambda)] = [-2(x+1) - 2\lambda_1(x-1),\ -2(y+1)+\lambda_2] = [0\ 0]$;
2. $i)$ $\lambda_1((x-1)^2-4) = 0$ and $ii)$ $\lambda_2 y = 0$;
3. $i)$ $(x-1)^2 \leq 4$; $ii)$ $y \geq 0$; and $\lambda_i \geq 0,\ i = 1,2$.

From $\lambda_2 = 2(y+1) \geq 2$, it follows from item 2.ii) that $y = 0$. Consequently, $\lambda_2 = 2$.

Next, suppose that g_1 is not binding, so that $\lambda_1 = 0$. Then, by item 1., we obtain $x = -1$, which contradicts the assumption that g_1 is not binding.

Hence, g_1 must be binding. This implies that either $x = -1$ or $x = 4$. From item 1., we find that in the first case $\lambda_1 = 0$, while in the second case $\lambda_1 = -2$.

Therefore, the only solution to the problem is $(x,y) = (-1,0)$.

c. Let $L(x, y, \lambda_1, \lambda_2) = -(x+1)^2 - (y+1)^2 - \lambda_1((x-1)^2 - 4) + \lambda_2 y$. Then

$$L(x, y, 0, 2) = -(x+1)^2 - (y+1)^2 + 2y = -(x+1)^2 - y^2 - 1 \leq -1 = L(-1, 0, 0, 2).$$

On the other hand, $L(-1, 0, \lambda_1, \lambda_2) = -1 \geq L(-1, 0, 0, 2)$.

EXERCISE 5.4 Use Theorem 5.13 to solve the next constrained optimization problems.

a. Maximize $U(x, y) = x + \ln(1 + y)$, subject to the constraints $x + 2y \geq 4$, $x \geq 0$ and $y \geq 0$.

b. Minimize $\tilde{U}(x, y) = x - \ln(1 + y)$, subject to the constraints $x + 2y \geq 4$, $x \geq 0$ and $0 \leq y \leq 10$.

c. Show that the Lagrangian $L(x, y, \lambda) := -x + \ln(1 + y) - \lambda_1(4 - x - 2y) + \lambda_2 x + \lambda_3 y - \lambda_4(y - 10)$, associated with the problem considered in item b., has a saddle-point at $((x, y), (\lambda_1, \lambda_2, \lambda_3, \lambda_4)) = ((0, 10), (0, 1, 0, \frac{1}{11}))$.

ANSWER 5.4

a. Note that

$$DU = [1 \ \frac{1}{1+y}] \text{ and } HU = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{1}{(1+y)^2} \end{bmatrix}.$$

Clearly, the Hessian HU is negative semi-definite on its domain, and thus U is a concave function. Moreover, all constraints are linear. Therefore, by Theorem 5.13, the problem has a solution iff. the Kuhn-Tucker conditions (1-3) are satisfied, or equivalently, if the corresponding Lagrange function has a saddle-point.

Therefore, consider

$$L(x, y, \lambda) := x + \ln(1 + y) - \lambda_1(4 - x - 2y) + \lambda_2 x + \lambda_3 y.$$

The problem has a solution only if the next system of equations admits a solution on the domain $\lambda_i \geq 0$:

$$DL = [1 + \lambda_1 + \lambda_2 \ \frac{1}{1+y} + 2\lambda_1 + \lambda_3] = [0 \ 0].$$

However, this system has no solution. Therefore, the problem has no solution.

b. To minimize $\tilde{U}$ we consider the maximization of $U(x, y) := -\tilde{U}(x, y)$. Note that

$$DU = [-1 \ \frac{1}{1+y}] \text{ and } HU = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{1}{(1+y)^2} \end{bmatrix}.$$

Clearly, the Hessian HU is negative semi-definite on its domain, and thus U is a concave function. Moreover, all constraints are linear. Therefore, by Theorem 5.13, the problem has a solution iff. the Kuhn-Tucker conditions (1-3) are satisfied.

Therefore, consider

$$L(x, y, \lambda) := -x + \ln(1 + y) - \lambda_1(4 - x - 2y) + \lambda_2 x + \lambda_3 y - \lambda_4(y - 10).$$

Then, the problem has a solution iff. the next system of equations admits a solution:

1. $DL = [-1 + \lambda_1 + \lambda_2 \frac{1}{1+y} + 2\lambda_1 + \lambda_3 - \lambda_4] = [0\ 0]$;
2. i) $\lambda_1(4 - x - 2y) = 0;\ ii)\ \lambda_2 x = 0;\ iii)\ \lambda_3 y = 0;\ iv)\ \lambda_4(y - 10) = 0$;
3. i) $4 - x - 2y \leq 0;\ ii)\ x \geq 0;\ iii)\ y \geq 0;\ iv)\ y \leq 10;\ v)\ \lambda_i \geq 0$.

Case 1: Assume $\lambda_1 = 0$:

From $DL_x = 0$, it follows that $\lambda_2 = 1$. Then, by 2.ii), we obtain $x = 0$. From $DL_y = 0$, we have $\lambda_4 > 0$, which implies $y = 10$. Moreover, by 3.i), $y \geq 2$, so $\lambda_3 = 0$. Therefore, $(x, y, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = (0, 10, 0, 1, 0, \frac{1}{11})$ is a solution to the problem.

Case 2: Assume $\lambda_1 > 0$:

Then $x + 2y = 4$ and, since $\lambda_4 > 0$, $y = 10$. Which yields $x = -16$, which violates 3.ii).

c. $L(0, 10, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \ln(11) + 16\lambda_1 + 10\lambda_3 \geq \ln(11)$ on its domain.
Furthermore, $L(x, y, 0, 1, 0, \frac{1}{11}) = \ln(1 + y) - \frac{1}{11}(y - 10)$. Then, $L'(x, y, 0, 1, 0) = [0\ 0]$ iff. $y = 10$. Furthermore, $L(x, y, 0, 1, 0)$ is concave (consider HL and notice this matrix is negative semi-definite). So, $L(x, y, 0, 1, 0)$ has a maximum at $(0, 10)$.

5.3 Linear Programming and its Dual Problem

EXERCISE 5.5 Rewrite next linear programming problems into the standard form and next solve them.

a. Maximize $f(x, y) = x+3y$, subject to the constraints $x \leq 4$, $y \leq 4$, $2x+y \leq 9$, $x+2y \leq 9$, $x+y \leq 6$, $x \geq 0$ and $y \geq 0$.

b. Minimize $f(x, y) = 6x + y$, subject to the constraints $x + y \geq 1$, $x + 100y \geq 100$, $100x + y \geq 100$, $x \geq 0$ and $y \geq 0$.

c. Maximize $U(x, y) = x + y$, subject to the constraints $x + 3y \leq 10$, $3x + y \leq 10$, $2x + 2y \leq 8$ $x \geq 0$ and $y \geq 0$. Solve the problem first by using Theorem 5.14. Next, solve the problem using Corollary 5.15.

d. Maximize $f(x, y) = x - y$ subject to the constraints $x + 3y \leq 40$, $3x + y \leq 16$, $x + y \leq 15$ $x \geq 1$ and $y \geq 10$.

ANSWER 5.5

a. Define $L(x, \lambda, \mu) = c^T x - \lambda^T(Ax - b) + \mu^T x$, with $c^T = [1\ 3]$; $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}$ and $b^T = [4\ 4\ 9\ 9\ 6]$.

The cornerpoints of the constraint set are: $(0, 0)$, $(0, 4)$, $(4, 0)$, $(4, 1)$, $(1, 4)$ and $(3, 3)$. The values

of the objective function f at these points are: 0, 12, 4, 7, 13 and 12, respectively. Note that the constraint set is compact. Therefore the maximum is attained at $(x, y) = (1, 4)$.

b. After reformulating the problem into a maximization problem subject to " $\leq$ " inequality constraints, corresponding Lagrange function $L(x, \lambda, \mu)$ is:

$$Lc^T x - \lambda^T(Ax - b) + \mu^T x, \text{ with } c^T = [-6\ -1];\ A = \begin{bmatrix} -1 & -1 \\ -1 & -100 \\ -100 & -1 \end{bmatrix};\ b^T = [-4\ -100\ -100].$$

The cornerpoints of the constraint set are: $(100, 0)$, $(0, 100)$, $(\frac{96}{99}, 3\frac{3}{99})$ and $(3\frac{3}{99}, \frac{96}{99})$.
The values of the objective function f at these points are: $600,\ 100,\ 6 * \frac{96}{99} + 3\frac{3}{99}$ and $\frac{96}{99} + 6 * 3\frac{3}{99}$, respectively. Thus, the minimum value is attained at $(x, y) = (0, 100)$. Note that the constraint set is not compact. Therefore, we verify whether the three conditions mentioned in Theorem 5.14 are satisfied at this point. Observe that the constraints $x = 0$ and $x + 100y \geq 10$ are binding at this point, while the other constraints are not. Hence, $\mu_2 = \lambda_1 = \lambda_3 = 0$. Using this, it follows readily that, with $\lambda_2 = 0.01$ and $\mu_1 = 5.99$, next three conditions are met.
i. $c^T - \lambda^T A + \mu^T = 0$;
ii. $\lambda_i(a_i^T x - b_i) = 0,\ i = 1, 2, 3$, and $\mu_i x_i = 0,\ i = 1, 2$;
iii. $\lambda \geq 0$; $\mu \geq 0$; $Ax \leq b$ and $x \geq 0$.
Therefore, the problem is solved at $(x, y) = (0, 100)$.

c. Consider

$$L(x, \lambda, \mu) := c^T x - \lambda^T(Ax - b) + \mu^T x, \text{ with } c^T = [1\ 1];\ A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \\ 2 & 2 \end{bmatrix};\ b^T = [10\ 10\ 8].$$

By Theorem 5.14, this LP problem has a maximum at x if and only if there exist $\lambda,\ \mu$ and x satisfying next three conditions:
i. $c^T - \lambda^T A + \mu^T = 0$;
ii. $\lambda_i(a_i^T x - b_i) = 0,\ i = 1, 2, 3$, and $\mu_i x_i = 0,\ i = 1, 2$;
iii. $\lambda \geq 0$; $\mu \geq 0$; $Ax \leq b$ and $x \geq 0$.
That is, iff next set of (in)equalities has a solution.
i. a) $1 - \lambda_1 - 3\lambda_2 - 2\lambda_3 + \mu_1 = 0$ and b) $1 - 3\lambda_1 - \lambda_2 - 2\lambda_3 + \mu_2 = 0$;
ii. a) $\lambda_1(x + 3y - 10) = 0$, b) $\lambda_2(3x + y - 10) = 0$, c) $\lambda_3(2x + 2y - 8) = 0$ and $\mu_i x_i = 0,\ i = 1, 2$;
iii. $\lambda_i \geq 0,\ i = 1, 2, 3$, $\mu_i \geq 0,\ i = 1, 2$, $x + 3y \leq 10$, $3x + y \leq 10$, $2x + 2y \leq 8$ and $x_i \geq 0,\ i = 1, 2$.
To solve this set of (in)equalities we proceed as follows:
1. precisely one constraint is binding.
   1. $x = 0$, $\lambda_1 = \lambda_2 = \lambda_3 = \mu_2 = 0$. Substitution of this into item i. yields a contradiction.
   2. $y = 0$, $\lambda_1 = \lambda_2 = \lambda_3 = \mu_1 = 0$. Substitution of this into item i. yields a contradiction.
   3. $x + 3y = 10$, $\lambda_2 = \lambda_3 = \mu_1 = \mu_2 = 0$. Substitution of this into item i. yields a contradiction.

4. $3x + y = 10$, $\lambda_1 = \lambda_3 = \mu_1 = \mu_2 = 0$. Substitution of this into item i. yields a contradiction.
5. $2x + 2y = 8$, $\lambda_1 = \lambda_2 = \mu_1 = \mu_2 = 0$. Substitution of this into item i. yields $\lambda_3 = \frac{1}{2}$. Substitution of $y = 4 - x$ into the inequalities $x + 3y < 10$, $3x + y < 10$ shows that all points $(x, 4 - x, 0, 0, \frac{1}{2}, 0, 0)$ with $1 < x < 3$ satisfy items i-iii.

2. precisely two constraints are binding. Here we obtain the additional two solutions $(x, y) = (1, 3)$ and $(x, y) = (3, 1)$.

Hence, all points $(x, 4 - x)$ with $1 \leq x \leq 3$ are solutions to the problem.

Another way to solve this problem is to evaluate the objective function U at the corner points: $(0, 0)$, $(0, \frac{10}{3})$, $(\frac{10}{3}, 0)$, $(3, 1)$ and $(1, 3)$. The corresponding objective values are $0,\ \frac{10}{3},\ \frac{10}{3},\ 4$ and 4, respectively.

Since the constraint set is compact, Weierstrass's theorem guarantees that a solution exists. Notice that the maximum value is attained at two corner points simultaneously. Therefore, by Corollary 5.15, the maximum is also attained at all points on the line segment connecting these two cornerpoints.

d. Let

$$g_1(x, y) := x + 3y,\ g_2(x, y) := 3x + y \text{ and } g_3(x, y) := x + y.$$

Next, consider the linear transformation of the variables: $z_1 := x - 1$ and $z_2 := y - 10$. Then, with $x := z_1 + 1$ and $y := z_2 + 10$, the LP problem in (x, y) coordinates is equivalent to the following LP problem in (z_1, z_2) coordinates.

$$\text{Maximize } f(z_1, z_2) := z_1 - z_2 - 9$$

subject to

$$g_1(z_1, z_2) := z_1 + 3z_2 \leq 9, g_2(z_1, z_2) := 3z_1 + z_2 \leq 3,\ g_3(z_1, z_2) := z_1 + z_2 \leq 4, z_i \geq 0,\ i = 1, 2.$$

The feasible set is described by $g_2(z_1, z_2) \leq 4$ and $z_i \geq 0$ for $i = 1, 2$. This set is compact. The corner points of this feasible set are: $(0, 0),\ (0, 3)$ and $(1, 0)$. The corresponding objective values are: $-9,\ -12$ and -8, respectively. Hence, the maximum is attained at $(z_1, z_2) = (1, 0)$. Therefore, the LP problem has a unique solution at $(x, y) = (2, 10)$.

EXERCISE 5.6 Formulate the corresponding dual linear programming problems of next LP problems. Next solve both these dual and LP problems.

a. Maximize $U(x, y) = x + y$, subject to the constraints $x + 3y \leq 10$, $3x + y \leq 10$, $2x + 2y \leq 8$ $x \geq 0$ and $y \geq 0$.
b. Maximize $f(x, y, z) = x - y + z$, subject to the constraints $x + y + z \leq 1$, $x \geq 0$, $y \geq 0$ and $z \geq 0$.
c. Minimize $f(x, y, z) = x + y + 2z$, subject to the constraints $x + y \geq 1$, $y + z \geq 2$, $x + z \geq 3$ and $x + y + 2z \geq 3$.

ANSWER 5.6

a. Using the notation of the standard LP formulation, we have

$$c^T = [1\ 1],\ A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \\ 2 & 2 \end{bmatrix} \text{ and } b^T = [10\ 10\ 8].$$

Thus, the dual LP problem is to minimize $b^T\lambda$, subject to

$$A^T\lambda \geq c \text{ and } \lambda \geq 0.$$

The corner points of the feasible set are found by setting three of the five inequalities to equality and then checking whether all constraints are satisfied. This procedure yields the following candidate corner points $(\lambda_1, \lambda_2, \lambda_3)$:

$$(0,0,0),\ (0,0,\frac{1}{2}),\ (0,1,0),\ (0,\frac{1}{3},0),\ (1,0,0) \text{ and } (\frac{1}{3},0,0).$$

However, checking feasibility shows that $(0,0,0)$, $(0,\frac{1}{3},0)$ and $(\frac{1}{3},0,0)$ do not satisfy all constraints. Hence, the actual corner points are

$$(0,0,\frac{1}{2}),\ (0,1,0) \text{ and } (1,0,0).$$

It is straightforward to verify that the minimum value of the dual objective function is attained at $(0,0,\frac{1}{2})$. Moreover, larger values are obtained in the interior of the feasible set. Therefore, the dual problem has a unique solution at $(\lambda_1, \lambda_2, \lambda_3) = (0,0,\frac{1}{2})$.

Substituting this result into the KKT conditions for the original LP problem shows that the maximum of the primal problem is attained at all points $(x, 4-x)$ with $x \in [2,3]$.

b. Using the notation of the standard LP formulation, we have

$$c^T = [1\ -1\ 1],\ A = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}, \text{ and } b = [1].$$

Thus, the dual LP problem is to minimize $b^T\lambda$, subject to

$$A^T\lambda \geq c, \text{ and } \lambda \geq 0.$$

The corner points of the feasible set are $(0,0,0)$, $(1,0,0)$ and $(0,1,0)$ and $(0,0,1)$. Note that the constraint set is compact. Evaluating the objective function at these corner points shows that the LP problem attains a maximum value of 1 at all points on the line segment connecting $(1,0,0)$ and $(0,0,1)$.

Clearly, the minimum of the dual LP problem is attained at $\bar{\lambda} = 1$.

Observe that $b\bar{\lambda} = 1 = c^T\bar{x}$ for any point $\bar{x}$ on the line segment between $(1,0,0)$ and $(0,0,1)$.

c. First, we rewrite the problem in standard form. Let

$$x_1 := 1 - x - y,\ x_2 := 2 - y - z \text{ and } x_3 := 3 - x - z.$$

Then

$$x = \frac{1}{2}(-x_1 + x_2 - x_3 + 2),\ y = \frac{1}{2}(-x_1 - x_2 + x_3) \text{ and } z = \frac{1}{2}(x_1 - x_2 - x_3 + 4).$$

The objective function can then be rewritten as

$$\tilde{f}(x_1, x_2, x_3) := \frac{1}{2}(x_1 + x_2 + x_3 - 6)$$

to be maximized subject to the constraints

$$x_i \geq 0 \text{ for } i = 1, 2, 3, \text{ and } g_1(x_1, x_2, x_3) := x_2 + x_3 \leq 2.$$

Note that the maximizer of this problem coincides with that of

$$\hat{f}(x_1, x_2, x_3) := \frac{1}{2}(x_1 + x_2 + x_3)$$

over the same feasible set. Hence, the standard LP problem is to maximize

$$c^T \bar{x}, \text{ where } c^T = \frac{1}{2}[1\ 1\ 1] \text{ and } \bar{x}^T = [x_1\ x_2\ x_3],$$

subject to

$$A\bar{x} \leq 2, \text{ with } A = [0\ 1\ 1] \text{ and } b = [2].$$

The corresponding dual LP problem is to minimize

$$b\lambda, \text{ subject to } A^T\lambda \geq c \text{ and } \lambda \geq 0.$$

Clearly, this dual problem has no solution. Therefore, the original LP problem also has no solution.

Mixed Exercises

EXERCISE 5.7 Let $b \in \mathbb{R}$. Use Lemma 5.4 to show that next sets G are convex.

a. $G := \{x \in \mathbb{R}^n \mid a^T x \leq b\}$, where $a \in \mathbb{R}^n,\ a \neq 0$.

b. $G := \{(x, y) \in \mathbb{R}^2_{++} \mid \frac{x^2}{y} \leq b\}$, where $b > 0$.

c. $G := \{(x, y) \in \mathbb{R}^2 \mid \log(e^x + e^y) \leq b\}$.

d. $G := \{(x, y, z) \in \mathbb{R}^3_{+++} \mid \frac{1}{x^2} \leq 1,\ y^3 \leq 9 \text{ and } \sqrt{z} \geq 4\}$.

ANSWER 5.7

a. Consider $g(x) := a^T x$. $g(x)$ is a convex function. So, by Lemma 5.4, the level set G is a convex set.

b. Consider $g(x,y) := \frac{x^2}{y}$ for $x > 0$ and $y > 0$. Note that the positive quadrant C is a convex set. Indeed, if $x, y \in C$, then for any $\alpha \in (0,1)$, the point $z := \alpha x + (1-\alpha)y > 0$, so $z \in C$ as well.
Next, observe that

$$g' = [\frac{2x}{y} \quad -\frac{x^2}{y^2}] \text{ and } Hg = \begin{bmatrix} \frac{2}{x} & -\frac{2x}{y^2} \\ -\frac{2x}{y^2} & \frac{2x^2}{y^3} \end{bmatrix}.$$

Thus, the Hessian Hg is positive semi-definite on C, which implies that $g(x,y)$ is convex on C. Consequently, by Lemma 5.4, the level set G is a convex set.

c. Consider $g(x,y) := \log(e^x + e^y)$. Then

$$g' = [\frac{e^x}{e^x + e^y} \quad \frac{e^y}{e^x + e^y}] \text{ and } Hg = \begin{bmatrix} \frac{e^x e^y}{(e^x+e^y)^2} & -\frac{e^x e^y}{(e^x+e^y)^2} \\ \frac{-e^x e^y}{(e^x+e^y)^2} & \frac{e^x e^y}{(e^x+e^y)^2} \end{bmatrix}.$$

Thus, the Hessian Hg is positive semi-definite, which implies that $g(x,y)$ is convex. Consequently, by Lemma 5.4, the level set G is a convex set.

d. Consider

$$g_1(x,y,z) := \frac{1}{x^2},\ g_2(x,y,z) := y^3 \text{ and } g_3(x,y,z) := -\sqrt{z} \text{ for } x > 0,\ y > 0 \text{ and } z > 0.$$

As in item b, the set where all coordinates are positive is convex.
Moreover, straightforward differentiation shows that the Hessians Hg_i, for $i = 1,2,3$, are positive semi-definite on this set. Hence, each $g_i,\ i = 1,2,3$, is convex on this domain.
Therefore, by Lemma 5.4, all corresponding level sets are convex. Since their intersection is clearly nonempty, it follows from Proposition 5.5 that the set G is convex.

EXERCISE 5.8 Solve next constrained optimization problems.

a. Maximize $f(x,y,z) = x + y + z$ s.t. $x + y \leq 1$, $y + z \leq 2$, $x + z \leq 3$, $x \geq 0$, $y \geq 0$ and $z \geq 0$.

b. Minimize $f(x,y) = e^{-(x+y)}$ s.t. $x \leq 4 + y$.

c. Maximize $f(x,y,z) = x + y - z$ s.t. $x + y + z \geq 1$, $x + 4y \leq 5$, $x \geq 0$, $y \geq 0$ and $z \geq 0$.

d. Minimize $f(x,y,z) = e^x + y^2$, s.t. $x^2 + y^2 + z^2 \leq 1$ and $x + y + z^2 \leq 3$.

e. Maximize $f(x,y) = -(x-1)^2 - (y-1)^2$ on $C := \{(x,y) \mid x > 0,\ y > 0\}$ s.t. $x + y \leq 1$.

f. Reconsider the maximization problem in item c, but now on the set $C := \{(x,y) \mid x < 0,\ y < 0\}$.

g. Minimize $f(x,y) = (x-1)^2 + (y-1)^2$ s.t. $x + y \geq 3$ and $x^2 + y^2 \leq 9$.

h. Maximize $f(x,y) = y - x$ s.t. $x + 2y \geq 5$, $2x + y \geq 4$ and $y + x \leq 4$.

i. Minimize $f(x,y,z) = 2x + y + z$ s.t. $\ln(x + y + z) \geq 1$ and $x + y + z \geq 0$.

j. Maximize $f(x_1,x_2,x_3,x_4,x_5) = x_1 + 2x_2 - 3x_3 + 4x_4 - 5x_5$ s.t. $5x_1 - 4x_2 + 3x_3 - 2x_4 - x_5 \leq 10$, $x_1 + x_2 + x_3 + x_4 + x_5 \leq 2$ and $x_i \geq 0$.

ANSWER 5.8

a. Note that the constraint set is compact. Since f is continuous, the problem therefore has a solution. According to Algorithm 5.16, we can find this solution by evaluating f at the corner points. These corner points are obtained by considering the intersection of any three binding constraints. For example, setting $x = 0$, $y = 0$ and $z = 0$ yields the corner point $(0, 0, 0)$. Setting $x = 0$, $y = 0$ and $x + z = 2$ yields $(0, 0, 2)$, and so on.
Proceeding in this way yields the following six corner points:

$$(0,0,0),\ (0,0,2),\ (0,1,0),\ (1,0,0),\ (0,1,1) \text{ and } (1,0,2).$$

It is straightforward to verify that f attains its maximum value at $(1, 0, 2)$.

b. The problem is equivalent to maximizing

$$f(x,y) = -e^{-(x+y)} \text{ subject to } x - y \leq 4.$$

It is straightforward to verify that the Hessian Hf is negative semi-definite, so f is concave. Moreover, $g(x, y) = x - y$ is convex, and the constraint set has an interior point. Therefore, by Corollary 5.11, the problem has a solution iff. the KKT conditions are satisfied.
With $L(x, y, \lambda) := -e^{-(x+y)} - \lambda(x - y - 4)$, we obtain $DL = [e^{-(x+y)} - \lambda,\ e^{-(x+y)} + \lambda]$. Thus, the system $DL = [0\ 0]$ has no solution. Consequently, the problem has no solution.

c. Note that the constraint set is not compact, since $z \geq 0$ can take arbitrarily large values. According to Algorithm 5.16, we find the solution - if it exists - by evaluating f at the corner points.
These corner points are obtained by considering the intersection point of any three binding constraints. For example, setting $x = 0$, $y = 0$ and $z = 0$ yields $(0, 0, 0)$ while setting $x = 0$, $y = 0$ and $x + y + z = 1$ yields $(0, 0, 1)$, and so on. This procedure results in the following four 4 corner points:

$$(0,0,0),\ (0,0,1),\ (0,1,0) \text{ and } (1,0,0).$$

It is clear that f attains a maximum value of 1 at $(0, 1, 0)$ and $(1, 0, 0)$. Hence, all points on the line segment connecting these two points are candidate solutions.
To verify whether these points are indeed optimal, we check the conditions of Theorem 5.14. With

$$c^T := [1\ 1\ -1],\ A := \begin{bmatrix} -1 & -1 & -1 \\ 1 & 4 & 0 \end{bmatrix} \text{ and } b^T := [-1\ 5],$$

we must verify whether the following two conditions hold at any point $(\theta, 1 - \theta, 0),\ 0 \leq \theta \leq 1$ on the line segment: there exist $\lambda \geq 0$ and $\mu \geq 0$ such that

i) $c^T - \lambda^T A + \mu^T = 0$ and ii) $\lambda_1(x + y + z - 1) = 0,\ \lambda_2(x + 4y) = 0$ and $\mu_1 x = \mu_2 y = \mu_3 z = 0$.

At any point on the line segment, the constraint $x + y + z \geq 1$ is binding, while the constraint $x + 4y \leq 5$ is not. Hence, from ii), $\lambda_2 = \mu_1 = \mu_2 = 0$ for $\theta \in (0, 1)$.
Substiting into i) gives:

$$[1\ 1\ -1] - [\lambda_1\ 0] \begin{bmatrix} -1 & -1 & -1 \\ 1 & 4 & 0 \end{bmatrix} + [0\ 0\ \mu_3] = [0\ 0\ 0].$$

It is easily verified that this condition holds for $\lambda_1 = 1$ and $\mu_3 = 0$. A similar verification shows that conditions i) and ii) also hold at the corner points $(1, 0, 0)$ and $(0, 1, 0)$.
Therefore, all points on the line segment between $(1, 0, 0)$ and $(0, 1, 0)$ solve the problem.

d. It is straightforward to verify that this is a convex optimization problem and that the constraint set has an interior point. By Corollary 5.11, the problem has a solution at some point $(\bar{x}, \bar{y}, \bar{z})$ iff the KKT conditions are satisfied at that point.
Thus, $(\bar{x}, \bar{y}, \bar{z}) \in C$ solves the problem iff. it satisfies the following conditions, with

$$L(x, y, z, \lambda_1, \lambda_2) := -e^x - y^2 - \lambda_1(x^2 + y^2 + z^2 - 1) - \lambda_2(x + y + z^2 - 3) :$$

i) $D_x L = D_y L = Dz L = 0$;
ii) $\lambda_1(x^2 + y^2 - 1) = 0$, $\lambda_2(x + y + z^2 - 3) = 0$ and
iii)$x^2 + y^2 \leq 1$, $x + y + z^2 \leq 0$ and $\lambda_i \geq 0$ for $i = 1, 2$.
Note that

$$D_x L = -e^x - 2\lambda_1 x - \lambda_2,\ D_y L = -2y(1 + \lambda_1) - \lambda_2 \text{ and } D_z L = -2z(\lambda_1 + \lambda_2).$$

From $D_z L = 0$ it follows that either $\lambda_1 = \lambda_2 = 0$ or $z = 0$. From $D_x L = 0$ we see that $\lambda_1 = 0$ is not possible. Hence, $z = 0,\ \lambda_1 > 0$ and, consequently, $x^2 + y^2 = 1$.
If $\lambda_2 = 0$, then from $D_y L = 0$ it follows that $y = 0$, so $x^2 = 1$. Using $D_x L = 0$, we obtain $x = -1$ and $\lambda_1 = \frac{1}{2}e^{-1}$. Since $x + y + z^2 \leq 3$, the point $(-1, 0, 0)$ is feasible and yields a solution.
If $\lambda_2 > 0$, then it follows that $x + y = 3$, which is incompatible with $x^2 + y^2 = 1$.

e. Note that this is a convex optimization problem. By Theorem 5.13, the problem has a solution at some point $(\bar{x}, \bar{y})$ iff the KKT conditions are satisfied at that point.
Thus, $(\bar{x}, \bar{y}) \in C$ solves the problem iff. it satisfies the following conditions, with

$$L(x, y, \lambda) := -(x - 1)^2 - (y - 1)^2 - \lambda(x + y - 1) :$$

i) $D_x L = D_y L = 0$;
ii) $\lambda(x + y - 1) = 0$ and iii) $x + y \leq 1$ and $\lambda \geq 0$.
Note that

$$D_x L = -2(x - 1) - \lambda \text{ and } D_y L = -2(y - 1) - \lambda.$$

Hence, condition i) holds iff $x = y$ and $\lambda = -2(x-1)$.

From condition ii), either $\lambda = 0$ or $x + y = 1$. If $\lambda = 0$, then $x = y = 1$, which violates condition iii). Therefore, we obtain the unique solution $(x, y, \lambda) = (\frac{1}{2}, \frac{1}{2}, 1)$.

f. We obtain the same conditions as in item c, but we now need to check whether this set of conditions has a solution within C. It is easy to verify that no such solution exists. Therefore, by Theorem 5.13, we conclude that the problem has no solution.

Notice, however, that the problem does have a supremum at $(x, y) = (0, 0)$ (for example, consider the optimization problem with constraints $x \leq 0$, $y \leq 0$ and $C = \mathbb{R}^2$).

g. Note that this is a convex optimization problem. Moreover, the constraint set has an interior point (for example $(2, 2)$ satisfies both constraints strictly). Therefore, Slater's condition holds, and by Corollary 5.11, the problem has a solution at some point $(\bar{x}, \bar{y})$ iff. the KKT conditions are satisfied at that point.

Thus, $(\bar{x}, \bar{y}) \in C$ solves the problem iff. it satisfies the following, with

$$L(x, y, \lambda) := -(x-1)^2 - (y-1)^2 + \lambda_1(x + y - 3) - \lambda_2(x^2 + y^2 - 9) :$$

i) $D_x L = D_y L = 0$;

ii) $\lambda_1(x + y - 3) = 0$ and $\lambda_2(x^2 + y^2 - 9) = 0$ and

iii)$x + y \geq 3$, $x^2 + y^2 \leq 9$ and $\lambda_i \geq 0$ for $i = 1, 2$.

Note that

$$D_x L = -2(x-1) + \lambda_1 - 2\lambda_2 x \text{ and } D_y L = -2(y-1) + \lambda_1 - 2\lambda_2 y.$$

To solve this set of equations consider next 3 possibilities:

1. No constraint is binding. Then $\lambda_1 = \lambda_2 = 0$, which gives $(x, y) = (1, 1)$, but this is not feasible.
2. Precisely one constraint is binding.
   1. $x + y = 3$. Then $\lambda_2 = 0$. From i), $x = y$ and $\lambda_1 = 2(x-1)$. Substituting into the other conditions gives the candidate solution $(x, y, \lambda_1, \lambda_2) = (\frac{3}{2}, \frac{3}{2}, 1, 0)$.
   2. $x^2 + y^2 = 9$. Then $\lambda_1 = 0$. From i), $x = y$ and $(1 + \lambda_2)x = 1$, giving $(x, y) = (\frac{3}{\sqrt{2}}, \frac{3}{\sqrt{2}})$. However, this results in $\lambda_2 < 0$, which violates iii).
3. Both constraints are binding. Then $(x, y) = (0, 3)$. But this gives $\lambda_1 < 0$, which also violates iii).

Hence, the minimum is attained at $(x, y) = (\frac{3}{2}, \frac{3}{2})$.

h. This is a linear programming problem. Moreover, the constraint set is compact. To find the maximum, we evaluate the objective function at the corner-points $(1, 2)$, $(0, 4)$ and $(3, 1)$. The corresponding function values are: 1, 4 and -2, respectively. Hence, the LP problem has a unique solution at $(x, y) = (0, 4)$.

i. Note that $-f(x, y, z)$ is concave, and both constraint functions $g_1 := -\ln(1 + x + y + z)$ and $g_2(x, y, z) := -x - y - z$ are convex. Moreover, the constraint set has an interior point. Hence, by Corollary 5.11, the problem has a solution iff. the KKT conditions are satisfied. With

$$L(x, y, \lambda) := -(2x + y + z) + \lambda_1(\ln(1 + x + y + z) - 1) + \lambda_2(x + y + z),$$

we obtain

$$DL = [-2 + \frac{\lambda_1}{1+x+y+z} + \lambda_2,\ -1 + \frac{\lambda_1}{1+x+y+z} + \lambda_2,\ -1 + \frac{\lambda_1}{1+x+y+z} + \lambda_2].$$

Obviously, the system $DL = [0\ 0\ 0]$ has no solutions. Consequently, the problem has no solution.

j. This is an LP-problem with

$$c = [1\ 2\ -3\ 4\ -5]^T,\ A = \begin{bmatrix} 5 & -4 & 3 & -2 & -1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix},\ \text{and } b = [10\ 2]^T.$$

Clearly, the constraint set is compact, so the problem has a solution.
The corresponding dual LP-problem is to minimize

$$f_D(\lambda_1, \lambda_2) = 10\lambda_1 + 2\lambda_2$$

subject to

$$5\lambda_1 + \lambda_2 \geq 1,\ -4\lambda_1 + \lambda_2 \geq 2,\ 3\lambda_1 + \lambda_2 \geq -3,\ -2\lambda_1 + \lambda_2 \geq 4 \text{ and } -\lambda_1 + \lambda_2 \geq -5.$$

In fact the feasible set is determined by the 3 inequalities

$$\lambda_1 \geq 0,\ \lambda_2 \geq 2 + 4\lambda_1 \text{ and } \lambda_2 \geq 4 + 2\lambda_1.$$

Since the LP-problem has a solution, we know that the solution of the dual LP-problem is attained at least at one corner point. The corner points of the feasible set are $(0, 4)$ and $(1, 6)$. Evaluating f_D at these points shows that the minimum is attained at $(0, 4)$, with value 8. Hence, the dual problem has a unique solution at $(0, 4)$.
The Lagrange function of the primal LP-problem is

$$L := c^T x - \lambda^T (Ax - b) + \mu^T x.$$

From the KKT conditions, substituting $\lambda^T = [0\ 4]$, we obtain

$$\mu_1 = 3,\ \mu_2 = 2,\ \mu_3 = 7,\ \mu_4 = 0 \text{ and } \mu_5 = 9.$$

Thus, $x_i = 0$ for $i = 1, 2, 3, 5$. Since the optimal value of f is 8, it follows that $x_4 = 2$. Therefore, the unique solution of the LP-problem is attained at $(0, 0, 0, 2, 0)$.

EXERCISE 5.9 Show there exists a $\lambda \in (0, 1)$ such that $\lambda \ln(x) \leq (1 - \lambda)(x - 1)$ for all $x > 0$.

ANSWER 5.9 Note that both $f_1(x) := \ln(x)$ and $f_2(x) := 1 - x$ are concave for $x > 0$. Moreover, the set $S = \{x > 0 \mid f_1(x) > 0 \text{ and } f_2(x) > 0\}$ is empty.
Therefore, by Lemma 5.8, there exists a $\lambda \in (0, 1)$ such that $\lambda f_1(x) + (1 - \lambda) f_2(x) \leq 0$, for all $x > 0$. This implies the stated result.

EXERCISE 5.10 Assume $a,\ b,\ c_1,\ c_2 \in \mathbb{R}$ are some given constants. Consider the maximization of $\sqrt{x_1 + x_2}$ subject to the constraints $\max\{x_1, x_2\} \leq a$ and $\frac{x_1^2}{x_2} \leq b$, $x_i \geq c_i > 0,\ i = 1, 2$. Show that if above optimization problem has a solution, its solution will provide a global maximum.

ANSWER 5.10 According to Theorem 5.6, any solution to the problem is a global maximum if we can show that the problem is a convex programming problem. To establish this, first consider $f(x_1, x_2) := \sqrt{x_1 + x_2}$. A straightforward calculation shows that

$$f" = \frac{-1}{(x_1 + x_2)^{\frac{3}{2}}} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Hence, $-f$" is positive semi-definite, which implies that the objective function f is concave.
Clearly, the linear functions $g_i(x_1, x_2) := -x_i$ for $i = 1, 2$, are convex.
Next consider $g_3(x_1, x_2) := \frac{x_1^2}{x_2}$. Straightforward calculations yield

$$g_3' = [\frac{2x_1}{x_2} \ -\frac{x_1^2}{x_2^2}] \text{ and therefore } g_3" = \begin{bmatrix} \frac{2}{x_2} & \frac{-2x_1}{x_2^2} \\ \frac{-2x_1}{x_2^2} & \frac{2x_1^2}{x_2^3} \end{bmatrix}.$$

Note that both $\frac{2}{x_2}$ and $\frac{2x_1^2}{x_2^3}$ are positive on the domain. Moreover, the determinant of $g_3"$ equals zero. Hence, g_3 is also convex function on its domain.
Finally, consider $g_4(x_1, x_2) := \max\{x_1, x_2\}$. To show that this function is convex, we use the definition of convexity. Let $x = (x_1, x_2)^T$, $y = (y_1, y_2)^T$, and $\lambda \in [0, 1]$. Then,

$$\begin{aligned} g_4(\lambda x + (1 - \lambda)y) &= \max\{\lambda x_1 + (1 - \lambda)y_1, \lambda x_2 + (1 - \lambda)y_2\} \\ &\leq \lambda \max\{x_1, x_2\} + (1 - \lambda) \max\{y_1, y_2\} \\ &= \lambda g_4(x) + (1 - \lambda) g_4(y). \end{aligned}$$

Thus, the objective function is concave and all constraint functions are convex. Therefore, the problem is a convex programming problem.

EXERCISE 5.11 Consider the problem to minimize $f(x, y) = x^2 + 4x - 2xy + y^2 - 4y$ subject to the constraints $x \geq 0$; $y \geq 0$ and $(x - 3)^2 + (y - 3)^2 \geq 4$.
a. Show the constraint set is not compact.
b. Show that this is not a convex optimization problem.
c. Show, using Fritz-John's theorem, that the problem has an infinite number of global minima.

d. Show that the problem has no local minima.

ANSWER 5.11

a. For arbitrarily large values of a and b, all constraints are satisfied at the point (a, b). Therefore, the feasible set is unbounded and, consequently, not compact.

b. $g_3(x, y) := -(x-3)^2 - (y-3)^2 + 4$ is not convex (in fact it is concave).

c. Let $\tilde{f}(x, y) := -x^2 - 4x + 2xy - y^2 + 4y$; $g_1(x, y) := -x$; $g_2(x, y) := -y$ and $g_3(x, y) := -(x-3)^2 - (y-3)^2 + 4$.

By inspecting the partial derivatives of $\tilde{f}$ and the functions g_i, it is straightforward to verify that all functions are continuously differentiable.

Consider the Lagrange function

$$L(x, y, \lambda_0, \lambda_1, \lambda_2, \lambda_3) := \lambda_0(-x^2 - 4x + 2xy - y^2 + 4y) + \lambda_1 x + \lambda_2 y + \lambda_3((x-3)^2 + (y-3)^2 - 4).$$

By Theorem 4.32, candidate optimal solutions must satisfy:

1. $[D_1 L(x, y, \lambda)\ D_2 L(x, y, \lambda)] = [\lambda_0(-2x - 4 + 2y) + \lambda_1 + 2\lambda_3(x-3),\ \lambda_0(2x - 2y + 4) + \lambda_2 + 2\lambda_3(y-3)] = [0,\ 0]$;
2. $i)\ \lambda_1 x = 0$; $ii)\ \lambda_2 y = 0$; and $\lambda_3((x-3)^2 + (y-3)^2 - 4) = 0$;
3. $i)\ x \geq 0$; $ii)\ y \geq 0$; $iii)\ (x-3)^2 + (y-3)^2 \geq 4$ and $iv)\ \lambda_i \geq 0,\ i = 1, 2, 3$;
4. $(\lambda_0, \lambda_1, \lambda_2, \lambda_3) \neq (0, 0, 0, 0)$, where $\lambda_0 \in \{0, 1\}$.

To solve this system, we distinguish two cases.

Case 1: $\lambda_0 = 1$.

1. No constraint is binding.

   Then $\lambda_i = 0$ for $i = 1, 2, 3$. From item 1. we obtain $y = x + 2$. Thus, for $x > 0$ and $(x-3)^2 + (x-1)^2 > 4$ we obtain feasible solutions. This holds when either $0 < x < 1$ or $x > 3$. Hence $(x, 2 + x, 1, 0, 0, 0, 0)$, with either $0 < x < 1$ or $x > 3$, are candidate solutions.
2. Exactly one constraint is binding:
   1. $g_1 = 0$. That is: $x = 0$ and $\lambda_2 = \lambda_3 = 0$. Then, by item 1, $y = 2$ and $\lambda_1 = 0$. This point satisfies all constraints, so $(0, 2, 1, 0, 0, 0, 0)$ is a candidate.
   2. $g_2 = 0$. That is: $y = 0$ and $\lambda_1 = \lambda_3 = 0$. Then, by item 1, $x = -2$, which violates the constraint on g_1.
   3. $g_3 = 0$. That is: $(x-3)^2 + (y-3)^2 = 4$ and $\lambda_1 = \lambda_2 = 0$. Adding both derivatives in item 1 shows $\lambda_3(x + y - 6) = 0$.
      1. If $x + y - 6 = 0$, substitution into $g_3 = 0$ gives either $x = 3 + \sqrt{2}$ or $x = 3 - \sqrt{2}$. This leads to the points $(3 + \sqrt{2}, 3 - \sqrt{2}, 1, 0, 0, 2 + \sqrt{2})$ and $(3 - \sqrt{2}, 3 + \sqrt{2}, 1, 0, 0, 2 - \sqrt{2})$, both of which are feasible.
      2. If $\lambda_3 = 0$, we obtain the candidate solutions $(3, 5, 1, 0, 0, 0, 0)$ and $(1, 3, 1, 0, 0, 0, 0)$.
3. Exactly two constraints are binding:

   The only feasible solution $(x, y) = (0, 0)$ does not satisfy the necessary conditions.

Case 2: $\lambda_0 = 0$:

This case always leads to an inconsistent system, and hence no feasible solutions arise.

So, the candidate solutions are

$$(x, 2+x, 1, 0, 0, 0, 0), \text{ with } 0 \leq x \leq 1 \text{ or } x \geq 3; (3+\sqrt{2}, 3-\sqrt{2}, 1, 0, 0, 2+\sqrt{2}); \text{and}$$
$$(3-\sqrt{2}, 3+\sqrt{2}, 1, 0, 0, 2+\sqrt{2}).$$

To determine whether one of these candidates yields the solution of the problem, we consider

$$L(x, y, 1, 0, 0, 0, 0) = -x^2 - 4x + 2xy - y^2 + 4y.$$

Straightforward calculations show

$$DL(x, y, 1, 0, 0, 0, 0) = [-2x - 4 + 2y,\ 2x - 2y + 4] \text{and } HL(x, y, 1, 0, 0, 0, 0) = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}.$$

This Hessian matrix is negative semi-definite, so the function $L(x, y, 1, 0, 0, 0, 0)$ is concave.

By Proposition 4.11, the problem attains a global minimum at all points $(x, 2+x, 1, 0, 0, 0, 0)$, with $0 \leq x \leq 1$ or $x \geq 3$. At these points, the value of f equals -4.

Finally, for the remaining candidates $(3+\sqrt{2}, 3-\sqrt{2}, 1, 0, 0, 2+\sqrt{2})$ and $(3-\sqrt{2}, 3+\sqrt{2}, 1, 0, 0, 2-\sqrt{2})$ the value of the objective function is $f(3+\sqrt{2}, 3-\sqrt{2}) = 8(1+\sqrt{2})$ and $f(3-\sqrt{2}, 3+\sqrt{2}) = 8(1-\sqrt{2})$, respectively. Hence, these points do not yield a global minimum.

d. To determine whether f has a local minimum at the two points $(3+\sqrt{2}, 3-\sqrt{2}, 1, 0, 0, 2+\sqrt{2})$ and $(3-\sqrt{2}, 3+\sqrt{2}, 1, 0, 0, 2-\sqrt{2})$ we apply Theorem 4.19.

To this end, first note that both g_1 and g_2 are non-binding at these points, whereas g_3 is binding. Moreover, $Dg_3 = [-2(x-3),\ -2(y-3)]$. Hence, $Dg_3(3+\sqrt{2}, 3-\sqrt{2}) = [-2\sqrt{2},\ 2\sqrt{2}]$. Following the notation in Theorem 4.19 we next consider:

$$\begin{aligned} \tilde{L} &:= HL(3+\sqrt{2}, 3-\sqrt{2}, 1, 0, 0, 2+\sqrt{2}) \\ &= \begin{bmatrix} -2\lambda_0 + 2\lambda_3 & 2\lambda_0 \\ 2\lambda_0 & -2\lambda_0 + 2\lambda_3 \end{bmatrix} = \begin{bmatrix} 2+2\sqrt{2} & 2 \\ 2 & 2+2\sqrt{2} \end{bmatrix}. \end{aligned}$$

Consequently, with $B := [1,\ -\frac{D_y g_3(3+\sqrt{2}, 3-\sqrt{2})}{D_x g_3(3+\sqrt{2}, 3-\sqrt{2})}]$,

$$B\tilde{L}B^T = [1\ 1] \begin{bmatrix} 2+2\sqrt{2} & 2 \\ 2 & 2+2\sqrt{2} \end{bmatrix} [1\ 1]^T > 0.$$

Thus, by Theorem 4.19, f does not attain a local minimum at $(3+\sqrt{2}, 3-\sqrt{2})$.
Next, consider the point $(3-\sqrt{2}, 3+\sqrt{2})$. At this point,

$$\begin{aligned}\tilde{L} &:= HL(3-\sqrt{2}, 3+\sqrt{2}, 1, 0, 0, 2-\sqrt{2}) \\ &= \begin{bmatrix} -2\lambda_0 + 2\lambda_3 & 2\lambda_0 \\ 2\lambda_0 & -2\lambda_0 + 2\lambda_3 \end{bmatrix} = \begin{bmatrix} 2-2\sqrt{2} & 2 \\ 2 & 2-2\sqrt{2} \end{bmatrix}.\end{aligned}$$

Hence, with $B := [1, \ -\frac{D_y g_3(3-\sqrt{2}, 3+\sqrt{2})}{D_x g_3(3-\sqrt{2}, 3+\sqrt{2})}]$,

$$B\tilde{L}B^T = [1 \ 1] \begin{bmatrix} 2-2\sqrt{2} & 2 \\ 2 & 2-2\sqrt{2} \end{bmatrix} [1 \ 1]^T > 0.$$

Therefore, again by Theorem 4.19, f does not have a local minimum at $(3-\sqrt{2}, 3+\sqrt{2})$.

Advanced, Deepening and Broadening Exercises

EXERCISE 5.12 *(Importance differentiability assumptions)*
This exercise demonstrates that in Step 1 of Algorithm 4.9 it is important to include any points where either f or g are not differentiable too in the search for potential optima.
Consider the maximization of $f(x) = \sqrt{1-x^2}, \ x \in [-1, 1]$ subject to $g(x) = x - 1 \geq 0$.

a. Use the theory presented in this section to show that this problem has a maximum at $x = 1$.
b. Let $X = [-1, 1]$ and $\Lambda = \{\lambda \mid \lambda \geq 0\}$. Show that the with this problem corresponding Lagrange function $L(x, \lambda) := \sqrt{1-x^2} - \lambda(1-x)$ has not a saddle point at $(1, \bar{\lambda})$, whatever $\bar{\lambda} \geq 0$ is, on the set $X \times \Lambda$.
c. Compare the result obtained in item b. with Theorem 5.13. Conclude.

ANSWER 5.12

a. Following Algorithm 4.9, we find the candidate solutions as follows.
   Step 1. Note, f and g are twice differentiable on the open convex set $(-1, 1)$, but f is not differentiable at $x^2 = 1$. As $x \geq 1$, $x = 1$ is a candidate point for an extremum.
   Step 2. $g' = 1$. So there is no extra candidate.
   Step 3. $L(x) = \sqrt{1-x^2} - \lambda(1-x)$ on $x \in (-1, 1)$.
   Step 4. Find all points satisfying:

$$1. \ \frac{-x}{\sqrt{1-x^2}} + \lambda = 0, \ 2. \ \lambda(1-x) = 0, \text{ and } 3. \ x \geq 1 \text{ and } \lambda \geq 0.$$

   From 2. it follows that $\lambda = 0$. From 1 it follows then that $x = 0$, which violates 3. So, we find no additional candidate points for an extremum.

So, the only candidate solution is $x = 1$. Clearly $x = 1$ solves the problem.

b. With $L(x, \lambda) = \sqrt{1 - x^2} - \lambda(1 - x)$, we have $L(1, \lambda) = 0$ for all $\lambda \geq 0$. Now suppose $\bar{\lambda} > 0$. Then it is clear that the inequality $L(x, \bar{\lambda}) \leq L(1, \bar{\lambda}) = 0$ does not hold for all $x \in [-1, 0)$. In the case $\bar{\lambda} = 0$, it follows that the inequality $L(x, \bar{\lambda}) \leq L(1, \bar{\lambda}) = 0$ does not hold for all $x \in (-1, 1)$. Therefore, $L(x, \lambda)$ does not have a saddle-point on $X \times \Lambda$ at $(1, \lambda)$, for any $\lambda \geq 0$.

c. The only condition that fails in Theorem 5.13 is that f is not differentiable on an open set X. In this case, the set X is closed, and at the boundary points $x = \pm 1$, the derivative of f cannot be defined in an appropriate sense - namely, there is no open neighborhood around these points on which the derivative of f can be properly defined.

From item b., we see that the result stated in the theorem does not hold. Hence, the requirement that f be differentiable on an open set X is essential to the theorem to apply.

EXERCISE 5.13 *(A function is convex iff its epigraph is a convex set).*

Consider the set of points located above the graph of a function $f : \mathbb{R}^n \to \mathbb{R}$. This set is called the *epigraph* of f. Formal: epi $f := \{(x, \gamma) \mid x \in$ domain f and $f(x) \leq \gamma\}$.

a. Determine epi f if $f(x) = x^2$ with $x \in (0, 2)$.

b. Let $S \subset \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R}$. Show that $f(x)$ is convex on S if and only if its epigraph on S is a convex set.

c. Formulate a similar result for function g if g is concave.

ANSWER 5.13

a. epi $f = \{(x, \gamma) \mid \gamma \geq x^2, \text{ with } 0 < x < 2\}$.

b. " $\Rightarrow$ " Assume (x, γ_x) and (y, γ_y) belong to epi f. Then $f(x) \leq \gamma_x$ and $f(y) \leq \gamma_y$.

Since f is convex, it follows that for all $\alpha \in [0, 1]$,

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$$

Hence, $f(\alpha x + (1 - \alpha)y) \leq \alpha\gamma_x + (1 - \alpha)\gamma_y$, which implies that

$$(\alpha x + (1 - \alpha)y, \alpha\gamma_x + (1 - \alpha)\gamma_y) \in \text{ epi } f.$$

Equivalently, $\alpha(x, \gamma_x) + (1 - \alpha)(y, \gamma_y) \in$ epi f.

" $\Leftarrow$ " Assume $(x, f(x))$ and $(y, f(y))$ belong to epi f. Then for all $\alpha \in [0, 1]$,

$$\alpha(x, f(x)) + (1 - \alpha)(y, f(y)) \in \text{ epi } f,$$

which implies $f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$.

c. g is concave $\leftrightarrow f := -g$ is convex $\leftrightarrow$ epi f is convex $\leftrightarrow$ the set $\{(x, \gamma) \mid x \in$ domain g and $-g(x) \leq \gamma\}$ is convex $\leftrightarrow \{(x, \gamma) \mid x \in$ domain g and $g(x) \geq -\gamma\}$ is convex $\leftrightarrow \{(x, -\gamma) \mid x \in$ domain g and $g(x) \geq \gamma\}$ is convex $\leftrightarrow \{(x, \gamma) \mid x \in$ domain g and $g(x) \geq \gamma\}$ is convex.

In other words, g is concave if and only if the set $\{(x, \gamma) \mid x \in \text{domain } g \text{ and } g(x) \geq \gamma\}$ is convex. That is, the set of points lying below the graph of g is convex.

EXERCISE 5.14 *(Separating Hyperplane Theorem)*

Let $S \subset \mathbb{R}^n$ be a closed convex set. Let $y \notin S$. Below you are asked to show that there exists a plane $H: \{x \in \mathbb{R}^n \mid a_1 x_1 + \cdots + a_n x_n = b\}$ that separates y and S. That is, there is a plane H such that y is located to one side of H and S to the other side. Without loss of generality we will assume that $n = 2$.

a. Let $y = (0, 0)$ and $S = \{(x_1, x_2) \mid x_2 \geq x_1^2 + 1\}$. Construct geometrically a line that separates y and S.

b. Consider the distance function $d(x) = \|x - y\|$, where $x \in S$. Motivate why $d(x)$ is a continuous function on S.

c. Consider a closed ball B with center y and finite radius r, $B = \{(x_1, x_2) \mid (x_1 - y_1)^2 + (x_2 - y_2)^2 \leq r\}$, where r is such that $I := B \cap S$ is nonempty. Show that I is a compact set.

d. Show that $d(x)$ attains a global minimum on I. Let x_0 be the argument where $d(x)$ attains this minimum.

e. Let $a := x_0 - y$ and $b := a^T x_0$. Show that $a^T y < b$ and $a^T x_0 = b$.

f. Let $x \in S$. Let $x(\lambda) := \lambda x + (1 - \lambda) x_0, \ \lambda \in [0, 1]$. Note that, since S is convex, $x(\lambda) \in S$, for all $\lambda \in [0, 1]$. Motivate why $\|x(\lambda) - y\|^2 \geq \|x_0 - y\|^2$, for all $\lambda \in [0, 1]$. Use this inequality to show that $(x - x_0)^T (x_0 - y) \geq 0$.

g. Use item e. to conclude that $a^T x \geq b$, for all $x \in S$.

h. Assume $0 \notin S$. Show there exists a vector $a \in \mathbb{R}^n$ and $b > 0$ such that $a^T x \geq b$, for all $x \in S$.

ANSWER 5.14

b. The function $d(x) = (x - y)^T (x - y)$ is quadratic in x, and therefore continuous.

c. The set I is bounded since B is bounded. Moreover, both S and B are closed, hence I is closed. Therefore, I is compact.

d. The function $d(x)$ is continuous on the compact set I. Hence, by Weierstrass' theorem, $d(x)$ attains a global minimum on I.

e. We have $a^T y - b = a^T (y - x_0) = -a^T a < 0$. Moreover, $a^T x_0 = b$ follows directly from the definition.

f. Since $d(x_0) \leq d(x)$ for all $x \in S$, it follows that

$$\|x(\lambda) - y\|^2 \geq \|x_0 - y\|^2 \text{ for all } \lambda \in [0, 1].$$

That is, $\|\lambda x + (1 - \lambda) x_0 - y\|^2 \geq \|x_0 - y\|^2$, or equivalently, $\|\lambda (x - x_0) + x_0 - y\|^2 \geq \|x_0 - y\|^2$. Expanding and simplifying yields $\lambda^2 \|x - x_0\|^2 + 2\lambda (x - x_0)^T (x_0 - y) \geq 0$, or

$$\lambda (\lambda \|x - x_0\|^2 + 2(x - x_0)^T (x_0 - y)) \geq 0.$$

Since this inequality holds for all $\lambda \in [0, 1]$, it follows that $2(x - x_0)^T (x_0 - y) \geq 0$.

g. We obtain, by item f,

$$a^T x - b = a^T x - a^T x_0 = a^T(x - x_0) = (x_0 - y)^T(x - x_0) \geq 0.$$

h. Taking $y = 0$, it follows from item e. that $a := x_0 \neq 0$ and $b := a^T x_0 = x_0^T x_0 > 0$. The result then follows from item g.

EXERCISE 5.15 *(Proof Lemma 5.8)*.

In this exercise you are asked to provide a proof for Lemma 5.8. Use the notation from this Lemma 5.8.

a. Define for $x \in X$, the set $Z_x := \{z \in \mathbb{R}^m \mid f_i(x) > z_i,\ i = 1, \cdots, m\}$. Let $Z := \cup_{x \in X} Z_x$. Show that $0 \notin Z$.

b. Show that Z is a convex set.

c. Show that with $z \in Z_x$ also for every $\epsilon \geq 0$, $z - \epsilon \in Z$. Also show that if ϵ_i is defined by $\epsilon_i := f_i(x) - z_i - \delta_i$, for arbitrary $\delta_i > 0$, $\tilde{z} := z + \epsilon \in Z$.

d. Use Lemma 5.7 to conclude that there exists a $p \leq 0$, $p \neq 0$, such that $p^T z \geq 0$ for all $z \in Z$.

e. Let $\lambda_i := \frac{p_i}{\sum_{i=1}^m p_i}$. Use item c to conclude that there exists a $\lambda \in \Lambda$, $\lambda \neq 0$, such that $\sum_{i=1}^m \lambda_i f_i(x) \leq 0$, for all $x \in X$.

ANSWER 5.15

a. By assumption the set $S = \{x \mid f_i(x) > 0,\ i = 1, \cdots, m\}$ is empty. Or, stated differently, for every fixed $x \in X$, $0 \notin Z_x$. So, $0 \notin Z$ too.

b. Let $z, \bar{z} \in Z$. Then, with $\lambda \in [0, 1]$, there exist x and $\bar{x}$ in Z such that

$$\lambda z_i + (1 - \lambda)\bar{z}_i < \lambda f_i(x) + (1 - \lambda) f_i(\bar{x}) \leq f_i(\lambda x + (1 - \lambda)\bar{x}).$$

Hence, $\lambda z + (1 - \lambda)\bar{z} \in Z_{\lambda x + (1-\lambda)\bar{x}}$.

c. With $z \in Z_x := \{z \in \mathbb{R}^m \mid f_i(x) > z_i,\ i = 1, \cdots, m\}$ and $\epsilon \geq 0$ it follows that also $f_i(x) - (z_i - \epsilon_i) > 0,\ i = 1, \cdots, m$. That is, also $z - \epsilon \in Z_x$.

Similarly it follows that with $\tilde{z}$ as defined, $f_i(x) - \tilde{z}_i = f_i(x) - (z_i - \epsilon_i - \delta_i) = \delta_i > 0$.

d. Since $0 \notin Z$ and Z is convex, by Lemma 5.7 there exists $p \neq 0$ such that $p^T z \geq p^T 0 = 0$ for all $z \in Z$.

Now assume $p_i > 0$ for some i and $z \in Z$. Then there exists a $x \in X$ such that $z \in Z_x$. Now consider entry z_i of z. From the definition of $Z_x := \{z \in \mathbb{R}^n \mid f_i(x) > z_i,\ i = 1, \cdots, m\}$ it is clear that then also for every $\bar{z}_i < z_i$, $\bar{z} := (z_1, \cdots, z_{i-1}, \bar{z}_i, z_{i+1}, \cdots, z_m) \in Z$. But then there exists a $\bar{z} \in Z$ such that $p^T \bar{z} < 0$. Which violates our observation that $p^T z \geq = 0$ for all $z \in Z$. So, p_i must be nonpositive.

e. With $\lambda_i := \frac{p_i}{\sum_{i=1}^m p_i}$, it follows from item d that $\lambda^T z \leq 0$ for all $z \in Z$. Now, let $z \in Z_x$. Then by definition of Z_x, we have that $f_i(x) - z_i =: \epsilon_i > 0$. By item c we have that for arbitrary $\delta > 0$ then

also $\tilde{z}$, where $\tilde{z}_i = f_i(x) - z_i - \delta_i \in Z$. So, for arbitrary $\delta > 0$, we have $0 \geq \lambda^T \tilde{z} = \lambda^T(f - z - \delta) \geq \lambda^T f - \lambda^T \delta$. From which we conclude that $0 \leq \lambda^T f$.

EXERCISE 5.16 *(counterexample of Theorem 5.10 if Slater's condition fails).*
Consider the maximization of $f(x) = x$ subject to $g(x) = -x^2 \geq 0$.

a. Use the theory presented in this section to show that this problem has a maximum at $x = 0$.
b. Show that the with this problem corresponding Lagrange function $L(x, \lambda) = x - \lambda x^2$ has not a saddle point at $(0, \bar{\lambda})$, whatever $\bar{\lambda} \geq 0$ is.

ANSWER 5.16

a. Consider Algorithm 4.9 to determine candidate solutions for the problem:
   Step 1. Note that both f and g are twice differentiable.
   Step 2. Since $g^{'} = -2x$, we have $g^{'}(x) = 0$ at $x = 0$. Moreover, $g(0) = 0$. Hence, $x = 0$ is a candidate point for an extremum.
   Step 3. Consider $L(x) := x - \lambda x^2$.
   Step 4. Determine all points satisfying:

$$1.\ 1 - 2\lambda x = 0,\ 2.\ \lambda x^2 = 0,\ \text{and } 3.\ x^2 \leq 0,\ \lambda \geq 0.$$

   From 2., it follows that either $x = 0$ or $\lambda = 0$. However, both cases violate condition 1. Therefore, no additional candidate points for an extremum are found.
   Hence, the only candidate solution is $x = 0$, which clearly solves the problem.
b. Note that $L(0, \bar{\lambda}) = 0 \leq 0 = L(0, \lambda)$ for all $\bar{\lambda},\ \lambda \in \Lambda = \{\lambda \mid \lambda \geq 0\}$. However, for every $\bar{\lambda} \in \Lambda$ with $\bar{\lambda} > 0$,

$$L(x, \bar{\lambda}) = x(1 - \bar{\lambda}x) > 0$$

   for example at $x = \frac{1}{2\bar{\lambda}}$. Whereas for $\bar{\lambda} = 0$ we have, for example, $L(1, 0) > 0$. Thus, for any $\bar{\lambda} \in \Lambda$, there exists a x such that $L(x, \bar{\lambda}) > L(0, \bar{\lambda})$.

EXERCISE 5.17 *(Any saddle-point of Lagrange function solves KKT conditions)*
Assume f and g_i are all differentiable on an open set $X \subset \mathbb{R}^n$. Let $\Lambda := \{\lambda \in \mathbb{R}^m \mid \lambda_i \geq 0\}$.
Show that if the Lagrange function $L(x, \lambda) = f(x) - \lambda^T(g(x) - b)$ has a saddle point equilibrium at $(\bar{x}, \bar{\lambda}) \in X \times \Lambda$, $(\bar{x}, \bar{\lambda})$ satisfy the necessary conditions 1-3 of the Kuhn-Tucker Theorem 4.15.

ANSWER 5.17
From the inequality $L(x, \bar{\lambda}) \leq L(\bar{x}, \bar{\lambda})$ for all $x \in X$, it follows that $L(x, \bar{\lambda})$ attains a global maximum at $x = \bar{x}$. Since $L(x, \bar{\lambda})$ is differentiable on X, we obtain

$$L_x(\bar{x}, \bar{\lambda}) = 0.$$

Thus, condition 1 of Theorem 4.15 is satisfied.
Furthermore, by item ii) of Theorem 5.9, condition 2 of Theorem 4.15 is also satisfied.
Finally, the fact that $\bar{\lambda}^T(g(\bar{x}) - b) = 0$, together with the inequality $L(\bar{x}, \bar{\lambda}) \leq L(\bar{x}, \lambda)$, for all $\lambda \in \Lambda$, implies that

$$\lambda^T(g(\bar{x}) - b) \leq 0 \text{ for all } \lambda \in \Lambda.$$

Hence, $g_i(\bar{x}) - b_i \leq 0$, which establishes condition 3 of Theorem 4.15.

EXERCISE 5.18 *(Slater's constraint qualification in convex programming problems)*
Assume the concave function f and convex functions g_i are all differentiable on an open convex set $X \subset \mathbb{R}^n$.

a. Let $(\bar{x}, \bar{\lambda})$ satisfy the necessary conditions 1-3 of the Kuhn-Tucker Theorem 4.15. Show that the corresponding Lagrange function $L(x, \lambda) = f(x) - \lambda^T(g(x) - b)$ then has a saddle point equilibrium at $(\bar{x}, \bar{\lambda})$.
b. Assume Slater's condition (5.1) applies. Use Exercise 5.17 to show that the constrained optimization problem (4.8) has a solution if and only if conditions 1-3 of Theorem 4.15 apply. (Notice the rank condition in Theorem 4.15 is dismissed here.)

ANSWER 5.18

a. First notice that since $g_i(x)$ is convex, $b_i - g_i(x)$ is a concave function. Therefore

$$L(x, \bar{\lambda}) = f(x) - \bar{\lambda}^T(g(x) - b) = f(x) + \bar{\lambda}^T(b - g(x))$$

is concave, because $\bar{\lambda}^T(b - g(x))$ is a nonnegative linear combination of concave functions and $f(x)$ itself is concave.
Since, by assumption, $L_x(\bar{x}, \bar{\lambda}) = 0$, the concavity of $L(x, \bar{\lambda})$ implies that $L(x, \bar{\lambda})$ attains a global maximum at $\bar{x}$. Hence,

$$L(x, \bar{\lambda}) \leq L(\bar{x}, \bar{\lambda}) \text{ for all } x \in X.$$

On the other hand, from the facts that $\bar{\lambda}^T(g(\bar{x}) - b) = 0$, $g_i(\bar{x}) - b_i \leq 0$ and $\lambda_i \geq 0$, it follows directly that

$$L(\bar{x}, \bar{\lambda}) \leq L(\bar{x}, \lambda) \text{ for all } \lambda \in \Lambda.$$

b. Assume the constrained optimization problem (4.8) has a solution. Then, under the stated assumptions, Theorem 5.10 guarantees that the Lagrange function $L(x, \lambda)$ possesses a saddle point. By Exercise 5.17, this in turn implies that conditions 1-3 of Theorem 4.15 are satisfied.
Conversely, suppose that conditions 1-3 of Theorem 4.15 hold. Then, by item a., the Lagrange function $L(x, \lambda)$ has a saddle point. By item i) of Theorem 5.9, this implies that the constrained optimization problem has a solution.

EXERCISE 5.19 *(Proof Theorem 5.9)*

Adopt the notation used in Theorem 5.9.

a. Consider the inequality $L(\bar{x}, \bar{\lambda}) \leq L(\bar{x}, \lambda)$. Show that, since this equality holds $\forall \lambda \in \Lambda$, we can conclude that $g(\bar{x}) \leq b$.

b. Use item a. to show that $\bar{\lambda}^T(g(\bar{x}) - b) = 0$.

c. Use item b. and the fact that $L(x, \bar{\lambda}) \leq L(\bar{x}, \bar{\lambda})$, $\forall x \in X$, to show that $\bar{x}$ solves the constraint optimization problem.

ANSWER 5.19

a. Since $L(\bar{x}, \bar{\lambda}) \leq L(\bar{x}, \lambda)$ for all $\lambda \in \Lambda$, it follows immediately that

$$\lambda^T(g(\bar{x}) - b) \leq \bar{\lambda}^T(g(\bar{x}) - b) \text{ for all } \lambda \in \Lambda.$$

Now, suppose there exists an index i such that $g_i(\bar{x}) - b_i > 0$. By choosing λ_i arbitrarily large while setting $\lambda_j = 0$, for all $j \neq i$, we could make $\lambda^T(g(\bar{x}) - b)) > \bar{\lambda}^T(g(\bar{x}) - b)$, which contradicts the previous inequality. Therefore it must hold that

$$g_i(\bar{x}) - b_i \leq 0 \text{ for all } i.$$

In other words $\bar{x}$ satisfies the constraints.

b. Since, by item a. $g_i(\bar{x}) - b_i \leq 0$ for all i, it follows that

$$\bar{\lambda}^T(g(\bar{x}) - b) \leq 0 \text{ for all } \lambda \in \Lambda.$$

As, by item a., the inequality $\lambda^T(g(\bar{x}) - b) \leq \bar{\lambda}^T(g(\bar{x}) - b)$ holds $\forall \lambda \in \Lambda$ it follows, by taking $\lambda = 0$, that

$$\bar{\lambda}^T(g(\bar{x}) - b) \geq 0.$$

Combining these two results, we conclude that $\bar{\lambda}^T(g(\bar{x}) - b) = 0$.

c. From the inequality $L(x, \bar{\lambda}) \leq L(\bar{x}, \bar{\lambda})$ and the fact that $\bar{\lambda}^T(g(\bar{x}) - b) = 0$, it immediately follows that

$$f(x) - \bar{\lambda}^T(g(x) - b) \leq f(\bar{x}) \text{ for all } x \in X.$$

In particular, for any $x \in X$ that satisfies $g_i(x) - b_i \leq 0$ for all i, we have

$$f(\bar{x}) - f(x) \geq -\bar{\lambda}^T(g(x) - b) \geq 0,$$

since $\bar{\lambda} \geq 0$. Or, equivalently, $f(\bar{x}) \geq f(x)$, for all $x \in X$ satisfying $g_i(x) - b_i \leq 0$ for all i.

EXERCISE 5.20 *(Proof Theorem 5.10)*

Adopt the notation used in Theorem 5.10. Assume in the items a-d below that all assumptions in the theorem apply except Slater's condition.

a. Show that the intersection of the sets $\{x \in X \mid f(x) - f(\bar{x}) > 0\}$ and $\{x \in X \mid b_i - g_i(x) \geq 0\},\ i = 1, \cdots, m$ is empty.

b. Use Lemma 5.8 to show that there exist coefficients $\bar{\lambda}_i \geq 0,\ i = 0, 1, \cdots, m$, not all zero, such that $\bar{\lambda}_0(f(x) - f(\bar{x})) - \sum_{i=1}^{m} \bar{\lambda}_i(g_i(x) - b_i) \leq 0$, for all $x \in X$.

c. Use item b. to show that there exist coefficients $\bar{\lambda}_i \geq 0,\ i = 0, 1, \cdots, m$, not all zero, such that with $L_0(x, \lambda) := \lambda_0 f(x) - \sum_{i=1}^{m} \lambda_i(g_i(x) - b_i)$, $L_0(x, \bar{\lambda}) \leq L_0(\bar{x}, \bar{\lambda})$ if $\bar{x}$ solves the constraint optimization problem.

d. Use item b. to show that $\sum_{i=1}^{m} \bar{\lambda}_i(g_i(\bar{x}) - b_i) = 0$.

e. Use item b. to show by contradiction that, if Slater's condition holds, $\bar{\lambda}_0 \neq 0$.

f. Assume Slater's condition applies. Show that in that case the Lagrange function $L(x, \lambda) := f(x) - \sum_{i=1}^{m} \lambda_i(g_i(x) - b_i)$ has a saddle point at $(\bar{x}, \bar{\lambda})$, for some $\bar{\lambda} \in \Lambda$.

ANSWER 5.20

a. This follows directly from the fact that f has a maximum at $\bar{x}$ on the constraint set $\{x \in X \mid g_i(x) \leq b_i\},\ i = 1, \cdots, m$.

b. Let

$$\tilde{f}_0(x) := f(x) - f(\bar{x}) \text{ and } \tilde{f}_i(x) := b_i - g_i(x) \text{ for } i = 1, \cdots, m.$$

Then $\tilde{f}_i,\ i = 0, 1, \cdots, m$, are concave on the convex set X. Furthermore, by item a., the set

$$\tilde{S} := \{x \in X \mid \tilde{f}_0(x) > 0 \text{ and } \tilde{f}_i(x) \geq 0,\ i = 1, \cdots, m\}$$

is empty. Therefore, also the set $S := \{x \in X \mid \tilde{f}_i(x) > 0\ i = 0, \cdots, m\}$ is empty.
Consequently, by Lemma 5.8, the advertized result follows.

c. From item b., we have

$$\bar{\lambda}_0 f(x) - \sum_{i=1}^{m} \bar{\lambda}_i(g_i(x) - b_i) \leq \bar{\lambda}_0 f(\bar{x}) \leq \bar{\lambda}_0 f(\bar{x}) - \sum_{i=1}^{m} \bar{\lambda}_i(g_i(\bar{x}) - b_i).$$

In other words, $L_0(x, \bar{\lambda}) \leq L_0(\bar{x}, \bar{\lambda})$.

d. From item b., we have

$$\bar{\lambda}_0(f(x) - f(\bar{x})) - \sum_{i=1}^{m} \bar{\lambda}_i(g_i(x) - b_i) \leq 0 \text{ for all } x \in X.$$

By setting $x = \bar{x}$ in this inequality we obtain $\sum_{i=1}^{m} \bar{\lambda}_i(g_i(\bar{x}) - b_i) \geq 0$.

On the other hand, since $\bar{\lambda}_i \geq 0$ and $g_i(\bar{x}) \leq b_i,\ i = 1, \cdots, m$, it also holds that

$$\sum_{i=1}^{m} \bar{\lambda}_i (g_i(\bar{x}) - b_i) \leq 0.$$

Combining these two observations we conclude that $\sum_{i=1}^{m} \bar{\lambda}_i (g_i(\bar{x}) - b_i) = 0$, which proves the stated result.

e. Assume $\bar{\lambda}_0 = 0$ and that Slater's condition holds at some point $\hat{x} \in X$; that is, $g_i(\hat{x}) < b_i$ for all $i = 1, \cdots, m$.

From item b., we have

$$\sum_{i=1}^{m} \bar{\lambda}_i (g_i(\hat{x}) - b_i) \geq 0.$$

However, since not all $\bar{\lambda}_i$ are zero, Slater's condition implies that for at least one index i,

$$\bar{\lambda}_i (g_i(\hat{x}) - b_i) < 0.$$

Hence, also $\sum_{i=1}^{m} \bar{\lambda}_i (g_i(\hat{x}) - b_i) < 0$, which contradicts the previous inequality.

Therefore, the assumption that $\bar{\lambda}_0$ is false, and the conclusion follows.

f. Since $\bar{\lambda}_0 \neq 0$ (see item e.), it follows directly from item c. that by defining $\tilde{\lambda}_i := \frac{\bar{\lambda}_i}{\bar{\lambda}_0}$, we have

$$L(x, \tilde{\lambda}) \leq L(\bar{x}, \tilde{\lambda}).$$

Moreover, from item d. and the fact that for any $\lambda \in \Lambda$, $\lambda_i (g_i(\bar{x}) - b_i) \geq 0$, it follows that

$$f(\bar{x}) - \sum_{i=1}^{m} \lambda_i (g_i(\bar{x}) - b_i) \geq f(\bar{x}) - \sum_{i=1}^{m} \tilde{\lambda}_i (g_i(\bar{x}) - b_i),$$

or, equivalently, $L(\bar{x}, \tilde{\lambda}) \leq L(\bar{x}, \lambda)$.

EXERCISE 5.21 *(Concave function on a convex set has at most one (global) maximum)*
Consider a concave function $f(x)$ defined on a convex set C.

a. Show that if $f(x)$ attains a maximum at $\bar{x}$, $f(\bar{x})$ is a global maximum for f on S.

b. Assume f is strictly concave on C. Show that if f has a global maximum at $\bar{x}$, there is no other point in C where this maximum is attained.

ANSWER 5.21

a. Assume that there exists $c \in C$ such that $f(c) > f(\bar{x})$. Consider the line segment

$$\theta c + (1 - \theta)\bar{x},\ 0 \leq \theta \leq 1.$$

Since C is convex, every point on this line segment belongs to C.

Now fix $\theta \in (0,1)$ and define $t := \theta c + (1-\theta)\bar{x}$. Since f is concave, it follows that

$$f(t) \geq \theta f(c) + (1-\theta) f(\bar{x}) > \theta f(\bar{x}) + (1-\theta) f(\bar{x}) = f(\bar{x}).$$

Thus, in every neighborhood of $\bar{x}$, there exists a point in C at which f attains a higher value. This contradicts the assumption that f has a maximum at $\bar{x}$. Therefore, the assumption that there exists $c \in C$ such that $f(c) > f(\bar{x})$ must be false.

b. Assume that the maximum it attained at both $\bar{x}$ and s. Then, as in item a., define

$$t := \theta s + (1-\theta)\bar{x},\ \theta \in (0,1).$$

By concavity of f, we obtain

$$f(t) > \theta f(s) + (1-\theta) f(\bar{x}) = \theta f(\bar{x}) + (1-\theta) f(\bar{x}) = f(\bar{x}).$$

This contradicts the assumption that f attains a global maximum at $\bar{x}$. Hence, the assumption must be false. That is, there is no other point in C where the maximum is attained.

EXERCISE 5.22 *(First-order criterion for a maximum if concave function is differentiable)*
Assume the concave function f is differentiable on the convex set C. Below you are asked to show that f has a maximum at $\bar{x}$ on C iff $Df^T(\bar{x})(x-\bar{x}) \leq 0$ for all $x \in C$.

a. Use the first order characterization of a concave function to show that if $Df^T(\bar{x})(x-\bar{x}) \leq 0$, for all $x \in C$, f has a maximum at $\bar{x}$ on C.
b. Again, use the first order characterization of a concave function to show that if f has a maximum at $\bar{x}$ on C, $Df^T(\bar{x})(x-\bar{x})$ can never be positive for some $x \in C$.
c. Give a geometric interpretation of the condition $Df^T(\bar{x})(x-\bar{x}) \leq 0$ for all $x \in C$.

ANSWER 5.22

a. Recall from Chapter 3 that f is concave on C iff.

$$f(y) \leq f(x) + Df^T(x)(y-x) \text{ for all } x, y \in C.$$

Since $Df^T(\bar{x})(y-\bar{x}) \leq 0$ for all $y \in C$, it follows from the concavity of f that

$$f(y) \leq f(\bar{x}) + Df^T(\bar{x})(y-\bar{x}) \leq f(\bar{x}) \text{ for all } y \in C.$$

Thus, f attains a global maximum at $\bar{x}$.

b. Again, since f is concave on C, we have

$$f(y) \leq f(x) + Df^T(x)(y-x) \text{ for all } x, y \in C.$$

Now suppose $Df^T(\bar{x})(\bar{y}-\bar{x}) > 0$ for some $\bar{y} \in C$. Define

$$s := \lambda\bar{y} + (1-\lambda)\bar{x},\ \lambda \in [0,1].$$

Since C is convex, it follows that $s \in C$. Moreover,

$$\frac{df(s)}{d\lambda} = Df(s)(\bar{y}-\bar{x}) > 0$$

in a neighborhood of $\lambda = 0$. Thus, f increases from $f(\bar{x})$ when moving away from $\bar{x}$ in the direction of $\bar{y}$, which contradicts the assumption that f attains a maximum at $\bar{x}$.

c. The condition

$$Df^T(\bar{x})(x-\bar{x}) \leq 0 \text{ for all } x \in C$$

can be rewritten as

$$Df^T(\bar{x})x \leq Df^T(\bar{x})\bar{x} \text{ for all } x \in C.$$

In other words (see the Separating Hyperplane Theorem, Lemma 5.7), there exists a point $\bar{x} \in C$ such that all points in C lie on one side of the hyperplane defined by the derivative of f at $\bar{x}$.

EXERCISE 5.23 *(Proof Theorem 5.13)*

Adopt the notation used in Theorem 5.13. Let $g_i(x) := a_i^T x - b_i$, with $a_i \in \mathbb{R}^n,\ b_i \in \mathbb{R}$, with $a_i \neq 0$, $i = 1, \cdots, n$.

a. Consider item b. from Exercise 5.20. Show that choosing $\bar{\lambda}_0 = 0$ here yields a contradiction.

b. Follow the lines of the proof of Exercise 5.20 to conclude that the Lagrange function $L(x,\lambda) := f(x) - \sum_{i=1}^m \lambda_i(g_i(x) - b_i)$ has a saddle point at $(\bar{x}, \bar{\lambda})$, for some $\bar{\lambda} \in \Lambda$.

ANSWER 5.23

a. By item b. of Exercise 5.20, there exist coefficients $\bar{\lambda}_i \geq 0,\ i = 0, 1, \cdots, m$, not all zero, such that

$$\bar{\lambda}_0(f(x) - f(\bar{x})) - \sum_{i=1}^m \bar{\lambda}_i(a_i^T x - b_i) \leq 0, \text{ for all } x \in X. \tag{5.1}$$

Now suppose that $\bar{\lambda}_0 = 0$ in (5.1). Then it follows that

$$\sum_{i=1}^m \bar{\lambda}_i(a_i^T x - b_i) \geq 0 \text{ for all } x \in X.$$

Since $a_i^T x - b_i \leq 0$ and $\bar{\lambda}_i \geq 0,\ i = 1, \cdots, m$, we must have

$$\bar{\lambda}_i(a_i^T x - b_i) = 0, \text{ for all } x \in X,\ i = 1, \cdots, m.$$

Recall that at least one coefficient $\bar{\lambda}_i$ is strictly positive. Without loss of generality, assume this holds for $i = 1$. Then it follows that

$$a_1^T x - b_1 = 0, \text{ for all } x \in X.$$

Since X is an open subset of $\mathbb{R}^n$, this implies that $a_1 = 0$ and $b_1 = 0$, which contradicts the assumption that g_1 is a nontrivial linear function.

Therefore, the assumption that $\bar{\lambda}_0 = 0$ must be false.

b. As in the proof of Exercise 5.20, item f., we conclude from item b. of that exercise that

$$\bar{\lambda}_0 f(x) - \sum_{i=1}^{m} \bar{\lambda}_i (a_i^T x - b_i) \leq \bar{\lambda}_0 f(\bar{x}) \leq \bar{\lambda}_0 f(\bar{x}) - \sum_{i=1}^{m} \bar{\lambda}_i (a_i^T \bar{x} - b_i).$$

The last inequality follows from the facts that $\bar{\lambda}_i \geq 0$ and $a_i^T x - b_i \leq 0$. Thus, defining $\tilde{\lambda}_i := \frac{\bar{\lambda}_i}{\bar{\lambda}_0}$, we obtain

$$L(x, \tilde{\lambda}) \leq L(\bar{x}, \tilde{\lambda}).$$

Similarly, following the argument in item f., we conclude that $L(\bar{x}, \tilde{\lambda}) \leq L(\bar{x}, \lambda)$ for all $\lambda \in \Lambda$.

EXERCISE 5.24 *(Equivalent necessary and sufficient conditions for solving LP problem)*
In this exercise you are asked to prove next equivalent statement for Theorem 5.14 for the LP problem (5.2).
The LP problem (5.2) has a maximum at x if and only if next conditions apply:

i. $c^T - \lambda^T A \leq 0$; $(c^T - \lambda^T A)x = 0$;

ii. $\lambda_i (a_i^T x - b_i) = 0, \ i = 1, \cdots, m$;

iii. $\lambda \geq 0$; $Ax \leq b$ and $x \geq 0$.

Prove this result using Theorem 5.14.

ANSWER 5.24

$\Rightarrow$ From Theorem 5.14, item i., it follows immediately (since $\mu \geq 0$) that

$$c^T - \lambda^T A \leq 0.$$

Moreover, by postmultiplying the equality $c^T - \lambda^T A + \mu^T = 0$ with x, and using item ii of Theorem 5.14, we obtain

$$(c^T - \lambda^T A)x = 0.$$

$\Leftarrow$ From the first condition i., it follows that there exists a vector $\mu^T \geq 0$ such that

$$c^T - \lambda^T A + \mu^T = 0.$$

Postmultiplying this equality by x yields $(c^T - \lambda^T A + \mu^T)x = 0$. Since $(c^T - \lambda^T A)x = 0$, it follows that $\mu^T x = 0$. Because, both μ and x are nonnegative, it follows that $\mu_i x_i = 0$ for all i.

EXERCISE 5.25 In this exercise you are asked to show that the function values of an affine function monotonically increase if one evaluates them on a line.

a. Show that one can parameterize a line l in $\mathbb{R}^n$ as: $l(t) = u + tv$, where $u,\ v \in \mathbb{R}^n$ and $t \in \mathbb{R}$.

b. Show, using item a., that the function values of the affine function $f(x) = a^T x + b$, with $a \in \mathbb{R}^n,\ b \in \mathbb{R}$, either remain constant, strictly increase or strictly decrease if one considers them on the line $l(t)$, where t increases.

ANSWER 5.25

a. Consider a line l in $\mathbb{R}^n$, and let c and d be two points on this line.
Define $u := c^T$ and $v := d^T - c^T$. Then u is a vector that points to the line and v is a vector parallel to the line. Hence, every point on the line can be expressed as $u + tv$ for some $t \in \mathbb{R}$.

b. Any point x on the line l can be written, as shown in item a., as

$$x(t) = u + tv.$$

Hence

$$f(x(t)) = a^T x(t) + b = a^T(u + tv) + b = a^T u + b + a^T v t.$$

Thus, f is a linear function of t. Consequently, depending on the sign of $a^T v$, the values of f either remain constant, increase strictly monotonically, or decrease strictly monotonically as t increases.

EXERCISE 5.26 *(Proof items a-c of Corollary 5.15)*
Consider the LP problem (5.2). Show next properties hold.

a. Show, using Theorem 5.14, that a solution cannot occur at the interior of the constraint set. Therefore, if the LP has a solution, it is attained at the border of the constraint set.

b. If the LP problem is solved at $\bar{x}$ and $\hat{x}$ then it is also solved at any point on the line segment between $\bar{x}$ and $\hat{x}$. Hint: consider the parameterization of an arbitrary point on the line segment and show this point belongs to the constraint set and the objective value attained at this point is everywhere the same. More general, if the LP problem has a solution at n different points, then the LP problem is also solved at any point in the convex hull of these points.

c. The LP problem has either no, a unique or an infinite number of solutions.

ANSWER 5.26

a. Assume that the LP problem has an interior solution. Then, by items ii. and iii. of Theorem 5.14, we have $\lambda = \mu = 0$. Substituting this into item i. of the theorem then leads to a contradiction (note that, by assumption, $c \neq 0$).

b. Assume that all conditions in Theorem 5.14 are satisfied at $(\bar{x}, \bar{\lambda}, \bar{\mu})$ and $(\hat{x}, \hat{\lambda}, \hat{\mu})$. Let $\tilde{x} := \tau\bar{x} + (1-\tau)\hat{x}$, with $\tau \in (0,1)$ be an arbitrary point on the line segment connecting $\bar{x}$ and $\hat{x}$. Then

$$A\tilde{x} = \tau A\bar{x} + (1-\tau)A\hat{x} \leq \tau b + (1-\tau)b = b.$$

Moreover,

$$\tilde{x} = \tau\bar{x} + (1-\tau)\hat{x} \geq 0.$$

Finally,

$$c^T\tilde{x} = c^T(\tau\bar{x} + (1-\tau)\hat{x}) = \tau c^T\bar{x} + (1-\tau)c^T\hat{x}.$$

By assumption, $c^T\bar{x} = c^T\hat{x}$, and therefore $c^T\tilde{x} = c^T\bar{x}$.

It follows that all points on the line segment lie within the feasible set, and the objective function takes the same maximal value at each of these points. Hence, every point on the line segment is also an optimal solution to the LP problem.

The general case follows by the same argument.

c. This follows immediately from item b.

EXERCISE 5.27 *(Proof item d of Corollary 5.15)*

In this exercise you are asked to provide a proof for the last statement d. made in Corollary 5.15. That is, in case the LP problem has a solution there will be at least one corner point where this solution is attained.

To prove this assertion assume $\bar{x}$, not a corner point, is a solution for the LP problem.

a. Assume $n-1$ independent constraints are binding at $\bar{x}$. Show that the line defined by these $n-1$ binding constraints intersects another constraint $g_n(x)$ which is independent from these $n-1$ binding constraints.

b. Use Exercise 5.25 to show that item a. implies that there is a corner point where the LP attains its maximum.

c. Prove statement d. in Corollary 5.15 by proceeding by induction along the lines of items a. and b. above.

ANSWER 5.27

a. All points simultaneously satisfying the $n-1$ independent binding constraints form a line $l(x)$ in $\mathbb{R}^n$. Since every line in $\mathbb{R}^n$ intersects at least one of the hyperplanes $x_i = 0$ for $i = 1, \cdots, n$, it follows that $l(x)$ intersects one of these binding constraints, say, $x_1 = 0$, at a point $\tilde{x}$.

If $\tilde{x}$ does not lie in the feasible set, then $l(x)$ must have crossed the boundary of the feasible set at an earlier point.; in other words, another constraint became active beforehand. Consequently, $l(x)$ intersects the boundary of the feasible set at some corner point c.

b. Since $\bar{x}$ is, by assumption, a point on $l(x)$ that is not a corner point of the feasible set, it follows from Exercise 5.25 that the objective function must be constant along this line. Otherwise this would

contradict the fact that the objective function attains a maximum at $\bar{x}$ over the feasible set. In particular, it then follows that the objective function also attains this maximum value at the corner point c.

c. If $\bar{x}$ lies on a boundary of the feasible set where exactly i constraints, with $0 < i < n-1$, are binding, it follows by a similar argument that the objective function must attain its maximum at all points on this boundary. Likewise, as before, this boundary set must intersect another constraint, meaning that the maximum is also attained on the restricted set where $i+1$ constraints are active. By applying induction, the final result then follows.

EXERCISE 5.28 *(Max-Min (in)equality and saddle-point existence)*

Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ and $f: X \times Y \to \mathbb{R}$. Assume all maxima and minima mentioned below exist (otherwise one may consider suprema and infima in items a. and b.).

a. Show that $\underline{f} := \max_{x\in X} \min_{y\in Y} f(x,y) \leq \min_{y\in Y} \max_{x\in X} f(x,y) =: \bar{f}$.

b. Let $\bar{x} \in \arg\max_{x\in X} \min_{y\in Y} f(x,y)$ and $\bar{y} \in \arg\min_{y\in Y} \max_{x\in X} f(x,y)$. Show that for all $x \in X$ $f(x,\bar{y}) \leq f(\bar{x},\bar{y}) \leq f(\bar{x},y)$ for all $y \in Y$ if and only if $\underline{f} = \bar{f}$. That is, f has a saddle-point equilibrium iff. a max-min equality holds.

c. Show that if $f(x,y) \in C^2$ and $f(x,y)$ is strictly convex in one variable and strictly concave in the other variable the max-min equality holds.

ANSWER 5.28

a. For a fixed $x \in X$, define $L(x) := \min_{y\in Y} f(x,y)$. Then, for all $y \in Y$, we have $L(x) \leq f(x,y)$. Consequently, for all $y \in Y$,

$$\max_{x\in X} L(x) \leq \max_{x\in X} f(x,y).$$

In other words, for all $y \in Y$, $\underline{f} \leq \max_{x\in X} f(x,y)$. In particular, this inequality holds for the value of y that provides the minimum on the right-hand side, i.e. the y that minimizes $\max_{x\in X} f(x,y)$. Therefore, we conclude $\underline{f} \leq \bar{f}$.

b. " $\Rightarrow$ " Assume that $(\bar{x},\bar{y})$ provides a saddle-point as stated. Then,

$$\bar{f} = \min_{y\in Y} \max_{x\in X} f(x,y) \leq \max_{x\in X} f(x,\bar{y}) \leq f(\bar{x},\bar{y}) \leq \min_{y\in Y} f(\bar{x},y) = \max_{x\in X} \min_{y\in Y} f(x,y) = \underline{f}.$$

Since, by item a., we always have $\bar{f} \geq \underline{f}$, it follows that $\bar{f} = \underline{f}$.

" $\Leftarrow$ " Let $\bar{x}$ and $\bar{y}$ be as stated. Following the argument in item a., we have

$$\begin{aligned} \underline{f} &= \max_{x\in X} \min_{y\in Y} f(x,y) = \min_{y\in Y} f(\bar{x},y) \\ &\leq f(\bar{x},\bar{y}) \leq \max_{x\in X} f(x,\bar{y}) = \min_{y\in Y} \max_{x\in X} f(x,y) = \bar{f}. \end{aligned}$$

By assumption, equality holds, so we conclude that

$$\min_{y\in Y} f(\bar{x},y) = f(\bar{x},\bar{y}) = \max_{x\in X} f(x,\bar{y}).$$

From this, it follows that for all $y \in Y$, $f(\bar{x},y) \geq f(\bar{x},\bar{y}) \geq f(x,\bar{y})$ for all $x \in X$.

c. Since by assumption, the minimum of $f(x,y)$ exists for all x, it follows that for a fixed x, this minimum satisfies $\frac{\partial f}{\partial y} = 0$. Let $\bar{y}(x)$ denote this minimum value. By assumption, the maximum of $f(x,\bar{y}(x))$ also exists. Therefore,

$$\frac{df(x,\bar{y}(x))}{dx} = 0.$$

Note that

$$\frac{df(x,\bar{y}(x))}{dx} = \frac{\partial f}{\partial x}(x,\bar{y}(x)) + \frac{\partial f}{\partial y}(x,\bar{y}(x))\bar{y}^{'}(x).$$

From this, we conclude that if $\max_{x\in X} \min_{y\in Y} f(x,y)$ is attained at some point $(\bar{x},\bar{y})$, then both

$$\frac{\partial f}{\partial y}(\bar{x},\bar{y}) = 0 \text{ and } \frac{\partial f}{\partial y}(\bar{x},\bar{y}) = 0.$$

Furthermore, due to the strict concavity/convexity assumptions on f, the Hessian $Hf(x,y)$ is strictly indefinite ($f^{"}_{xx}$ is negative definite and $f^{"}_{yy}$ is positive definite, or vice versa). This implies, in particular, that the stationary point $(\bar{x},\bar{y})$ is a saddle-point of f. Hence, by item b., the max-min equality holds.

EXERCISE 5.29 *(Duality gap Lagrange dual function)*
Assume Problem (4.8) has a solution $f(x^*) = p$. Consider for each $\lambda \geq 0$ and $x \in C$ the with problem (4.8) associated Lagrange function $L(x,\lambda) := f(x) - \lambda^T(g(x) - b)$. Let for each fixed $\lambda \geq 0$ $L_D(\lambda) := \sup_{x\in C} L(x,\lambda)$. Let $d := \inf_{\lambda\geq 0} L_D(\lambda)$.
Show that $p \leq d$. Hint: use Exercise 5.28.

ANSWER 5.29
For a fixed $\lambda \geq 0$, define $L(x,\lambda) := f(x) - \lambda^T(g(x) - b)$.
First, recall from Remark 5.18 that

$$p = \max_{x\in C} \min_{\lambda\geq 0} L(x,\lambda).$$

By Exercise 5.28, we have

$$\max_{x\in C} \min_{\lambda\geq 0} L(x,\lambda) \leq \min_{\lambda\geq 0} \max_{x\in C} L(x,\lambda),$$

from which it follows that

$$p \leq \min_{\lambda\geq 0} L_D(\lambda) = d.$$

EXERCISE 5.30 *(Convexity Lagrange dual function)*
Consider for each $\lambda \geq 0$ and $x \in C$ the with problem (4.5) associated Lagrange function $L(x, \lambda) := f(x) - \lambda(g(x) - b)$. Define for each fixed $\lambda \geq 0$, $L_D(\lambda) := \sup_{x \in C} L(x, \lambda)$. In this exercise you are asked to prove that this Lagrange dual function, $L_D(\lambda)$, is convex.
a. Show that $\sup_{x \in C} f(x) + g(x) \geq \sup_{x \in C} f(x) + \sup_{x \in \mathcal{D}} g(x)$.
b. Show that $L_D(\tau\lambda_1 + (1-\tau)\lambda_2) \geq \tau L_D(\lambda_1) + (1-\tau)L_D(\lambda_2)$.

ANSWER 5.30
a. Let $\bar{f} := \sup_{x \in C} f(x)$ and $\bar{g} := \sup_{x \in C} g(x)$. Then, for all $x \in C$,

$$\bar{f} + \bar{g} \geq f(x) + g(x).$$

In particular, it follows that $\bar{f} + \bar{g} \geq \sup_{x \in C} f(x) + g(x)$.
b. Using item a. we have:

$$\begin{aligned}
L_D(\tau\lambda_1 + (1-\tau)\lambda_2) &= \sup_{x \in C} f(x) - (\tau\lambda_1 + (1-\tau)\lambda_2)^T (g(x-b)) \\
&= \sup_{x \in C} \tau(f(x) - \lambda_1^T(g(x) - b)) + (1-\tau)(f(x) - \lambda_2^T(g(x) - b)) \\
&\geq \sup_{x \in C} \tau(f(x) - \lambda_1^T(g(x) - b)) + \sup_{x \in C}(1-\tau)(f(x) - \lambda_2^T(g(x) - b)) \\
&= \tau \sup_{x \in C} f(x) - \lambda_1^T(g(x) - b) + (1-\tau) \sup_{x \in C} f(x) - \lambda_2^T(g(x) - b) \\
&= \tau L_D(\lambda_1) + (1-\tau)L_D(\lambda_2).
\end{aligned}$$

EXERCISE 5.31 *(Lagrange dual function LP problem)*
Consider the with the standard LP problem associated Lagrange function $L(x, \lambda, \mu) := c^T x - \lambda^T(Ax - b) + \mu^T x$. Consider for fixed $\lambda \geq 0$ and $\mu \geq 0$ the corresponding Lagrange dual function $L_D(\lambda, \mu) := \sup_{x \in \mathbb{R}^n} L(x, \lambda, \mu)$.
a. Show that $L_D(\lambda, \mu) = \begin{cases} b^T\lambda \text{ if } A^T\lambda - c - \mu = 0 \\ \quad \infty \text{ elsewhere} \end{cases}$
b. Show that $\inf_{\lambda \geq 0, \mu \geq 0} L_D(\lambda, \mu) = \inf_{\lambda \geq 0} b^T\lambda$ subject to $A^T\lambda - c \geq 0$.

ANSWER 5.31
a. Define

$$L(x, \lambda, \mu) := c^T x - \lambda^T(Ax - b) + \mu^T x = b^T\lambda - x^T(A^T\lambda - c - \mu).$$

This is a linear function of x, so its supremum is ∞, unless $A^T\lambda - c - \mu = 0$. In that case $L(x, \lambda, \mu)$ reduces to the constant function $b^T\lambda$.
b. Since $\mu \geq 0$, the stated result follows immediately from item a.

EXERCISE 5.32 *(A Quadratic programming problem with linear equality constraints)*
Consider the problem to minimize $f(x) = x^T Qx$, where Q is a positive definite matrix, subject to the linear constraints $Ax = b$. Denote its solution by p.

a. Reformulate the problem into the standard form and show that the with this problem associated Lagrange function is $L(x, \lambda, \mu) = -x^T Qx - \lambda^T(Ax - b) - \mu^T(b - Ax)$ with $\lambda \geq 0$ and $\mu \geq 0$.
b. Determine for fixed $\lambda \geq 0$ and $\mu \geq 0$ the with the standard form corresponding Lagrange dual function $L_D(\lambda, \mu) := \sup_{x \in \mathbb{R}^n} L(x, \lambda, \mu)$. Show that the image of $L_D(\lambda, \mu)$ coincides with the image of $g(y) := \frac{1}{4} y^T AQA^T y + y^T b$.
c. Assume matrix A has full row rank. Show that $p \geq d$, where $d = b^T (AQA^T)^{-1} b$.

ANSWER 5.32

a. The problem is equivalent to solving the maximization problem $-x^T QX$ subject to the inequality constraints $Ax - b \leq 0$ and $b - Ax \leq 0$. If the minimization problem has solution p, then the solution to the corresponding maximization problem is $-p$.
   The Lagrange function associated with this maximization problem is

$$L(x, \lambda, \mu) = -x^T Qx - \lambda^T(Ax - b) - \mu^T(b - Ax),$$

   with $\lambda \geq 0$ and $\mu \geq 0$.
b. Define

$$L_D(\lambda, \mu) = \sup_{x \in \mathbb{R}^n} L(x, \lambda, \mu).$$

   Notice that $L(x, \lambda, \mu)$ is a quadratic function, and since Q is positive definite, the supremum is actually attained as a maximum at the stationary point of L. Since $L_x^{'} = -2Qx - A^T(\lambda - \mu)$, the stationary point is $x^* = -\frac{1}{2} Q^{-1} A^T (\lambda - \mu)$.
   Thus,

$$L_D(\lambda, \mu) = L(x^*, \lambda, \mu) = -\frac{1}{4}(\lambda - \mu)^T AQ^{-1} A^T (\lambda - \mu) + \frac{1}{2}(\lambda - \mu)^T AQ^{-1} A^T + (\lambda - \mu)^T b.$$

   Equivalently, letting $y := \lambda - \mu$, we have

$$L_D(\lambda, \mu) = g(y),\ y \in \mathbb{R}^n.$$

c. According to Remark 5.18, the maximum of $-x^T Qx$ subject to $Ax = b$ is bounded above by the minimum of $g(y)$. Since A has full row rank, the matrix AQA^T is invertible (and positive definite). Differentiating $g(y)$ gives

$$\frac{dg(y)}{dy} = \frac{1}{2} AQA^T y + b.$$

Hence, g attains its minimum at $y^* = -2(AQA^T)^{-1}b$ with $g(y^*) = -b^T(AQA^T)^{-1}b$.
It follows that the negative of the maximum of $-x^TQx$ subject to $Ax = b$ is at least $b^T(AQA^T)^{-1}b$, that is, $p \geq b^T(AQA^T)^{-1}b$.

EXERCISE 5.33 *(Quadratic programming problem)*
The quadratic programming problem (qpp) is concisely stated as follows: find $\min_{x\in\mathbb{R}^n} \frac{1}{2}x^TQx+q^Tx$ subject to the linear constraints $Ax + b \geq 0$ and $x \geq 0$, where $x, b, q \in \mathbb{R}^n$, $Q \in \mathbb{R}^{n\times n}$ is a symmetric matrix, and the constraint set is nonempty.

a. Show that the qpp has a solution if the constraint set is bounded.
b. Give an example of a qpp that has no solution.
c. Show that if Q is positive semi-definite any solution of the with this problem corresponding KKT conditions provides a global minimum.
d. Show that if the qpp has a solution, there exist $x \geq 0,\ \lambda_1 \geq 0$ and $\lambda_2 \geq 0$ that satisfy the set of (in)equalities:
   1. $-Qx - q + A^T\lambda_1 + \lambda_2 = 0$;
   2. $x \geq 0,\ \lambda_1 \geq 0,\ \lambda_2 \geq 0$ and $Ax + b \geq 0$;
   3. $\lambda_1^T(Ax + b) = 0$ and $\lambda_2^Tx = 0$.
e. Show that the advertized necessary conditions in item d. can be reformulated as that there exists a vector $z \geq 0$ such that with $w := Cz + d \geq 0$, $w^Tz = 0$, with $C = \begin{bmatrix} Q & -A^T \\ A & 0 \end{bmatrix}$ and $d = \begin{bmatrix} q \\ b \end{bmatrix}$.

ANSWER 5.33

a. Since the feasible set is a polyhedron, it is a convex, closed, and, by assumption, bounded set. Because the objective function is continuous, Weierstrass' theorem guarantees that the problem has a solution.
b. Choose, e.g., $Q = -1,\ A = 1$ and $q = b = 0$.
c. If Q is positive semi-definite, the objective function is convex. Therefore, by Theorem 5.13, the KKT conditions are both necessary and sufficient for existence of a global minimum.
c. Let $L(x, \lambda_1, \lambda_2) := -\frac{1}{2}x^TQx - q^Tx + \lambda_1^T(Ax + b) + \lambda_2^Tx$. The necessary conditions that result from the KKT conditions are then:
   1. $-Qx - q + A^T\lambda_1 + \lambda_2 = 0$;
   2. $x \geq 0,\ \lambda_1 \geq 0,\ \lambda_2 \geq 0$ and $Ax + b \geq 0$;
   3. $\lambda_1^T(Ax + b) = 0$ and $\lambda_2^Tx = 0$.
d. Let $z^T := [x^T\lambda_1^T]$ and $w^T := [(Qx + q - A^T\lambda_1)^T\ (Ax + b)^T]$.

EXERCISE 5.34 *(Equivalence linear complementarity problem and quadratic programming problem)*
Consider the so-called *linear complementarity problem* (lcp) to find a vector $x \geq 0$ satisfying $Ax+b \geq 0$ and $x^T(Ax + b) = 0$ (i.e.: either $x_i = 0$ or $(Ax+b)_i = 0$ for all i) (where $x, b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$).

Also, consider the with this problem associated so-called *quadratic programming problem* (qpp): $\min_{x\in\mathbb{R}^n} x^T(Ax+b)$ subject to the linear constraints $Ax+b\geq 0$ and $x\geq 0$.

a. Show that the solution of the linear complementarity problem (lcp) exists if and only if the quadratic programming problem (qpp) has a solution and the minimum value is 0.
b. Show that for any square matrix A, $x^TAx=\frac{1}{2}x^T(A+A^T)x$. (Hint: consider the transpose of the scalar x^TAx).
c. Formulate sufficient conditions on matrix A under which the problem has a solution if $b\geq 0$.

ANSWER 5.34

a. Since in the LCP we are seeking a solution for which both x and $Ax+b$ are nonnegative, it follows that the function $f(x)=x^T(Ax+b)\geq 0$. Moreover, if a solution x^* exists, $f(x^*)=0$ due to the complementarity condition.
   Conversely, if the quadratic optimization problem admits a solution $x^*\geq 0$ such that $Ax^*+b\geq 0$ and $f(x^*)=0$, it follows that x^* satisfies the complementarity condition. Therefore, x^* is a solution to the LCP.
b. Since x^TAx is a scalar, $x^TAx=(x^TAx)^T=x^TA^Tx$. From this, the stated equality follows immediately.
c. From item b., we observe that, w.l.o.g., we can consider the problem of minimizing

$$f(x)=\frac{1}{2}x^T(A+A^T)x+x^Tb.$$

   By item a., we need to determine the conditions under which the minimum value of this problem is zero. Note that $f(0)=0$. Therefore, for a solution to exist, it suffices that the matrix A satisfies

$$x^T(A+A^T)x\geq 0 \text{ for all } x\geq 0.$$

   For example, this property holds if $A+A^T$ is positive semi-definite. It also holds if $A+A^T$ has only nonnegative entries.

Problems with an economic content

EXERCISE 5.35 *(optimal consumption of products)*

Consider the optimal consumption x and y of two products. Assume minimal consumption of product i is $c_i,\ i=1,2$. The price of product i is $p_i, i=1,2$, and the maximum the consumer can spend is b which exceeds $p_1c_1+p_2c_2$. Assume utility derived from consumption of one unit of product i is $s_i,\ i=1,2$.

a. Formulate the optimization problem. Show that it is optimal to consume the minimal amount of product 1 and as much as possible from product 2 if $\frac{s_2}{p_2}>\frac{s_1}{p_1}$ and the other way around if this last inequality reverses.

b. Assume next that the maximum available amount of product i is m_i and that the excess of the products not consumed must be recycled at a cost of d_i (with $d_i < p_i$) per unit of product, $i = 1, 2$. Assume the disutility derived of recycling a unit of product i is t_i.
Reformulate the optimization problem. Show that the problem has only a solution iff. the budget b exceeds the cost of buying the minimal amount of the products (s)he needs and the cost of recycling the rest of the products. Solve the problem under this premise. Show that again the optimal solution is to spend as much as possible on one product and if the maximum quotum of this product is reached, spend the rest of the budget on the other product. Introduce the "shadow" prices $\tilde{p}_i := p_i - d_i$ and "shadow" preferences $\tilde{s}_i := s_i + t_i,\ i = 1, 2$. Show that again the sign of $\frac{\tilde{s}_1}{\tilde{p}_1} - \frac{\tilde{s}_2}{\tilde{p}_2}$ determines the location of the maximum.

ANSWER 5.35

a. The optimization problem is to maximize

$$U(x, y) := s_1 x + s_2 y \text{ subject to } x \geq c_1,\ y \geq c_2 \text{ and } p_1 x + p_2 y \leq b.$$

Note that this is a LP problem defined over a compact domain.
The corner points of the feasible set are (c_1, c_2), $(c_1, \frac{b - p_1 c_1}{p_2})$ and $(\frac{b - p_2 c_2}{p_1}, c_2)$. The maximum utility is uniquely attained at $(c_1, \frac{b - p_1 c_1}{p_2})$ iff.

$$s_1 c_1 + s_2 \frac{b - p_1 c_1}{p_2} > s_1 \frac{b - p_2 c_2}{p_1} + s_2 c_2,$$

or, equivalently,

$$(b - p_1 c_1 - p_2 c_2)(\frac{s_2}{p_2} - \frac{s_1}{p_1}) > 0.$$

If $\frac{s_2}{p_2} < \frac{s_1}{p_1}$ it is optimal to consume the minimal required amount of product 2 and spend the remainder of the budget on product 1. If $\frac{s_2}{p_2} = \frac{s_1}{p_1}$, any point on the budget line $p_1 x + p_2 y \leq b$ between the two points $(c_1, \frac{b - p_1 c_1}{p_2})$ and $(\frac{b - p_2 c_2}{p_1}, c_2)$ provides the maximum attainable utility.

b. The optimization problem is to maximize

$$U(x, y) := s_1 x + s_2 y - t_1(m_1 - x) - t_2(m_2 - x) \text{ subject to}$$

$$c_1 \leq x \leq m_1,\ c_2 \leq y \leq m_2,\ p_1 x + p_2 y + d_1(m_1 - x) + d_2(m_2 - x) \leq b.$$

Again, this is a LP problem over a compact domain.
To ensure the feasible set is not empty, it is necessary and sufficient that the corner point (c_1, c_2) satisfies the constraint $p_1 x + p_2 y + d_1(m_1 - x) + d_2(m_2 - x) \leq b$. That is

$$p_1 c_1 + p_2 c_2 + d_1(m_1 - c_1) + d_2(m_2 - c_2) \leq b.$$

Introduce the "shadow" prices $\tilde{p}_i := p_i - d_i,\ i = 1, 2$, "shadow" budget $\tilde{b} := b - d_1 m_1 - d_2 m_2$ and "shadow" preferences $\tilde{s}_i := s_i + t_i,\ i = 1, 2$. Next, we distinguish between four cases.

1. if $m_1 \geq \frac{\tilde{b} - \tilde{p}_2 c_2}{\tilde{p}_1}$ and $m_2 \geq \frac{\tilde{b} - \tilde{p}_1 c_1}{\tilde{p}_2}$ we have next 3 corner points (see item a.): (c_1, c_2), $(c_1, \frac{\tilde{b} - \tilde{p}_1 c_1}{\tilde{p}_2})$ and $(\frac{\tilde{b} - \tilde{p}_2 c_2}{\tilde{p}_1}, c_2)$.
2. if $m_1 \geq \frac{\tilde{b} - \tilde{p}_2 c_2}{\tilde{p}_1}$ and $m_2 \leq \frac{\tilde{b} - \tilde{p}_1 c_1}{\tilde{p}_2}$ we have next 4 corner points: (c_1, c_2), (c_1, m_2), $(c_1, \frac{\tilde{b} - \tilde{p}_1 c_1}{\tilde{p}_2})$ and $(\frac{\tilde{b} - \tilde{p}_2 m_2}{\tilde{p}_1}, m_2)$.
3. if $m_1 \leq \frac{\tilde{b} - \tilde{p}_2 c_2}{\tilde{p}_1}$ and $m_2 \geq \frac{\tilde{b} - \tilde{p}_1 c_1}{\tilde{p}_2}$ we have next 4 corner points: (c_1, c_2), (m_1, c_2), $(m_1, \frac{\tilde{b} - \tilde{p}_1 m_1}{\tilde{p}_2})$ and $(\frac{\tilde{b} - \tilde{p}_2 c_2}{\tilde{p}_1}, c_2)$.
4. if $m_1 \leq \frac{\tilde{b} - \tilde{p}_2 c_2}{\tilde{p}_1}$ and $m_2 \leq \frac{\tilde{b} - \tilde{p}_1 c_1}{\tilde{p}_2}$ we have next 5 corner points: (c_1, c_2), (m_1, c_2), (c_1, m_2), $(m_1, \frac{\tilde{b} - \tilde{p}_1 m_1}{\tilde{p}_2})$ and $(\frac{\tilde{b} - \tilde{p}_2 m_2}{\tilde{p}_1}, m_2)$.

In all cases the maximum is attained at either one of the two corners on the budgetline or, if the objective value at both these points coincides, the whole line segment between these two corner points on the budgetline. So, in the first case we arrive at the same conclusions as in item a., but then with price, preference and budget variables replaced by their "shadow" versions. By a simple inspection it follows that also in the other cases the sign of $\frac{\tilde{s}_1}{\tilde{p}_1} - \frac{\tilde{s}_2}{\tilde{p}_2}$ determines the location of the maximum.

EXERCISE 5.36 *(optimal portfolio selection)*

Consider an investor who has to make a decision on which fraction ω_b from his capital X he should invest in bonds, the fraction ω_s he should invest in stocks and the remaining fraction $1 - \omega_b - \omega_s$ he should put on his savings account. For this portfolio the expected returns μ are $(1 - \omega_b - \omega_s)(1 + r)X + \omega_b(1 + r_b)X + \omega_s(1 + r_s)X$ and the associated variance $\sigma^2 = (1 - \omega_b - \omega_s)^2\sigma_0^2 + \omega_b^2\sigma_b^2 + \omega_s^2\sigma_s^2$. Here r (σ_0^2), r_b (σ_b^2) and r_s (σ_s^2) are the expected returns (variance) for his saving accounts, bonds and stocks, respectively. Assume $r \leq r_b \leq r_s$ and $0 \leq \sigma_0^2 < \sigma_b^2 < \sigma_s^2$.

To make a choice in this decision problem the investor uses a utility function $U(\mu, \sigma) = \mu - \alpha\sigma^2$, where $\alpha \geq 0$ expresses his attitude towards risk, measured by σ^2. His optimal portfolio is determined as the fractions that maximize this utility. However, he does not want to invest more than a fraction $\bar{b}$ into bonds, a fraction $\bar{s}$ into stocks (where $\bar{b} \geq \bar{s}$) and the sum of these fractions should be at most $\bar{t}$, where $\bar{t} < \bar{b} + \bar{s}$.

a. Formulate the optimization problem.

b. Show this is a convex optimization problem with linear constraints.

c. Show that the solution is attained at the interior of the constraint set iff. with $d := (\sigma_0^2 + \sigma_s^2)(\sigma_0^2 + \sigma_b^2) - \sigma_0^4$ next conditions apply:

   i) $0 < 2\alpha\sigma_0^2\sigma_s^2 + ((r_b - r)\sigma_s^2 - (r_s - r_b)\sigma_0^2)X < 2\alpha d\bar{b}$;

   ii) $2\alpha\sigma_0^2\sigma_b^2 + ((r_s - r)\sigma_b^2 + (r_s - r_b)\sigma_0^2)X < 2\alpha d\bar{s}$ and

   iii) $2\alpha\sigma_0^2(\sigma_b^2 + \sigma_s^2) + ((r_b - r)\sigma_s^2 + (r_s - r)\sigma_b^2)X < 2\alpha d\bar{t}$.

   Maximum utility is then attained at:

   $(\omega_b, \omega_s) = \frac{1}{2\alpha d}(2\alpha\sigma_0^2\sigma_s^2 + ((r_b - r)\sigma_s^2 - (r_s - r_b)\sigma_0^2)X, 2\alpha\sigma_0^2\sigma_b^2 + ((r_s - r)\sigma_b^2 + (r_s - r_b)\sigma_0^2)X)$.

d. Assume $0 < r < r_b$. Show that if the maximum occurs on the constraint set then it is at either one of next locations:
$(\omega_b, \omega_s) = (\bar{b}, \frac{1}{2\alpha(\sigma_0^2+\sigma_s^2)}((r_s - r)X + 2\alpha\sigma_0^2(1-\bar{b})))$ provided $\omega_s \leq \bar{t} - \bar{b}$ and $(r_b - r)X - 2\alpha\bar{b}\sigma_b^2 + 2\alpha\sigma_0^2(1 - \bar{b} - \omega_s) \geq 0$.
$(\omega_b, \omega_s) = (\frac{1}{2\alpha(\sigma_0^2+\sigma_b^2)}((r_b - r)X + 2\alpha\sigma_0^2(1-\bar{s})), \bar{s})$ provided $\omega_b \leq \bar{t} - \bar{s}$ and $(r_s - r)X - 2\alpha\bar{s}\sigma_s^2 + 2\alpha\sigma_0^2(1 - \bar{s} - \omega_b) \geq 0$.
$(\omega_b, \omega_s) = \frac{1}{2\alpha(\sigma_b^2+\sigma_s^2)}((r_b - r_s)X + 2\alpha\sigma_s^2\bar{t}, (r_s - r_b)X + 2\alpha\sigma_b^2\bar{t})$ provided $\omega_b \leq \bar{b}$, $\omega_s \leq \bar{s}$ and $(r_b - r)X + 2\alpha\sigma_0^2(1-\bar{t}) - 2\alpha\sigma_b^2\omega_b \geq 0$.

e. Assume $\sigma_0 = 0$. Show that at an interior solution the fractions of cash, bonds and stocks are: $1 - \frac{r_b - r}{2\alpha\sigma_b^2}X - \frac{r_s - r}{2\alpha\sigma_s^2}X$, $\frac{r_b - r}{2\alpha\sigma_b^2}X$ and $\frac{r_s - r}{2\alpha\sigma_s^2}X$.
Show that for large choices of α an interior solution occurs if $r_b > r$. Provide an interpretation of α.

f. Assume $\sigma_0 = 0$; $\bar{t} > 1$ and α satisfies next inequalities: $\alpha > \frac{1}{\bar{b}}\frac{r_b - r}{2\sigma_b^2}X$; $\alpha > \frac{1}{\bar{s}}\frac{r_s - r}{2\sigma_s^2}X$; and $\frac{1}{\bar{t}}(\frac{1}{2}\frac{r_b - r}{\sigma_b^2}X + \frac{1}{2}\frac{r_s - r}{\sigma_s^2}X) < \alpha < \frac{1}{2}\frac{r_b - r}{\sigma_b^2}X + \frac{1}{2}\frac{r_s - r}{\sigma_s^2}X$. Determine the optimal solution for the problem and interpret it.

g. Assume $\sigma = 0$. Show that no investment in bonds occurs and maximal investment in stocks iff. $r_b = r$ and the investor is risk seeking. Quantify the notion of risk seeking here.

ANSWER 5.36

a. The optimization problem is to maximize $U(\omega_b, \omega_s)$ defined by

$$(1 - \omega_b - \omega_s)(1 + r)X + \omega_b(1 + r_b)X + \omega_s(1 + r_s)X - \alpha((1 - \omega_b - \omega_s)^2\sigma_0^2 + \omega_b^2\sigma_b^2 + \omega_s^2\sigma_s^2)$$

$$\text{subject to } 0 \leq \omega_b \leq \bar{b},\ 0 \leq \omega_s \leq \bar{s} \text{ and } \omega_b + \omega_s \leq \bar{t}.$$

b. Introduce for notational convenience $\omega_0 := 1 - \omega_b - \omega_s$. Then

$$U^{'} = [-(1 + r)X + (1 + r_b)X,\ -(1 + r)X + (1 + r_s)X] - 2\alpha[-\omega_0\sigma_0^2 + \omega_b\sigma_b^2,\ -\omega_0\sigma_0^2 + \omega_s\sigma_s^2]$$

and

$$U^{"} = -2\alpha \begin{bmatrix} \sigma_0^2 + \sigma_b^2 & \sigma_0^2 \\ \sigma_0^2 & \sigma_0^2 + \sigma_s^2 \end{bmatrix}.$$

One easily verifies that the Hessian matrix $U^{"}$ is negative semi-definite. So, U is concave.

c. Let $g_1(\omega_b, \omega_s) := \omega_b - \bar{b}$; $g_2(\omega_b, \omega_s) := \omega_s - \bar{s}$; $g_3(\omega_b, \omega_s) := \omega_b + \omega_s - \bar{t}$; $g_4(\omega_b, \omega_s) := -\omega_b$ and $g_5(\omega_b, \omega_s) := -\omega_s$.
Consider the Lagrange function $L(\omega_b, \omega_s, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5)$ defined by

$$\begin{aligned}&(1 - \omega_b - \omega_s)(1 + r)X + \omega_b(1 + r_b)X + \omega_s(1 + r_s)X - \alpha((1 - \omega_b - \omega_s)^2\sigma_0^2 + \omega_b^2\sigma_b^2 + \omega_s^2\sigma_s^2)\\ &-\lambda_1(\omega_b - \bar{b}) - \lambda_2(\omega_s - \bar{s}) - \lambda_3(\omega_b + \omega_s - \bar{t}) + \lambda_4\omega_b + \lambda_5\omega_s.\end{aligned}$$

According to Theorem 5.13, the solution of this optimization problem is obtained by solving the KKT conditions:

1. $[-(1+r)X+(1+r_b)X,\ -(1+r)X+(1+r_s)X]-2\alpha[-\omega_0\sigma_0^2+\omega_b\sigma_b^2,\ -\omega_0\sigma_0^2+\omega_s\sigma_s^2]+[-\lambda_1-\lambda_3+\lambda_4,\ -\lambda_2-\lambda_3+\lambda_5]=[0,\ 0]$;
2. $\lambda_i g_i(\omega_b,\omega_s)=0,\ i=1,\cdots,5$;
3. $g_i(\omega_b,\omega_s)\leq 0$ and $\lambda_i\geq 0$.

If all constraints are non-binding, all $\lambda_i=0,\ i=1,\cdots,5$. So, provided the constraints $g_i<0$ are satisfied the solution is then obtained by solving:

$$\begin{bmatrix} \sigma_0^2+\sigma_b^2 & \sigma_0^2 \\ \sigma_0^2 & \sigma_0^2+\sigma_s^2 \end{bmatrix}\begin{bmatrix} \omega_b \\ \omega_s \end{bmatrix}=\begin{bmatrix} \sigma_0^2+\frac{r_b-r}{2\alpha}X \\ \sigma_0^2+\frac{r_s-r}{2\alpha}X \end{bmatrix}.$$

From this, we obtain directly the stated result.

d. Consider the case in which exactly one constraint is binding. Under the assumption $0<r<r_b$, it follows immediately from the first KKT condition that neither g_4 nor g_5 can be the uniquely binding constraint.

1. Assume $g_1=0$. So, $\lambda_i=0,\ i=2,3,4,5$.
   Then
   $$(\omega_b,\omega_s)=(\bar{b},\frac{1}{2\alpha(\sigma_0^2+\sigma_s^2)}((r_s-r)X+2\alpha\sigma_0^2(1-\bar{b})))$$
   is obtained provided $\omega_s<\bar{t}-\bar{b}$ and $\lambda_1=(r_b-r)X-2\alpha\bar{b}\sigma_b^2+2\alpha\sigma_0^2(1-\bar{b}-\omega_s)\geq 0$.
2. Assume $g_2=0$. So, $\lambda_i=0,\ i=1,3,4,5$.
   Then
   $$(\omega_b,\omega_s)=(\frac{1}{2\alpha(\sigma_0^2+\sigma_b^2)}((r_b-r)X+2\alpha\sigma_0^2(1-\bar{s})),\bar{s})$$
   is obtained provided $\omega_b<\bar{t}-\bar{s}$ and $\lambda_2=(r_s-r)X-2\alpha\bar{s}\sigma_s^2+2\alpha\sigma_0^2(1-\bar{s}-\omega_b)\geq 0$.
3. Assume $g_3=0$. So, $\lambda_i=0,\ i=1,2,4,5$.
   Then
   $$(\omega_b,\omega_s)=\frac{1}{2\alpha(\sigma_b^2+\sigma_s^2)}((r_b-r_s)X+2\alpha\sigma_s^2\bar{t},(r_s-r_b)X+2\alpha\sigma_b^2\bar{t})$$
   is obtained provided $\omega_b<\bar{b}$, $\omega_s<\bar{s}$ and $\lambda_3=(r_b-r)X+2\alpha\sigma_0^2(1-\bar{t})-2\alpha\sigma_b^2\omega_b\geq 0$.

The case that precisely 2 constraints are binding occurs at the corner points and yields the advertized solutions (only the two corner points on the line $\omega_b+\omega_s=\bar{t}$ yield a solution).

e. Substitution of $\sigma_0=0$ into the result of item c. yields the stated result (notice the fraction of cash is $1-\omega_b-\omega_s$).

For large values of α, we observe that most of the cash is allocated to the savings account. As α increases, the investor allocates less to bonds and stocks. In other words, a higher α indicates that the investor is more risk-averse and prefers to avoid risky positions.

f. Under the specified conditions, the internal solution is realized, resulting in a negative allocation to the savings account. That is, the investor is borrowing funds and investing them in bonds and stocks.

g. Substituting $(\omega_b, \omega_s) = (0, \bar{s})$ into the (necessary and sufficient!) KKT conditions from item c. shows that they are satisfied iff. $r_b = r$ (with $\lambda_4 = 0$) and

$$\alpha \leq \frac{1}{2}\frac{1}{\bar{s}}\frac{r_s - r}{\sigma_s^2}.$$

EXERCISE 5.37 *(production under delivery constraints)*

Consider a price taking producer who manufactures 2 products. He has a contract to supply for product i a minimum of $\bar{y}_i,\ i = 1, 2$. The maximum amount of labor and capital he can get for the total production of both goods is $\bar{L}$ and $\bar{K}$, respectively. Assume the production functions of the goods are specified by $y_i = K_i^{\alpha_i} L_i^{1-\alpha_i},\ i = 1, 2$, with $\alpha_i \in (0, 1)$. Assume the cost of labor and capital is per unit ω and r, respectively. Assume the producer wants to maximize his profit and that the price he gets for product i is $p_i,\ i = 1, 2$..

a. Formulate the optimization problem.

b. Show this is a convex optimization problem with linear constraints.

c. Show that the necessary and sufficient conditions for a maximum are obtained as the solutions of next set of equations:

   1.i) $\alpha_1 y_1(p_1 + \lambda_1) = K_1(r + \lambda_4)$; ii) $\alpha_2 y_2(p_2 + \lambda_2) = K_2(r + \lambda_4)$; iii) $(1 - \alpha_1)y_1(p_1 + \lambda_1) = L_1(\omega + \lambda_3)$; iv) $(1 - \alpha_2)y_2(p_2 + \lambda_2) = L_2(\omega + \lambda_3)$;

   2. With $g_i := \bar{y}_i - y_i,\ i = 1, 2$; $g_3 := L_1 + L_2 - \bar{L}$ and $g_4 := K_1 + K_2 - \bar{K}$, $\lambda_i g_i = 0,\ i = 1, \cdots, 4$;

   3. $\lambda_i \geq 0$ and $g_i \leq 0,\ i = 1, \cdots, 4$.

d. Show that the problem has almost never a solution at the interior of the constraint set or at a point where precisely one constraint is binding.

e. Consider the case $\alpha_1 = \alpha_2$ and $p_1 > p_2$. Show, by contradiction, that the constraint g_2 is then always binding. That is the producer will produce precisely the amount of product two he is obliged too.

f. Show that the case that both production constraints are simultaneously binding occurs iff. prices satisfy $p_1 \leq \left(\frac{r}{\alpha_1}\right)^{\alpha_1}\left(\frac{\omega}{1-\alpha_1}\right)^{1-\alpha_1}$ and $p_2 \leq \left(\frac{r}{\alpha_2}\right)^{\alpha_2}\left(\frac{\omega}{1-\alpha_2}\right)^{1-\alpha_2}$. Explain this result.

g. Assume $\alpha_1 \neq \alpha_2$. Show that the case that both production factor constraints are simultaneously binding occurs iff. the prices of the production factors satisfy $\omega \leq \left(\frac{p_2^{\alpha_1}}{p_1^{\alpha_2}}\right)^{\frac{1}{\alpha_1-\alpha_2}} \omega_L$ and $r \leq \left(\frac{p_1^{1-\alpha_2}}{p_2^{1-\alpha_1}}\right)^{\frac{1}{\alpha_1-\alpha_2}} r_K$. Here $\omega_L = \left(\frac{\alpha_2^{\alpha_1\alpha_2}(1-\alpha_2)^{\alpha_1(1-\alpha_2)}}{\alpha_1^{\alpha_1\alpha_2}(1-\alpha_1)^{(1-\alpha_1)\alpha_2}}\right)^{\frac{1}{\alpha_1-\alpha_2}}$ and $r_K = \left(\frac{\alpha_1^{\alpha_1(1-\alpha_2)}(1-\alpha_1)^{(1-\alpha_1)(1-\alpha_2)}}{\alpha_2^{(1-\alpha_1)\alpha_2}(1-\alpha_2)^{(1-\alpha_1)(1-\alpha_2)}}\right)^{\frac{1}{\alpha_1-\alpha_2}}$.
   Does this solution always make sense?

   It can be shown that both ω_L and r_K are located between zero and one, where r_K is close to zero and ω_L close to one if both α_1 and α_2 are close to zero. Explain this last result intuitively if $p_1 = p_2$.

ANSWER 5.37

a. The producer wants to maximize

$$\Pi(K_1, K_2, L_1, L_2) = p_1 y_1 + p_2 y_2 - r(K_1 + K_2) - \omega(L_1 + L_2)$$

subject to the constraints

$$y_i \geq \bar{y}_i \text{ for } i = 1, 2;\ K_1 + K_2 \leq \bar{K} \text{ and } L_1 + L_2 \leq \bar{L}.$$

b. Clearly the constraints are linear. Furthermore, with $\beta_i := \alpha_i(1 - \alpha_i)$ for $i = 1, 2$,

$$\Pi^{"} = \begin{bmatrix} -\beta_1 K_1^{\alpha_1 - 2} L_1^{-\alpha_1 - 1} \begin{bmatrix} L_1^2 & -K_1 L_1 \\ -K_1 L_1 & K_1^2 \end{bmatrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & -\beta_2 K_2^{\alpha_2 - 2} L_2^{-\alpha_2 - 1} \begin{bmatrix} L_2^2 & -K_2 L_2 \\ -K_2 L_2 & K_2^2 \end{bmatrix} \end{bmatrix}.$$

One easily verifies that this matrix is negative semi-definite. Thus, Π is a concave function.

c. Since we have a convex optimization problem, by Theorem 5.13, we obtain the solutions for this optimization problem as the solutions that solve the KKT conditions.
Consider the Lagrange function $L(K_1, K_2, L_1, L_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4)$ defined by

$$\begin{aligned} &p_1 y_1 + p_2 y_2 - r(K_1 + K_2) - \omega(L_1 + L_2) - \lambda_1(\bar{y}_1 - y_1) - \lambda_2(\bar{y}_2 - y_2) \\ &-\lambda_3(L_1 + L_2 - \bar{L}) - \lambda_4(K_1 + K_2 - \bar{K}). \end{aligned}$$

Note that $y_1^{'} = [\alpha_1 \frac{y_1}{K_1}\ (1 - \alpha_1)\frac{y_1}{L_1}\ 0\ 0]$ and $y_2^{'} = [0\ 0\ \alpha_2 \frac{y_2}{K_2}\ (1 - \alpha_2)\frac{y_2}{L_2}]$. From this, the stated conditions mentioned in the first item follow directly. The remaining conditions are standard.

d. If $\lambda_i = 0,\ i = 1, \cdots, 4$, we have from 1.i) and 1.iii) that $K_1 = \frac{\alpha_1 p_1}{r} y_1$ and $L_1 = \frac{(1-\alpha_1)p_1}{\omega} y_1$. Substituting this into the definition of $y_1 = K_1^{\alpha_1} L_1^{1-\alpha_1}$ shows that necessarily

$$y_1 = \left(\frac{\alpha_1 p_1}{r}\right)^{\alpha_1} \left(\frac{(1 - \alpha_1)p_1}{\omega}\right)^{1-\alpha_1} y_1$$

should hold. Of course this equality only holds under very special parametric conditions.
The same conclusion results if either only g_1 or g_2 is binding.
In case where only g_3 is binding, we have $\lambda_i = 0$ for $i = 1, 2, 4$. Then from 1.i) it follows that $y_1 = \frac{r}{\alpha_1 p_1}$ and from 1.ii) that $y_2 = \frac{r}{\alpha_2 p_2}$.
Substituting these expressions into 1.iii) and 1.iv) yields

$$\omega + \lambda_3 = r\frac{K_1}{L_1}\frac{1 - \alpha_1}{\alpha_1} = r\frac{K_2}{L_2}\frac{1 - \alpha_2}{\alpha_2}.$$

Using the definition of y_1 from 1.i), we also obtain $\frac{K_1}{L_1} = \left(\frac{r}{\alpha_1 p_1}\right)^{\frac{1}{\alpha_1 - 1}}$, and similarly, from 1.ii): $\frac{K_2}{L_2} = \left(\frac{r}{\alpha_2 p_2}\right)^{\frac{1}{\alpha_2 - 1}}$.
Therefore, the condition

$$\left(\frac{r}{\alpha_1 p_1}\right)^{\frac{1}{\alpha_1 - 1}} \frac{1-\alpha_1}{\alpha_1} = \left(\frac{r}{\alpha_2 p_2}\right)^{\frac{1}{\alpha_2 - 1}} \frac{1-\alpha_2}{\alpha_2}$$

must hold for this case to arise, which is typically a rather restrictive condition and therefore unlikely to occur in practice.

The case that just g_4 is binding is left to the reader.

e. Assume that g_2 is not binding. Then $\lambda_2 = 0$. Combining conditions 1.i), iii) with 1.ii), iv), we obtain

$$\frac{\alpha}{1-\alpha} = \frac{r+\lambda_4}{\omega+\lambda_3}\frac{K_2}{L_2} = \frac{r+\lambda_4}{\omega+\lambda_3}\frac{K_1}{L_1}.$$

From this, we conclude that $\frac{K_2}{L_2} = \frac{K_1}{L_1}$. Substituting this into condition 1.ii) yields

$$\alpha\left(\frac{K_2}{L_2}\right)^{\alpha-1} p_2 = \alpha\left(\frac{K_1}{L_1}\right)^{\alpha-1} p_2 = r + \lambda_4.$$

On the other hand, from condition 1.i) we have

$$\alpha\left(\frac{K_1}{L_1}\right)^{\alpha-1} (p_1 + \lambda_1) = r + \lambda_4.$$

Hence, it follows that

$$p_2 = p_1 + \lambda_1.$$

Since $\lambda_1 \geq 0$ and, by assumption, $p_2 > p_1$, this leads to a contradiction. Therefore, g_2 must be binding.

f. With $y_i = \bar{y}_i,\ i = 1, 2$, and $\lambda_3 = \lambda_4 = 0$, conditions 1. in item c. reduce to:

$$\begin{aligned} &\text{1.i) } \alpha_1\bar{y}_1(p_1+\lambda_1) = K_1 r;\ \ \text{ii) } \alpha_2\bar{y}_2(p_2+\lambda_2) = K_2 r; \\ &\text{iii) } (1-\alpha_1)\bar{y}_1(p_1+\lambda_1) = L_1\omega;\ \text{iv) } (1-\alpha_2)\bar{y}_2(p_2+\lambda_2) = L_2\omega. \end{aligned}$$

From these we immediately obtain

$$K_1 = \frac{\alpha_1\bar{y}_1(p_1+\lambda_1)}{r},\ K_2 = \frac{\alpha_2\bar{y}_2(p_2+\lambda_2)}{r},\ L_1 = \frac{(1-\alpha_1)\bar{y}_1(p_1+\lambda_1)}{\omega}, L_2 = \frac{(1-\alpha_2)\bar{y}_2(p_2+\lambda_2)}{\omega}.$$

Substituting these relationships into the definition of y_i yields

$$\bar{y}_i = \left(\frac{\alpha_i}{r}\right)^{\alpha_i}\left(\frac{1-\alpha_i}{\omega}\right)^{1-\alpha_i}(p_i+\lambda_i)\bar{y}_i,$$

from which it follows that

$$\lambda_i = \left(\frac{r}{\alpha_i}\right)^{\alpha_i} \left(\frac{\omega}{1-\alpha_i}\right)^{1-\alpha_i} - p_i,\ i = 1, 2.$$

Since $\lambda_i \geq 0,\ i = 1, 2$, the stated result follows.

If prices are very low, the producer will only supply the minimum required quantities. Depending on whether production remains profitable under these conditions, the producer may even decide to shut down operations.

g. With $L_1 + L_2 = \bar{L},\ K_1 + K_2 = \bar{K}$ and $\lambda_1 = \lambda_2 = 0$, conditions 1. in item c. reduce to:

$$\text{1.i) } \alpha_1 y_1 p_1 = K_1(r + \lambda_4);\ \text{ii) } \alpha_2 y_2 p_2 = K_2(r + \lambda_4);$$
$$\text{iii) } (1-\alpha_1) y_1 p_1 = L_1(\omega + \lambda_3);\ \text{iv) } (1-\alpha_2) y_2 p_2 = L_2(\omega + \lambda_3).$$

From these, we obtain

$$K_1 = \frac{\alpha_1}{r + \lambda_4} p_1 y_1,\ K_2 = \frac{\alpha_2}{r + \lambda_4} p_2 y_2,\ L_1 = \frac{1-\alpha_1}{\omega + \lambda_3} p_1 y_1, L_2 = \frac{1-\alpha_2}{\omega + \lambda_3} p_2 y_2.$$

Substituting into the definition of y_i gives

$$y_i = \left(\frac{\alpha_i}{r + \lambda_4}\right)^{\alpha_i} \left(\frac{1-\alpha_i}{\omega + \lambda_3}\right)^{1-\alpha_i} p_i y_i,\ i = 1, 2,$$

which implies

$$\left(\frac{\alpha_i}{r + \lambda_4}\right)^{\alpha_i} \left(\frac{1-\alpha_i}{\omega + \lambda_3}\right)^{1-\alpha_i} p_i = 1,\ i = 1, 2.$$

These yield two equations in the unknowns $r + \lambda_4$ and $\omega + \lambda_3$. Solving them (after some algebra) gives

$$r + \lambda_4 = \left(\frac{p_1^{1-\alpha_2}}{p_2^{1-\alpha_1}} \frac{\alpha_1^{\alpha_1(1-\alpha_2)} (1-\alpha_1)^{(1-\alpha_1)(1-\alpha_2)}}{\alpha_2^{(1-\alpha_1)\alpha_2} (1-\alpha_2)^{(1-\alpha_1)(1-\alpha_2)}}\right)^{\frac{1}{\alpha_1 - \alpha_2}}$$

and

$$\omega + \lambda_3 = \left(\frac{p_2^{\alpha_1}}{p_1^{\alpha_2}} \frac{\alpha_2^{\alpha_1\alpha_2} (1-\alpha_2)^{\alpha_1(1-\alpha_2)}}{\alpha_1^{\alpha_1\alpha_2} (1-\alpha_1)^{(1-\alpha_1)\alpha_2}}\right)^{\frac{1}{\alpha_1 - \alpha_2}}.$$

The stated inequalities now follow directly, since $\lambda_i \geq 0,\ i = 3, 4$.

Moreover, from conditions 1. we have

$$\left(\frac{K_i}{L_i}\right)^{\alpha_i - 1} = \frac{r + \lambda_4}{\alpha_i p_i} \text{ and } \left(\frac{K_i}{L_i}\right)^{\alpha_i} = \frac{\omega + \lambda_3}{(1-\alpha_i) p_i},\ i = 1, 2.$$

Defining $\beta_i := \left(\frac{\omega+\lambda_3}{(1-\alpha_i)p_i}\right)^{\frac{1}{\alpha_i}},\ i = 1,2$, we obtain $K_i = \beta_i L_i,\ i = 1,2$. Since $K_1 + K_2 = \bar{K}$, it follows that $\beta_1 L_1 + \beta_2 L_2 = \bar{K}$. Together with $L_1 + L_2 = \bar{L}$ this yields

$$L_1 = -\frac{\bar{K} - \beta_2 \bar{L}}{\beta_2 - \beta_1}, L_2 = \frac{\bar{K} - \beta_1 \bar{L}}{\beta_2 - \beta_1}, \text{ and } K_i = \beta_i L_i,\ i = 1,2.$$

Note that requiring positive amounts of labor and capital imposes additional parametric restrictions for this case to apply.

Finally, if $\alpha_i,\ i = 1,2$, is small, production is capital-intensive. In that case, even low output levels may exhaust available capacity. For production to remain profitable, the cost of capital must therefore not be too high.

CHAPTER 6

Optimization problems with more than one decisionmaker

6.1 Full cooperation: the Pareto frontier

6.1.1 Sufficient Conditions for Pareto Efficiency

EXERCISE 6.1

Reconsider Example 6.2. Use Theorem 6.7 to show that the Pareto frontier is given by $\{(x, y) \mid x+y = 100,\ x \geq 0,\ y \geq 0\}$.

ANSWER 6.1

First, note that the feasible set of instruments, $\{(x, y) \mid x + y \leq 100,\ x \geq 0,\ y \geq 0\}$, is convex. Moreover, both $J_1(x, y) = x$ and $J_2(x, y) = y$ are concave functions. Hence, by Theorem 6.7, all Pareto solutions can be obtained by solving, for each $\alpha \in \mathcal{A}$, the parameterized maximization problem

$$J_\alpha(x, y) := \alpha x + (1 - \alpha)y.$$

Furthermore, by Theorem 6.4, for all $\alpha \in (0, 1)$, the solution of this optimization problem is Pareto efficient.

Observe that for all $\alpha \in (0, \frac{1}{2})$, the maximum is attained at $(x, y) = (0, 100)$, whereas for all $\alpha \in (\frac{1}{2}, 1)$, the maximum is attained at $(x, y) = (100, 0)$. The corresponding Pareto solutions are $(J_1, J_2) = (0, 100)$ and $(J_1, J_2) = (100, 0)$, respectively.

If $\alpha = \frac{1}{2}$, then $J = x + y$. In this case, all (x, y) such that $x + y = 100$ with $x \geq 0$ and $y \geq 0$ are Pareto efficient. The corresponding Pareto solutions are $(J_1, J_2) = (x, 100 - x)$, where $0 \leq x \leq 100$.

Finally, note that for $\alpha = 0$ and $\alpha = 1$, we recover the Pareto efficient solutions $(x, y) = (0, 100)$ and $(x, y) = (100, 0)$.

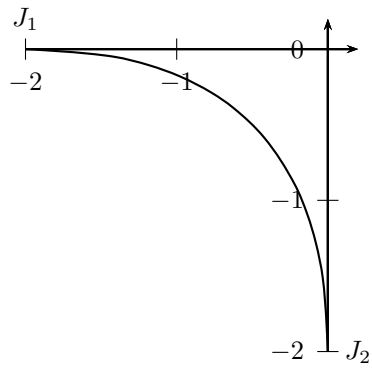


FIGURE 6.1 – *Pareto frontier Exercise 6.2.*

In conclusion, by Theorem 6.7, the set of Pareto solutions is given by $(J_1, J_2) = (x, 100 - x)$, where $0 \leq x \leq 100$. Equivalently, the Pareto frontier is $\{(x, y) \mid x + y = 100,\ x \geq 0,\ y \geq 0\}$.

EXERCISE 6.2

Consider $J_1(x_1, x_2) = -x_1^2 - x_2^2$ and $J_2(x_1, x_2) = -(x_1 - 1)^2 - (x_2 - 1)^2$ where $\Gamma = \mathbb{R}^2$. Calculate the Pareto frontier for this example. Plot the graph of it.

ANSWER 6.2

First, note that the set of instruments $\{(x_1, x_2) \in \mathbb{R}^2\}$ is convex. Moreover, both $J_1(x_1, x_2) = -x_1^2 - x_2^2$ and $J_2(x_1, x_2) = -(x_1 - 1)^2 - (x_2 - 1)^2$ are concave functions.
Hence, by Theorem 6.7, all Pareto solutions can be obtained by solving, for each $\alpha \in \mathcal{A}$, the parameterized maximization problem

$$J_\alpha(x_1, x_2) := \alpha J_1(x_1, x_2) + (1 - \alpha) J_2(x_1, x_2).$$

For a fixed $\alpha \in [0, 1]$, the stationary point of J_α is given by $(x_1, x_2) = (1 - \alpha, 1 - \alpha)$.
Therefore, the Pareto frontier is (see Figure 6.1)

$$(J_1, J_2) = (-2(1 - \alpha)^2, -2\alpha^2),\ \alpha \in [0, 1].$$

EXERCISE 6.3

Let $p = 100 - x_1 - x_2$, $\pi_1(x_1, x_2) = px_1 - (30x_1 + 150)$ and $\pi_2(x_1, x_2) = px_2 - (35x_2 + 50)$. Assume agent $i,\ i = 1, 2$, wants to maximize π_i w.r.t. x_i, subject to the constraints $x_i \geq 0$, $p \geq 0$ and $\pi_i \geq 0$.

a. Use your graphic calculator to make a plot of the set of potential allocations for the instruments (x_1, x_2).

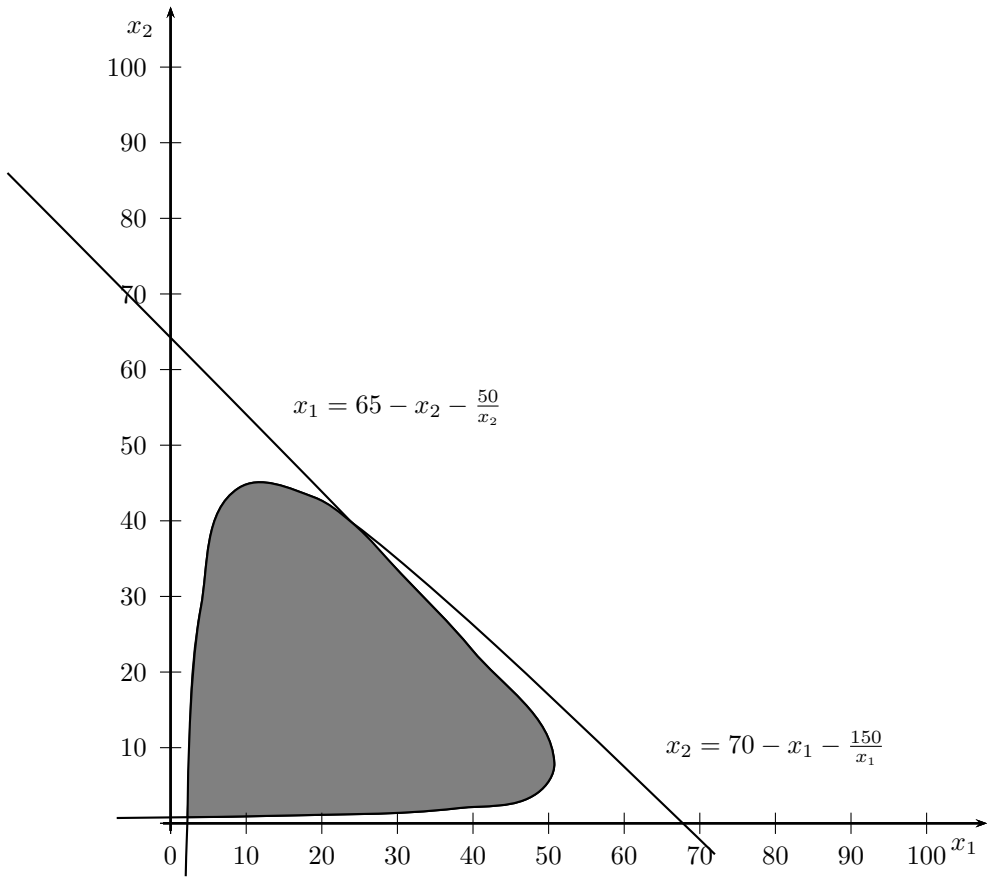


FIGURE 6.2 – *Constraint set of instruments Exercise 6.3.a.*

b. Find for all $\alpha \in (0,1)$ the maximum of $\alpha\pi_1(x_1,x_2) + (1-\alpha)\pi_2(x_1,x_2)$.

c. Plot for the set of potential allocations the solution set $(\pi_1(x_1,x_2), \pi_2(x_1,x_2))$. Indicate the Pareto frontier in this plot.

d. Does the set of Pareto solutions you found in item b coincide with the Pareto frontier you found in item c? Expound on the result.

ANSWER 6.3

a. Note that $\pi_1(x_1,x_2) \geq 0$ iff $x_2 \leq 70 - x_1 - \frac{150}{x_1}$ and $\pi_2(x_1,x_2) \geq 0$ iff $x_1 \leq 65 - x_2 - \frac{50}{x_2}$. The set of potential allocations for the instruments is illustrated in Figure 6.2.

b. Let

$$\begin{aligned} \pi_\alpha(x_1,x_2) &:= \alpha\pi_1(x_1,x_2) + (1-\alpha)\pi_2(x_1,x_2) \\ &= (100 - x_1 - x_2)(\alpha x_1 + (1-\alpha)x_2)) - 30\alpha x_1 - 35(1-\alpha)x_2 - 100\alpha - 50. \end{aligned}$$

Then $\pi_\alpha^{'} = [-2\alpha x_1 - x_2 + 70\alpha,\ -x_1 - 2(1-\alpha)x_2 + 65(1-\alpha)]$.
Provided that $\alpha \neq \frac{1}{2}$, there is a unique stationary point,

$$(x_1, x_2) = ((1-\alpha)\frac{140\alpha - 65}{4\alpha(1-\alpha) - 1}, \alpha\frac{60 - 130\alpha}{4\alpha(1-\alpha) - 1}).$$

However, examining the second-order derivative shows that π_α attains a saddle-point there. For $\alpha = \frac{1}{2}$, no maximum exists either (see also Exercise 2.34).
From the plot in item a, we observe that the feasible set is compact. Therefore, the maximum must be attained on the boundary. This occurs either where

$$x_2 = 70 - x_1 - \frac{150}{x_1} (\text{where } \pi_1 = 0),\ \text{with } 0 < x_1 < 65 - x_2 - \frac{50}{x_2},$$

or where

$$x_1 = 65 - x_2 - \frac{50}{x_2} (\text{where } \pi_2 = 0),\ \text{with } 0 < x_2 < 70 - x_1 - \frac{150}{x_1}.$$

The intersection points of these curves are approximately $(x_1, x_2) = (2.24, 0.796)$ and $(x_1, x_2) = (23.97, 39.77)$. First, consider the case $x_2 = 70 - x_1 - \frac{150}{x_1}$, with $2.24 \leq x_1 \leq 23.97$. Then

$$\begin{aligned}
\pi_\alpha(x_1, x_2(x_1)) &:= (100 - x_1 - x_2)(\alpha x_1 + (1-\alpha)x_2) - 30\alpha x_1 - 35(1-\alpha)x_2 - 100\alpha - 50 \\
&= (30 + \frac{150}{x_1})(\alpha x_1 + (1-\alpha)x_2) - 30\alpha x_1 - 35(1-\alpha)x_2 - 100\alpha - 50 \\
&= (1-\alpha)(-50 + x_2(\frac{150}{x_1} - 5)).
\end{aligned}$$

Thus,

$$\pi_\alpha^{'}(x_1) = (1-\alpha)(x_2^{'}(\frac{150}{x_1} - 5) - x_2\frac{150}{x_1^2}).$$

From this, it follows that the maximum is attained, for all α, at approximately $(x, y) = (4.03, 28.75)$ yielding $(\pi_1, \pi_2) = (0, 876.3)$.
Similarly, consider $x_1 = 65 - x_2 - \frac{50}{x_2}$, with $0.796 \leq x_2 \leq 39.77$. Then

$$\begin{aligned}
\pi_\alpha(x_1(x_2), x_2) &:= (100 - x_1 - x_2)(\alpha x_1 + (1-\alpha)x_2) - 30\alpha x_1 - 35(1-\alpha)x_2 - 100\alpha - 50 \\
&= (35 + \frac{50}{x_2})(\alpha x_1 + (1-\alpha)x_2) - 30\alpha x_1 - 35(1-\alpha)x_2 - 100\alpha - 50 \\
&= \alpha(-150 + x_1(\frac{50}{x_2} + 5)).
\end{aligned}$$

Hence,

$$\pi_\alpha^{'}(x_2) = \alpha(x_1^{'}(\frac{50}{x_2} + 5) - x_1\frac{50}{x_2^2}).$$

From this, we conclude that the maximum is attained, for all α, at approximately $(x, y) = (33.22, 1.66)$ yielding $(\pi_1, \pi_2) = (1016.7, 0)$.

c. Figure 6.3 shows the contour of the Pareto frontier.

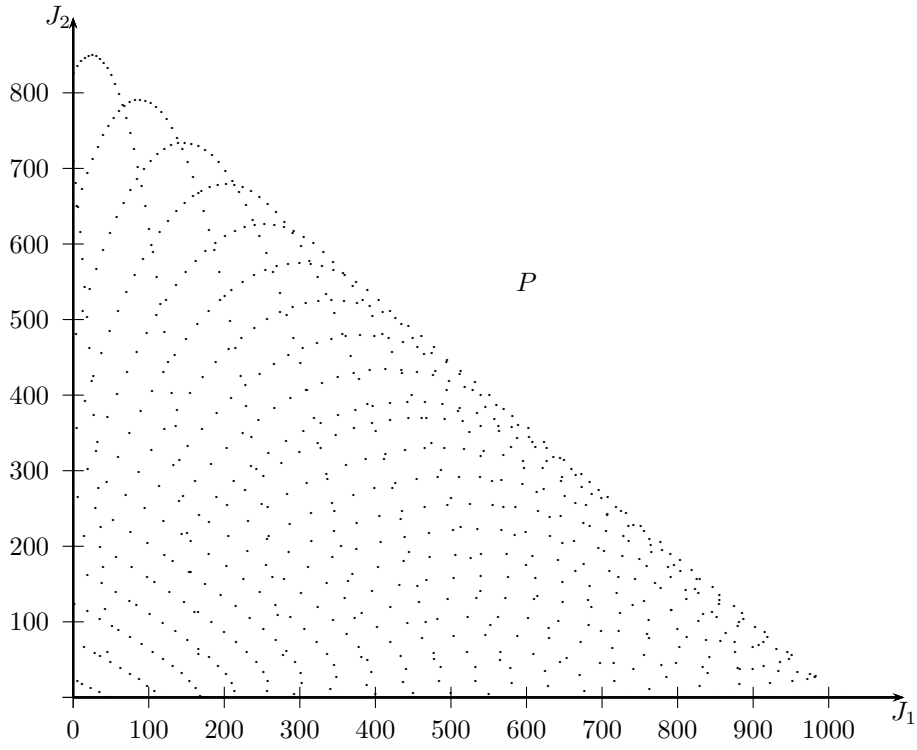


FIGURE 6.3 – *The Pareto frontier P for Exercise 6.3c.*

d. In item b, we only identified the corner points (0,876.3) and (1016.7,0) of the Pareto frontier. Since both profit functions are not convex, theory does not guarantee that all points on the Pareto frontier can be found by considering a weighted average of these two profit functions. This is clearly demonstrated in item c.

6.1.3 An Algorithm to find Pareto Solutions

EXERCISE 6.4 Find the Pareto frontier in the next cooperative optimization problems.

a. (*optimal production planning*)

i) A baker can produce both bread and cookies. He has just one oven and can bake either twenty breads or two hundred cookies. The production time of twenty breads is 1 hour and of two hundred cookies 10 minutes. Assume he can use the oven for 8 hours. Calculate the number of bread and cookies he can produce.

ii) Assume the price of one bread is 2 euro and that of 20 cookies is 1 euro. Calculate how many bread and cookies he should produce to maximize his revenues.

b. (*sharing limited production resources*)

Consider two brothers who own a sawing mill. One of the brothers produces shelves which he sells

to industry. The other one uses the mill to produce material for his furniture company. The time to produce one bundle, X, of shelves is 1 hour. The time to produce the specific material needed for the furniture, Y, is 2 hours. The mill can be used 16 hours a day. Assume the price of X is 20 euro and the price of Y is 50 euro. Calculate the number of shelves and specific material the brothers can optimally produce. What will be the optimal production if they decide to use the mill in such a way that their profits are the same.

c. (*fiscal coordination*)
Consider optimal fiscal policies in a large country 1 and a small country 2. Assume that output, y_i, in each country is affected by fiscal instruments, f_i, used by both countries (more precisely, f_i models the deviation of fiscal policies from their current equilibrium value). Let $y_1 = 4\bar{y} + 0.9f_1 + 0.1f_2$ and $y_2 = \bar{y} + 0.8f_2 + 0.2f_1$, where $\bar{y}$ is some given number. Assume both countries are in a recession and like to raise output to $5\bar{y}$ and $1.2\bar{y}$ respectively. This, of course, using as less as possible fiscal instruments. In country 1 this is formulated as the objective to minimize $(y_1 - 5\bar{y})^2 + f_1^2$ and in country 2 as the objective to minimize $(y_2 - 1.2\bar{y})^2 + f_2^2$. Find the Pareto frontier for this problem if both countries coordinate fiscal policies and $\bar{y} = 1$. What will be the outcome if both countries decide to minimize the sum of their objectives in that case (the so-called *social outcome*)?

d. (*sharing the cakes*)
Consider two consumers who like to spend (a part of) their income on the consumption of two goods. Their utility from consumption of an amount of x from good 1 and y from good 2 is $U_1(x, y) = 2\ln(x + 1) + \ln(y + 1)$ and $U_2(x, y) = 2\ln(x + 1) + \frac{4}{3}\ln(y + 1)$, whereas their budget constraints are $4x + 2y \leq 2$ and $4x + 2y \leq 4$, respectively.
Determine the solution(s) to this cooperative problem in case the maximum available amount of the goods is 10. The same question if this amount is 1 for both goods.

ANSWER 6.4

a. i) Let b denote the number of times he produces bread and c the number of times he produces cookies. Then, within 8 hours, he can produce $20b$ units of bread and $200c$ units of cookies, respectively. Here, b and c are nonnegative integers satisfying $60b + 10c = 480$.
ii) The revenue is given by $2 * 20b + 10 * c$. Substituting $c = 48 - 6b$, we obtain:

$$2 * 20b + 10 * (48 - 6b) = 480 - 40b.$$

Thus, revenue is maximized when $b = 0$ and $c = 48$. Therefore, the baker should produce only cookies.

b. All combinations (X, Y) satisfying $X + 2Y = 16$, where X and Y are nonnegative integers, are feasible.
The brothers' profit functions are $20X$ and $50Y$, respectively. Imposing the condition that both brothers earn equal profit yields $20X = 50Y$, or $X = \frac{5}{2}Y$. Substituting this into the constraint

gives $\frac{5}{2}Y + 2Y = 16$, which implies $Y = \frac{32}{9}$. However, this solution is not feasible since X and Y must be integers. The best feasible solution under the equal-profit condition is $Y = 2$ with $X = 5$, but this leaves a significant portion of the production capacity unused. Nearby feasible solutions that fully utilize capacity - such as $Y = 3$ with $X = 10$, or $Y = 4$ with $X = 8$ - yield higher profits for both brothers.

c. Country 1 aims to minimize its cost function

$$J_1 := (0.9f_1 + 0.1f_2 - \bar{y})^2 + f_1^2 \text{ with respect to } f_1,$$

while country 2 minimizes

$$J_2 := (0.2f_1 + 0.8f_2 - 0.2\bar{y})^2 + f_2^2 \text{ with respect to} f_2.$$

It can be readily verified that the second-order derivatives $J_i^{"}$, for $i = 1, 2,$ are positive definite. Hence, both J_1 and J_2 are convex. Therefore, the Pareto frontier can be obtained by, for each $\alpha \in [0, 1]$, minimizing the weighted sum

$$J_\alpha = \alpha J_1 + (1 - \alpha)J_2.$$

and determining the corresponding minimizers $(f_1^*(\alpha), f_2^*(\alpha))$. The derivative of J_α is given by

$$\begin{aligned} J_\alpha^{'} &= [(3.54\alpha + 0.08)f_1 + (0.32 - 0.14\alpha)f_2 - (0.08 + 1.72\alpha)\bar{y}, \\ &\quad (0.32 - 0.14\alpha)f_1 + (3.28 - 3.26\alpha)f_2 - (0.32 - 0.12\alpha)\bar{y}]. \end{aligned}$$

Solving the first-order conditions yields

$$\begin{aligned} (f_1^*(\alpha), f_2^*(\alpha)) &= (\frac{0.86\alpha + 0.04}{1.77\alpha + 0.04}\bar{y} - \frac{0.16 - 0.07\alpha}{1.77\alpha + 0.04}f_2^*, \\ &\quad \frac{(0.32 - 0.14\alpha)(0.08 + 1.72\alpha) - (3.54\alpha + 0.08)(0.32 - 0.12\alpha)}{(0.32 - 0.14\alpha)^2 - (3.54\alpha + 0.08)(3.28 - 3.26\alpha)}\bar{y}). \end{aligned}$$

Substituting into J_1 and J_2 then yields the Pareto solution (J_1, J_2). In Figure 6.4, we plot the set of Pareto solutions for all $\alpha \in \mathcal{A}$, which traces out the Pareto frontier P.

For $\alpha = 0.5$, we obtain $J_{0.5} = (0.5435, 0.008)$.

d. By straightforward differentiation of U_i, one can verify that the second-order derivatives $U_i^{"}$ are negative definite. Hence, both utility functions are concave. Moreover, the set of feasible instruments is given by

$$\Gamma = \{(x, y) \mid x \geq 0,\ y \geq 0,\ 2x + y \leq 1\},$$

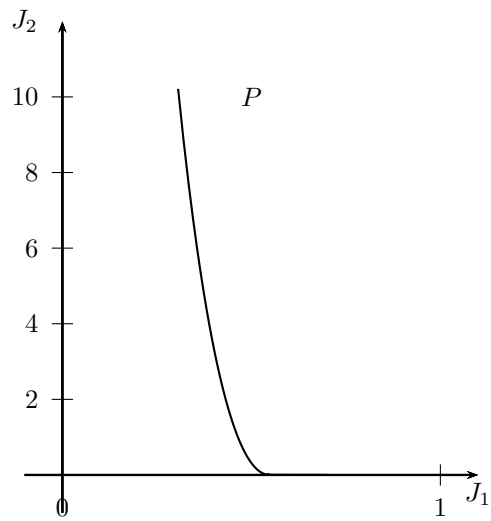


FIGURE 6.4 – *The Pareto frontier P for Exercise 6.4 item c.*

which is easily seen to be convex. Therefore, by Theorem 6.7, all Pareto efficient solutions can be obtained by solving, for each $\alpha \in [0,1]$, the problem

$$\max_{x,y} U_\alpha(x,y) := \alpha U_1 + (1-\alpha)U_2 \text{ subject to } (x,y) \in \Gamma.$$

Since Γ is compact the maximum is attained either in the interior or on the boundary. From the first-order condition $U^{'} = 0$, it follows that there are no interior stationary points.
Next, we check the boundaries.

1. if $x = 0$, then $U_\alpha(0,y) = \frac{1}{3}(4-\alpha)\ln(1+y)$, which is maximized at $y = 1$.
2. if $y = 0$, then $U_\alpha(x,0) = 2\ln(x+1)$, which is maximized at $x = \frac{1}{2}$.
3. if $2x + y = 1$, then $U_\alpha(x, 1-2x) = 2\ln(x+1) + \frac{1}{3}(4-\alpha)\ln(2-2x)$.
   Differentiating $U_\alpha(x, 1-2x)$ shows that it has a stationary point at $x = \frac{2+\alpha}{10-\alpha}$, which implies $y = 1 - 2x = \frac{6-3\alpha}{10-\alpha}$. At this point

$$U_\alpha(x, 1-2x) = 2\ln(\frac{12}{10-\alpha}) + \frac{1}{3}(4-\alpha)\ln(\frac{16-4\alpha}{10-\alpha}).$$

It follows that, for all $\alpha \in \mathcal{A}$, the maximum of $U_\alpha(x,y)$ is attained at

$$(x^*, y^*) := (\frac{2+\alpha}{10-\alpha}, \frac{6-3\alpha}{10-\alpha}).$$

Hence, the set of Pareto solutions is

$$(2\ln(x^*+1) + \ln(y^*+1), 2\ln(x^*+1) + \frac{4}{3}\ln(y^*+1)), \text{ where } 0 \leq \alpha \leq 1.$$

If, in addition, both x and y should not exceed $\frac{1}{2}$, the feasible set of instruments becomes

$$\Gamma = \{(x, y) \mid x \geq 0,\ 0 \leq y \leq \frac{1}{2},\ 2x + y \leq 1\}.$$

In this case, the solution (x^*, y^*) remains feasible only if

$$\frac{6 - 3\alpha}{10 - \alpha} \leq \frac{1}{2}, \text{ that is, if } \frac{2}{5} \leq \alpha \leq 1.$$

Otherwise, the maximum is attained at $y = \frac{1}{2}$ and $x = \frac{1}{4}$.
Thus, the set of Pareto solutions is reduced to

$$(2\ln(x^* + 1) + \ln(y^* + 1), 2\ln(x^* + 1) + \frac{4}{3}\ln(y^* + 1)),\ \frac{2}{5} \leq \alpha \leq 1.$$

6.2 Noncooperative Solution

EXERCISE 6.5 Reconsider Example 6.24.

a. Use Theorem 6.30 to show that this problem has a Nash equilibrium on every set $\Gamma := [0, a] \times [0, b]$, where a and b are some positive constants.

b. Plot the reaction functions and next calculate the Nash equilibrium if $a = b = 38$.

ANSWER 6.5

a. Note that Γ is a compact and convex set. Moreover, $\pi_i(p_1, p_2) \in C^2$ for $i = 1, 2$.
For a fixed $\bar{p}_2 \in [0, b]$, the function

$$\pi_1(p_1, \bar{p}_2) = (100 - \frac{5}{4}p_1 - \bar{p}_2)(p_1 - 30) - 120$$

is strictly concave on $[0, a]$, since $\pi_1^{"}(p_1, \bar{p}_2) = -\frac{5}{4} < 0$.
Similarly, for a fixed $\bar{p}_1 \in [0, a]$, the function

$$\pi_2(\bar{p}_1, p_2) = (80 - \bar{p}_1 - p_2)(p_2 - 35) - 5$$

is strictly concave on $[0, b]$, as $\pi_2^{"}(\bar{p}_1, p_2) = -1 < 0$.
Therefore, by Theorem 6.30, this problem admits a Nash equilibrium on every set $\Gamma := [0, a] \times [0, b]$.

b. From Figure 6.5 we see that the Nash efficient equilibrium prices are $(38, 38)$.

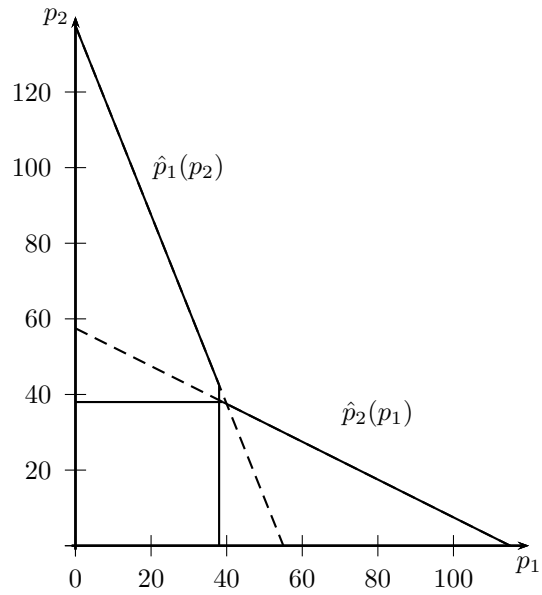


FIGURE 6.5 – *The reaction functions for Exercise 6.4 item b.*

6.3 Hierarchical Solution

EXERCISE 6.6 *(fiscal coordination)*
Consider three countries where country 1 is in a recession, country 2 is performing well and country 3's economy is overheated. All countries want to use their fiscal policies, f_i (see Exercise 6.4 item c.), to adapt output, y_i, in their country. Assume

$$y_1 = 5 + 0.8f_1 + 0.1f_2 + 0.1f_3,\ y_2 = 10 + 0.1f_1 + 0.8f_2 + 0.1f_3 \text{ and } y_3 = 15 + 0.1f_1 + 0.1f_2 + 0.8f_3.$$

The cost functions the different countries like to minimize are given by

$$J_1 = (y_1 - 10)^2 + f_1^2,\ J_2 = (y_2 - 10)^2 + f_2^2 \text{ and } J_3 = (y_3 - 10)^2 + f_3^2,$$

respectively. Assume the first two countries are willing to cooperate in order to tackle the economic problems.

a. Formulate this problem as a non-cooperative situation, where country 1 and 2 play in a coalition against country 3. Assume that the two cooperative countries formalize their cooperation in an agreement to minimize a weighted sum of their cost functions, i.e., $\alpha J_1 + (1 - \alpha)J_2$. Determine the reaction functions of the players.
b. Determine for fixed α the equations that have to be satisfied by a Nash equilibrium in the game in item a). Calculate, using e.g. a symbolic computerpackage, the Nash efficient equilibrium fiscal policies for this game.

c. Use a computerpackage, e.g. Matlab, to plot the Nash efficient fiscal policies in item b) and corresponding output for $\alpha \in (0,1)$ for country 1 and 2. Comment on the policies as a function of α.

d. Similarly, plot the Nash efficient fiscal policy and corresponding output for country 3 as a function of α. Explain again the results.

e. Plot the Nash solution $(J_1(\alpha), J_2(\alpha))$ for $\alpha \in (0,1)$. What is the outcome for $\alpha = 0.5$? Show that for approximately $\alpha = 0.9424$, $J_1(\alpha) = J_2(\alpha)$.

ANSWER 6.6

a. So countries 1 and 2 want to minimize w.r.t. f_1 and f_2:

$$\begin{aligned} J_{12}(\alpha) &:= \alpha J_1 + (1-\alpha)J_2 \\ &= \alpha(0.8f_1 + 0.1f_2 + 0.1f_3 - 5)^2 + \alpha f_1^2 + (1-\alpha)(0.1f_1 + 0.8f_2 + 0.1f_3)^2 + (1-\alpha)f_2^2. \end{aligned}$$

Whereas country 3 wants to minimize J_3 w.r.t. f_3.
Minima are obtained by the first-order conditions. Differentiation of J_{12} w.r.t $f_i,\ i = 1,2$ yields next condition should hold:

$$\begin{bmatrix} 0.02 + 3.26\alpha & 0.16 \\ 0.16 & 3.28 - 3.26\alpha \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} 8\alpha - (0.02 + 0.14\alpha)f_3 \\ \alpha\alpha - (0.16 - 0.14\alpha)f_3 \end{bmatrix}.$$

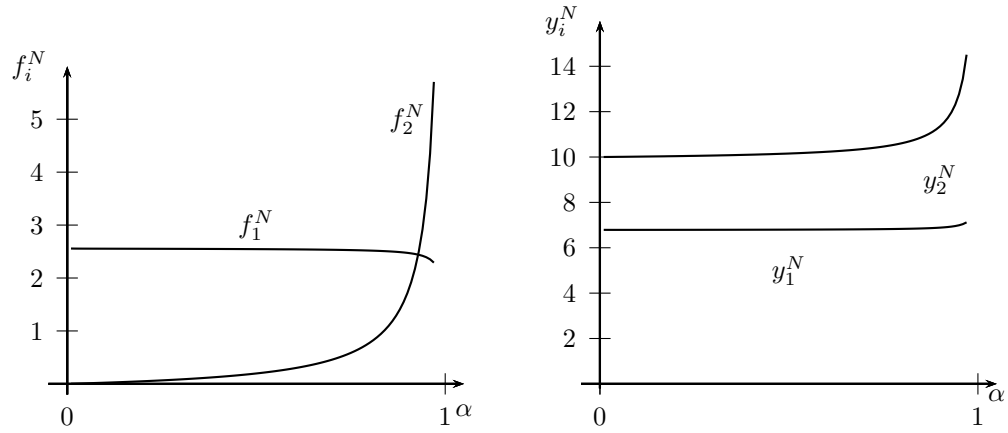
Moreover, we directly obtain by straightforward differentiation that

$$HJ_{12}(\alpha) = \begin{bmatrix} 0.02 + 3.26\alpha & 0.16 \\ 0.16 & 3.28 - 3.26\alpha \end{bmatrix}.$$

One easily verifies that, for all $\alpha \in [0,1]$, $HJ_{12}(\alpha)$ is positive definite. So, for all $\alpha \in [0,1]$, $J_{12}(\alpha)$ has a minimum at its stationary point.
Therefore the reaction functions for countries 1 and 2 are:

$$\begin{aligned} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} &= \begin{bmatrix} 0.02 + 3.26\alpha & 0.16 \\ 0.16 & 3.28 - 3.26\alpha \end{bmatrix}^{-1} \begin{bmatrix} 8\alpha - (0.02 + 0.14\alpha)f_3 \\ \alpha - (0.16 - 0.14\alpha)f_3 \end{bmatrix} \\ &= \frac{1}{d} \begin{bmatrix} 3.28 - 3.26\alpha & -0.16 \\ -0.16 & 0.02 + 3.26\alpha \end{bmatrix} \begin{bmatrix} 8\alpha - (0.02 + 0.14\alpha)f_3 \\ \alpha - (0.16 - 0.14\alpha)f_3 \end{bmatrix}, \end{aligned}$$

with $d = (3.28 - 3.26\alpha)(0.02 + 3.26\alpha) - 0.16^2$.
Twice differentiation of J_3 w.r.t. f_3 shows that also f_3 has a minimum at its stationary point, which gives rise to next reaction function for country 3:
$f_3 = -\frac{1}{3.28}(0.16f_1 + 0.16f_2 + 8)$.



(a) Fiscal policies.

(b) Output

FIGURE 6.6 – *The Nash efficient fiscal policies and output for Exercise 6.6 item c.*

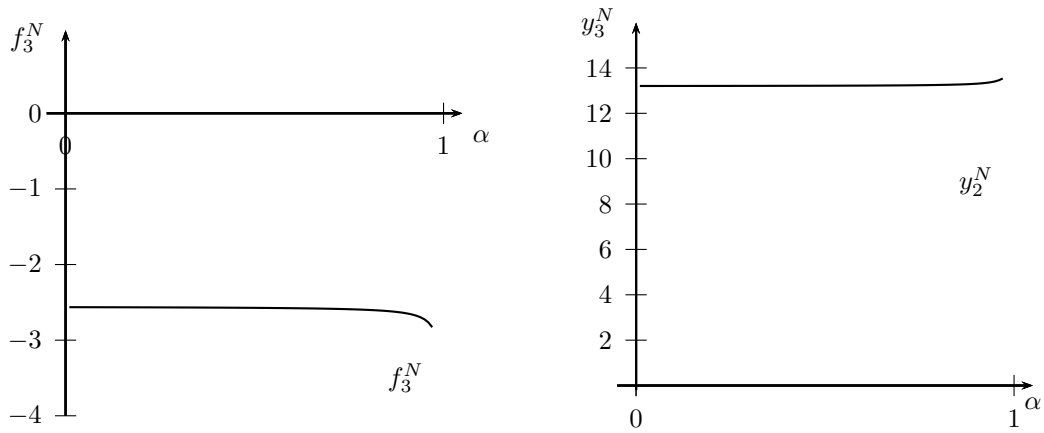
b. From item a. we have that the Nash equilibrium is obtained as the intersection of all reaction functions. That is, as the solution of next equation:
$\begin{bmatrix} 0.02+3.26\alpha & 0.16 & 0.02+0.14\alpha \\ 0.16 & 3.28-3.26\alpha & 0.16-0.14\alpha \\ 0.16 & 0.16 & 3.28 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} 8\alpha \\ \alpha \\ -8 \end{bmatrix}$. That is,

$$\begin{bmatrix} f_1^N \\ f_2^N \\ f_3^N \end{bmatrix} = \begin{bmatrix} 0.02+3.26\alpha & 0.16 & 0.02+0.14\alpha \\ 0.16 & 3.28-3.26\alpha & 0.16-0.14\alpha \\ 0.16 & 0.16 & 3.28 \end{bmatrix}^{-1} \begin{bmatrix} 8\alpha \\ \alpha \\ -8 \end{bmatrix}$$
$$= \frac{1}{d_2} \begin{bmatrix} 200(-2781\alpha^2+2771\alpha+10) \\ 250\alpha(171\alpha+1) \\ -200(-2771\alpha^2+2781\alpha+10) \end{bmatrix},$$

where $d_2 = -216953\alpha^2 + 216953\alpha + 780$.

c. From Figure 6.6 we see that if country 2's preferences have a substantial contribution to the cooperative cost function J_{12}, fiscal policy of country 2 will be almost negligible. This makes sense, as this country is satisfied with its target output value of 10. This changes, if the preferences of country 1 become more substantial in the cooperative cost function. Country 1 reduces then its fiscal policies somewhat at the expense of those of Country 2 which increases them substantially. That is, Country 2 tries to help Country 1 to get out of its recession. We see this impact is only minor. Country 2 primarily only overheats its own economy. This, since fiscal policies pursued by both countries do not strongly influence each other's economy.

d. From Figure 6.7 we see that Country 3 tries to decrease output by lowering fiscal policies. Which results indeed into a reduction of output into the direction of the target value. Moreover we observe,



(a) Fiscal policies. (b) Output

FIGURE 6.7 – *The Nash efficient fiscal policies and output for Exercise 6.6 item d.*

like in item c., that only for larger values of α Country 3 adapts its fiscal policy somewhat. Due to the substantial increase of Country 2's fiscal policy, Country 3 tries to decrease its fiscal policies even more to mitigate this expansionary effect on its output.

e. In Figure 6.8 we plotted the corresponding Nash solutions if α ranges from 0 to 1. For $\alpha = 0.5$ we obtain $(J_1^N(0.5), J_2^N(0.5)) = (16.7274, 0.0625)$. The value of α where approximately $J_1^N(\alpha) = J_2^N(\alpha)$ is $\alpha = 0.9424$ with $J_i^N = 15.0391,\ i = 1, 2$.

EXERCISE 6.7 Reconsider Example 6.24.

a. Determine the Stackelberg solution and its corresponding efficient actions if farmer 1 is the leader. Is this a realistic scenario?

b. In order to pursuade farmer 2 to satisfy the demand of his clients which price should farmer 1 charge at most for his bananas?

c. Answer the same question as in item a) if farmer 2 is the leader.

d. Plot the Pareto frontier, the Nash solution and the Stackelberg solutions of item b) and c) in one figure. Compare the distances of the last three solutions to the Pareto frontier and conclude.

ANSWER 6.7

a. The reaction function of farmer 2 is given by $p_2 = \frac{1}{2}(115 - p_1)$.
Substituting this into the profit function of farmer 1 yields

$$\pi_1(p_1) = (42.5 - \frac{3}{4}p_1)(p_1 - 30) - 120.$$

Farmer 1's profit is maximized at the stationary point of π_1, which occurs at $p_1 = 43\frac{1}{3}$. The corresponding reaction price of farmer 2 is $p_2 = \frac{1}{2}(115 - p_1) = 35\frac{5}{6}$.

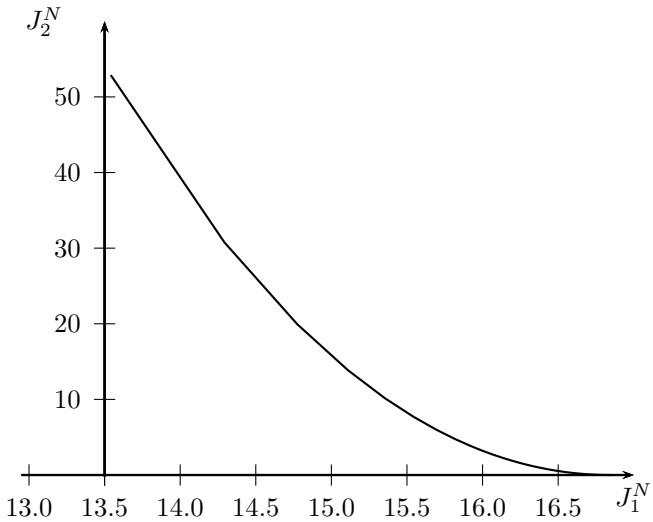


FIGURE 6.8 – *The Nash Solutions for Exercise 6.6 item e,* $\alpha \in (0, 1)$.

At these prices, farmer 1's profit equals $\pi_1 = 13\frac{1}{3}$, while farmer 2's profit is $\pi_2 = -4\frac{11}{36}$. Since farmer 2 incurs a loss, he would no longer be willing to sell his bananas.

b. Farmer 1's optimization problem is to maximize its profit, taking into account that farmer 2's reaction function is $p_2 = \frac{1}{2}(115 - p_1)$, and that farmer 2's profit π_2 must be nonnegative.
Thus farmer 2 seeks to maximize

$$\pi_1(p_1) = (42.5 - \frac{3}{4}p_1)(p_1 - 30) - 120,$$

subject to the constraint

$$\pi_2(p_1) = (80 - p_1 - \frac{1}{2}(115 - p_1))(\frac{1}{2}(115 - p_1) - 35) - 5 = (22\frac{1}{2} - \frac{1}{2}p_1)^2 - 5 \geq 0.$$

From item a., it follows that the maximum is attained at the boundary of the feasible set, i.e., for values of p_1 satisfying $(22\frac{1}{2} - \frac{1}{2}p_1)^2 = 5$. Solving this equation yields that $p_1 = 45 - \sqrt{20}$ maximizes profit. At this point, farmer 1's profit is approximately 7.43, while the corresponding price for farmer 2 is $p_2 = \frac{1}{2}(70 + \sqrt{20})$, at which farmer 2 earns zero profit.

c. If farmer 2 acts as the leader, the reaction function of farmer 1 is given by $p_1 = 55 - \frac{2}{5}p_2$.
Substituting this reaction function into farmer 2's profit function yields

$$\pi_2(p_2) = (25 - \frac{3}{5}p_2)(p_2 - 35) - 5.$$

Farmer 2 maximizes profit at the stationary point of π_2, which occurs at $p_2 = 38\frac{1}{3}$. The corresponding reaction price of farmer 1 is $p_1 = 55 - \frac{2}{5}p_2 = 39\frac{2}{3}$.
In this case, farmer 2's profit is $\pi_2 = \frac{5}{3}$, while farmer 1's profit is $\pi_1 = 0.0278$.

d. Consider the maximization of

$$\pi_\alpha(p_1, p_2) := \alpha\pi_1 + (1-\alpha)\pi_2,$$

where

$$\pi_1 = (100 - \frac{5}{4}p_1 - p_2)(p_1 - 30) - 120 \text{ and } \pi_2 = (80 - p_1 - p_2)(p_2 - 35) - 5.$$

Differentiation π_α yields:

$$\begin{aligned} \pi_\alpha^{'} &= [-\frac{5}{4}\alpha(p_1 - 30) + \alpha(100 - \frac{5}{4}p_1 - p_2) - (1-\alpha)(p_2 - 35), \\ &\quad -\alpha(p_1 - 30) - (1-\alpha)(p_2 - 35) + (1-\alpha)(80 - p_1 - p_2)]. \end{aligned}$$

From this, the first-order conditions can be written as

$$\begin{bmatrix} \frac{5}{2}\alpha & 1 \\ 1 & 2(1-\alpha) \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 35 + 102\frac{1}{2} \\ 115 - 85\alpha \end{bmatrix}.$$

From this expression, it can also easily be verified that the Hessian $H\pi_\alpha$ is negative definite for $\alpha \in [\frac{5-\sqrt{5}}{10}, \frac{5+\sqrt{5}}{10}]$. Therefore, π_α is concave on this interval.
For these values of α, the corresponding segment of the Pareto frontier is obtained by computing the stationary point of π_α, which gives

$$\begin{aligned} \begin{bmatrix} p_1^* \\ p_2^* \end{bmatrix} &= \begin{bmatrix} \frac{5}{2}\alpha & 1 \\ 1 & 2(1-\alpha) \end{bmatrix}^{-1} \begin{bmatrix} 35 + 102\frac{1}{2} \\ 115 - 85\alpha \end{bmatrix} \\ &= \frac{1}{1 - 5\alpha(1-\alpha)} \begin{bmatrix} 45 - 220\alpha + 205\alpha^2 \\ 35 - 185\alpha + \frac{425}{2}\alpha^2 \end{bmatrix}. \end{aligned}$$

This yields the Pareto solutions $(\pi_1(p_1^*, p_2^*), \pi_2(p_1^*, p_2^*))$.
Figure 6.9 illustrates the results for $\alpha \in [0.2825, 0.2842]$. Only within this range of parameters do both farmers choose to produce. Also in this exercise we observe that the Nash solution is somewhat closer to the Pareto frontier than the Stackelberg solution.

EXERCISE 6.8 Reconsider the fiscal policy coordination problem, Exercise 6.1.4.c.

a. Assume country 1 acts as a leader in this fiscal policy game. Determine the corresponding Stackelberg solution. Compare this solution with the Nash solution and conclude.
b. Answer the same questions as in item a) but now with country 2 acting as a leader.
c. Plot the Stackelberg solutions, the Nash solution and the Pareto frontier in one figure. Compare the distances of the different solutions to the Pareto frontier and conclude.

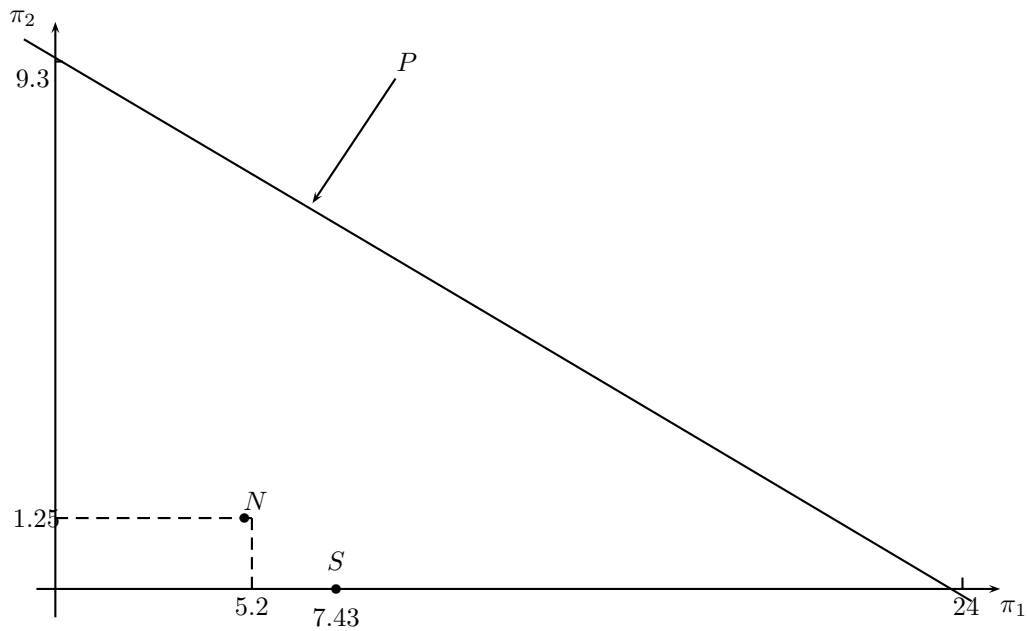


FIGURE 6.9 – *The Pareto (P), Nash (N) and Stackelberg (S) solution in Exercise 6.7 item d.*

ANSWER 6.8

a. With $f_1 = \bar{f}_1$, the reaction function of player 2 is $f_2 = 0.0976(\bar{y} - \bar{f}_1)$. Thus, player 1's optimization problem reduces to minimizing

$$J_1(f_1) = (-\bar{y} + 0.9f_1 + 0.1 * 0.0976 * (\bar{y} - f_1))^2 + 2f_1.$$

Straightforward calculations show that the minimum is attained at $f_1 = 0.5239\bar{y}$. This implies that the follower chooses $f_2 = 0.0465\bar{y}$.
Hence, the Stackelberg solution is $(J_1^S, J_2^S) = (0.5551, 0.0110)\bar{y}^2$.
The reaction function of player 1 is given by $f_1 = 0.5525(\bar{y} - 0.1f_2)$. The Nash efficient solutions are obtained as the intersection point of the two reaction functions, yielding $(f_1, f_2) = (0.5462, 0.1134)\bar{y}$. This leads to the Nash solution $(J_1^N, J_2^N) = (0.5454, 0.0129)\bar{y}^2$.
As expected, player 1 benefits while player 2 is worse off in the Stackelberg equilibrium compared to the Nash solution.

b. From item a., the reaction function of player 1 is given by $f_1 = 0.5525(\bar{y} - 0.1 * f_2)$. Thus, player 2's optimization problem reduces to minimizing

$$J_2(f_2) = (-0.2\bar{y} + 0.2 * 0.5525(\bar{y} - 0.1 * f_2) + 0.8 * f_2)^2 + f_2^2.$$

Straightforward calculations show that the minimum is attained at $f_2 = 0.0435\bar{y}$. This implies that player 1 will choose $f_1 = 0.5501\bar{y}$. Hence, the Stackelberg solution is $(J_1^S, J_2^S) = (0.5532, 0.0049)\bar{y}^2$.

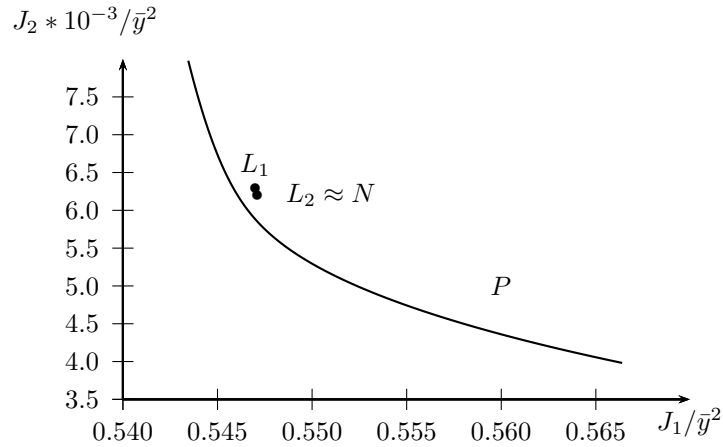


FIGURE 6.10 – *The Pareto frontier P for Exercise 6.8 item c.*

This is again consistent with theory: player 2 gains relative to the Nash solution, while player 1 looses compared to the Nash solution.

c. From Figure 6.10, we observe that all solutions nearly coincide. Therefore, appointing a leader in this case does not appear to provide any significant additional benefit.

EXERCISE 6.9 Consider two companies L and F who produce the same product and sell their product in the same market. The price they get for their product is $p = 5 - (q_L + q_F)$, where q_i is the amount of product company i supplies to the market, $i = L, F$. The cost function of the companies are $C_L(q_L) = q_L^4 - 8q_L^3 + 21\frac{1}{4}q_L^2 - 20q_L$ and $C_F = q_F + q_F^2$, respectively.

a. Determine the profit for company L if it is the leader in this Stackelberg game.

b. What will be the Stackelberg efficient action and corresponding profit of company F in this game if company L likes to minimize profit of its competitor?

ANSWER 6.9

a. The profit function of the follower company F is $\pi_F = p * q_F - C_F$. For a fixed value of q_L, straightforward calculations show that the follower's optimal reaction function is $q_F := 1 - \frac{1}{4}\bar{q}_L$. The leader therefore seeks to maximize

$$\pi_L = p * q_L - C_L = (4 - \frac{3}{4}q_L)q_L - q_L^4 + 8q_L^3 - 21\frac{1}{4}q_L^2 + 20q_L.$$

Simple calculations show that π_L attains global maxima at $q_L = 1$ and $q_L = 3$, while it has a local minimum at $q_L = 2$.

b. With $q_L = 1$, $q_F = \frac{3}{4}$, $p = 3\frac{1}{4}$ and $\pi_F = p * q_F - C_F = \frac{13}{4} * \frac{3}{4} - \frac{21}{16} = \frac{9}{8}$.
With $q_L = 3$, $q_F = \frac{1}{4}$, $p = 1\frac{3}{4}$ and $\pi_F = p * q_F - C_F = \frac{7}{4} * \frac{1}{4} - \frac{5}{16} = \frac{1}{8}$.
Thus, the leader will choose $q_L = 3$ in that case.

EXERCISE 6.10 *(worst-case growth optimization)*
Consider a farmer growing potatoes. Using artificial manure, f, he can affect the revenue of his crop, C. This revenue is, however, also subject to weather conditions, w. Assume $C = s_0\sqrt{f} - w$, where $w \in [-1, 1]$. Moreover, assume the farmer gets a price p for his crop potatoes and the cost per unit manure is λ.

a. Formulate the above problem as a game problem, where the farmer is the leader and nature is the follower and nature is indifferent in its choice of w.
b. Determine the max-min Stackelberg solution and corresponding efficient instruments. Interpret this solution in terms of the price λ of fertilizer.
c. Formalize the Stackelberg solution concept that yields the same solution for the farmer as in item b), in case nature is the leader of the game.

ANSWER 6.10
a. The farmer aims to maximize profit $\pi(f, w)$ with respect to f, where

$$\pi(f, w) := p * C - \lambda * f = p * (s_0\sqrt{f} - w) - \lambda * f \text{ and } w \in [-1, 1].$$

b. If the farmer seeks to maximize the worst-case profit, this occurs when $w = 1$. In fact, regardless of the realization of w, the farmer maximizes profit by using $f = \frac{1}{4}\frac{p^2 s_0^2}{\lambda^2}$ units of fertilizer.
The corresponding profit is $\pi(w) = p(\frac{1}{4}\frac{p^2 s_0^2}{\lambda} - w)$.
If the farmer is risk-averse, he will only produce if $\pi(1) \geq 0$, which holds when the price of fertilizer satisfies $\lambda \leq \frac{1}{4}ps_0^2$.
c. If nature acts as the leader in this game, we obtain the same outcome as in item b., with nature minimizing $\pi(f, w)$ with respect to w and the farmer maximizing profit with respect to fertilizer use.

EXERCISE 6.11 Consider player L who wants to maximize $J_L := -(x_L + x_F + 1)^2(x_L^2 + 1)^2$ w.r.t. x_L and player F who wants to maximize $J_F := -(x_F - x_L)^2(x_F + x_L)^2$ w.r.t. x_F. Assume player L is the leader in this game.
Determine both the max-max and the max-min Stackelberg solution in this game.

ANSWER 6.11 Consider for a fixed value of x_L, the maximization of J_F. Clearly, the maximum is attained at either $x_F = -x_L$ or $x_F = x_L$, which yields in both cases a value of 0 for the follower. The leader next has to maximize either $J_L = -(x_L^2 + 1)^2$ or $J_L = -(2x_L + 1)^2(x_L^2 + 1)^2$ w.r.t. x_L, respectively. In the first case $x_L = 0$ yields the maximum value of -1 for the leader, whereas in the second case $x_L = -\frac{1}{2}$ yields the maximum value of 0 for the leader.
So, the max-max Stackelberg solution is given by $(0, 0)$ and attained at $(x_L, x_F) = (-\frac{1}{2}, -\frac{1}{2})$ whereas the max-min Stackelberg solution is given $(-1, 0)$ and attained at $(x_L, x_F) = (0, 0)$. Obviously, $x_L = 0$ yields here a maximum value of -1 for the leader.

EXERCISE 6.12 Consider a game with a leader, L, and follower, F, who want to minimize their cost function

$$J_L(x_L, x_F) = (x_L - a)^2 + (x_F - b)^2 \text{ and } J_F(x_L, x_F) = (x_L - c)^2 + (x_F - d)^2, \text{ where } a \neq c,$$

respectively. Determine an affine incentive strategy $x_L = \alpha x_F + \beta$ for the leader such that his cost are zero.

ANSWER 6.12 Clearly, α and β must satisfy the conditions $\alpha x_F + \beta - a = 0$ and $x_F = b$.
The follower seeks to minimize

$$J_F = (\alpha x_F + \beta - c)^2 + (x_F - d)^2.$$

From the derivative of J_F we infer $x_F = \frac{\alpha(c-\beta)+d}{1+\alpha^2}$ minimizes J_F.
Therefore, α and β must satisfy the system:

$$\alpha b + \beta - a = 0 \text{ together with } \frac{\alpha(c-\beta)+d}{1+\alpha^2} = b.$$

Equivalently, we can write:

$$\beta = a - \alpha b \text{ and } \alpha(c-\beta) + d - (1+\alpha^2)b = 0.$$

Substituting $\beta = a - \alpha b$ into the second equation gives:

$$\alpha(c - a + \alpha b) + d - (1+\alpha^2)b = 0, \text{ or, } (c-a)\alpha + d - b = 0.$$

Hence, the leader can set the parameters of the affine strategy to achieve zero cost as follows:

$$\alpha = \frac{d-b}{a-c} \text{ and } \beta = a - \alpha b = \frac{a^2 + b^2 - ac - bd}{a-c}.$$

Mixed Exercises

EXERCISE 6.13 Assume player 1 wants to maximize $J_1(x_1, x_2) = 1$ w.r.t. x_1 and player 2 wants to maximize $J_2(x_1, x_2) = (x_2 - 1)^2(x_2 - 2)^2$ w.r.t. x_2, where $(x_1, x_2) \in \Gamma := \{(x_1, x_2) \,|1 \leq x_1 \leq 2 \text{ and } 0 \leq x_2 \leq 3, \ i = 1, 2\}$.

a. Determine the optimal reaction sets $\hat{x}_i(x_{-i}), \ i = 1, 2$.

b. Show that the game has a unique Nash equilibrium, whereas the game has an infinite number of Nash efficient locations.

ANSWER 6.13

a. Let $(\bar{x}_1, \bar{x}_2) \in \Gamma$. Then the sets

$$\Gamma_1(\bar{x}_2) = \{(x_1, \bar{x}_2) \mid (x_1, \bar{x}_2) \in \Gamma\} = [1, 2]$$

and

$$\Gamma_2(\bar{x}_1) = \{(\bar{x}_1, x_2) \mid (\bar{x}_2, x_1) \in \Gamma\} = [0, 3].$$

It follows that $J_1(x, \bar{x}_2)$ attains a global maximum on $\Gamma_1(\bar{x}_2)$ for all $x \in [1, 2]$.
Therefore, $\hat{x}_1(\bar{x}_2) = [1, 2]$ for all $\bar{x}_2 \in [0, 3]$.
Similarly, $J_2(\bar{x}_1, y)$ achieves a global maximum on $\Gamma_2(\bar{x}_1)$ at $y = 1$ and $y = 2$.
Thus, $\hat{x}_2(\bar{x}_1) = \{1, 2\}$ for all $\bar{x}_1 \in [1, 2]$.

b. From item a., the intersection points of the two reaction sets are

$$\{(x_1, x_2) = (x, 1), \text{ where } 0 \leq x \leq 2\} \cup \{(x_1, x_2) = (x, 2), \text{ where } 0 \leq x \leq 2\}.$$

At all of these points the unique Nash equilibrium is attained with $(J_1^N, J_2^N) = (1, 0)$.

EXERCISE 6.14 Use Algorithm 6.17 to find the Pareto frontier in the following cases.

a. If a player wants to maximize both $J_1(x_1, x_2, x_3) = x_1 - x_2$ and $J_2(x_1, x_2, x_3) = x_2 + x_3$ w.r.t. x_j, $j = 1, 2, 3$, where $(x_1, x_2, x_3) \in \Gamma := \{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 + x_3^2 \leq 1\}$.

b. If a player wants to maximize simultaneously $J_1(x_1, x_2) = x_1(1 - x_1 - x_2) - (x_1 + x_2)$ and $J_2(x_1, x_2) = x_2(1 - 2x_1 - 2x_2) - (x_1 + x_2)$, w.r.t. $x_j \geq 0$, $j = 1, 2$.

c. If a player wants to maximize simultaneously $J_1(x_1, x_2) = x_1(5 - x_1 - x_2) - (x_1 + x_2)$ and $J_2(x_1, x_2) = x_2(10 - 2x_1 - 2x_2) - (x_1 + x_2)$, w.r.t. $x_j \geq 0$, $j = 1, 2$.

d. If a government wants to spend its budget $b > 0$ on 3 projects with utilities: $U_i(x_i) = \beta_i x_i^2$, where $\beta_i > 0$ and $x_i \geq 0$, $i = 1, 2, 3$.

ANSWER 6.14

a. First, observe that the action set Γ is compact and convex. The functions $J_i(x_1, x_2, x_3)$, for $i = 1, 2$, are concave. Therefore, by Theorem 6.7, all Pareto efficient solutions can be obtained by solving

$$\max_{(x_1, x_2, x_3) \in \Gamma} \alpha(x_1 - x_3) + (1 - \alpha)(x_2 + x_3), \text{ with } 0 \leq \alpha \leq 1.$$

Since Γ is compact, this maximization problem has a solution, which occurs on the boundary of Γ.
Thus, we can skip Step 2 and Step 3 in Algorithm 6.17.
Step 1. From the corresponding Lagrange function we obtain next first-order conditions:

$$\text{1. } \alpha - 2\lambda x_1 = 0; \text{ 2. } (1 - \alpha) - 2\lambda x_2 = 0; \text{ 3. } 1 - 2\lambda x_3 = 0;$$
$$\text{4. } \lambda(x_1^2 + x_2^2 + x_3^2 - 1) = 0; \text{ 5. } \lambda \geq 0, \ x_1^2 + x_2^2 + x_3^2 \leq 1.$$

From condition 3, it follows that $\lambda > 0$. Then, by condition 4, we have $x_1^2 + x_2^2 + x_3^2 = 1$. Combining conditions 1 and 2 gives $x_1 + x_2 = x_3 = \frac{1}{2\lambda}$. Substituting this into $x_1^2 + x_2^2 + x_3^2 = 1$ yields:

$$x_1 = \frac{\frac{1}{\lambda} + \sqrt{8 - \frac{3}{\lambda^2}}}{4},\ x_2 = \frac{1}{2\lambda} - x_1 \text{ and } x_3 = \frac{1}{2\lambda},\ \text{with } \lambda \geq \sqrt{\frac{8}{3}}.$$

Step 4. The corresponding Pareto frontier is then given by

$$(J_1, J_2) = (x_1 - x_2, x_2 + x_3),$$

with $x_1,\ x_2$ and x_3 as above and $\lambda \geq \sqrt{\frac{8}{3}}$.

b. Following Algorithm 6.17, we proceed with the following steps.

Step 1. Consider the Lagrange function

$$L(x_1, x_2, \alpha_1, \alpha_2) = \alpha_1(x_1(1-x_1-x_2)-x_1-x_2)+\alpha_2(x_2(1-2x_1-2x_2)-x_1-x_2)+\lambda_1 x_1+\lambda_2 x_2.$$

The corresponding first-order conditions are

$$\text{1. } 2\alpha_1 x_1 + (2\alpha_2 + \alpha_1)x_2 = -\alpha_2 + \lambda_1;\ \ \text{2. } (2\alpha_2 + \alpha_1)x_1 + 4\alpha_2 x_2 = -\alpha_1 + \lambda_2;$$
$$\text{3. } \lambda_i x_i = 0,\ i = 1, 2;\ \ \text{4. } \alpha_i \geq 0,\ \lambda_i \geq 0,\ [\alpha_1\ \alpha_2\ \lambda_1\ \lambda_2] \neq 0,\ x_i \geq 0.$$

Assume $(x_1, x_2) = (0, \bar{x}_2)$ with $\bar{x}_2 > 0$ satisfies these conditions. Then $\lambda_2 = 0$, and from condition 2, we get $\alpha_1 = \alpha_2 = 0$. Substituting into condition 1 gives $\lambda_1 = 0$, which violates condition 4.
Similarly, $(x_1, x_2) = (\bar{x}_1, 0)$ with $\bar{x}_1 > 0$ also cannot satisfy the first-order conditions.
If $(x_1, x_2) = (\bar{x}_1, \bar{x}_2)$ with $\bar{x}_i > 0$, then $\lambda_i = 0$, which from conditions 1 and 2 implies $\alpha_i = 0$, again violating condition 4.
Finally, consider $(x_1, x_2) = (0, 0)$. By choosing, for example, $\lambda_1 = \alpha_2 > 0$ and $\lambda_2 = \alpha_1$, all first-order conditions are satisfied.

Step 2. We omit this step and proceed directly to Step 3.

Step 3. With $(\bar{x}_1, \bar{x}_2) = (0, 0)$, the corresponding restricted action sets are

$$\Gamma_1 = \{(x_1, x_2) \mid x_i \geq 0 \text{ and } J_2 \geq 0\} \text{ and } \Gamma_2 = \{(x_1, x_2) \mid x_i \geq 0 \text{ and } J_1 \geq 0\}.$$

Since $J_i(x_1, x_2) \leq 0$ on the constraint set, it follows that

$$0 = J_i(\bar{x}_1, \bar{x}_2) \geq J_i(x_1, x_2) \text{ for all } (x_1, x_2) \in \Gamma_i,\ i = 1, 2.$$

Thus, $(0, 0)$ is Pareto efficient.

Step 4. In this case, the Pareto frontier consists of a single point, namely $(0, 0)$.

c. Following Algorithm 6.17, we proceed with the following steps.
Step 1. Consider the Lagrange function

$$L(x_1, x_2, \alpha_1, \alpha_2) = \alpha_1(x_1(5-x_1-x_2)-x_1-x_2)+\alpha_2(x_2(10-2x_1-2x_2)-x_1-x_2)+\lambda_1 x_1+\lambda_2 x_2.$$

The corresponding first-order conditions are:

1. $2\alpha_1 x_1 + (2\alpha_2 + \alpha_1)x_2 = 4\alpha_1 - \alpha_2 + \lambda_1$; 2. $(2\alpha_2 + \alpha_1)x_1 + 4\alpha_2 x_2 = -\alpha_1 + 9\alpha_2 + \lambda_2$;
3. $\lambda_i x_i = 0,\ i = 1, 2$; 4. $\alpha_i \geq 0,\ \lambda_i \geq 0,\ [\alpha_1\ \alpha_2\ \lambda_1\ \lambda_2] \neq 0,\ x_i \geq 0$.

Assume that $(x_1, x_2) = (\bar{x}_1, \bar{x}_2)$, with $\bar{x}_i > 0$, satisfies these conditions. Then $\lambda_i = 0$, and conditions 1 and 2 reduce to

1b. $(2x_1 + x_2 - 4)\alpha_1 + (2x_2 + 1)\alpha_2 = 0$ and 2b. $(1 + x_1)\alpha_1 + (2x_1 + 4x_2 - 9)\alpha_2 = 0$.

In matrix form:

$$\begin{bmatrix} 2x_1 + x_2 - 4 & 2x_2 + 1 \\ 1 + x_1 & 2x_1 + 4x_2 - 9 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This system has the unique solution $\alpha_i = 0$ if and only if the determinant of the matrix is nonzero. By condition 4, the vector $[\alpha_1\ \alpha_2\ \lambda_1\ \lambda_2]$ cannot be zero. Hence, a solution exists only if

$$\det \begin{bmatrix} 2x_1 + x_2 - 4 & 2x_2 + 1 \\ 1 + x_1 & 2x_1 + 4x_2 - 9 \end{bmatrix} = (2x_1 + x_2 - 4)(2x_1 + 4x_2 - 9) - (2x_2 + 1)(1 + x_1) = 0,$$

or equivalently, $4(x_1 + x_2 - 10)(x_1 + x_2 - \frac{7}{4}) = 0$.
Morever, in this case the multipliers must satisfy

$$\alpha_1 = \frac{2x_2 + 1}{4 - 2x_1 - x_2}\alpha_2.$$

Since $\alpha_i > 0$ and $x_i > 0$, it follows that $4 - 2x_1 - x_2 > 0$. This immediately excludes the case $x_1 + x_2 = 10$.
Therefore, conditions 1-4 admit a solution $(x_1, x_2) = (\bar{x}_1, \bar{x}_2)$, with $\bar{x}_i > 0$, if and only if

$$\bar{x}_1 + \bar{x}_2 = \frac{7}{4}.$$

It is easily verified that the boundary points $(\frac{7}{4}, 0)$ and $(0, \frac{7}{4})$ also satisfy conditions 1-4 with above choice of parameters.
Next, consider the case $x_1 = 0$ and $x_2 > 0$. Then conditions 1 and 2 reduce to

$$\begin{bmatrix} x_2 - 4 & 2x_2 + 1 \\ 1 & 4x_2 - 9 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ 0 \end{bmatrix}.$$

Assuming the equations are linearly independent (the dependent case has, basically, already been discussed above), we obtain

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 4x_2 - 9 & -(2x_2 + 1) \\ -1 & x_2 - 4 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ 0 \end{bmatrix}.$$

From this, it follows that $\alpha_2 < 0$ if $\lambda_1 > 0$, which contradicts condition 4. If $\lambda_1 = 0$, then $\alpha_i = \lambda_i = 0$, again violating condition 4.

A similar argument shows that the cases $(0, \bar{x}_2)$, with $\bar{x}_2 \neq \frac{7}{2}$, and $(0, 0)$, do not satisfy conditions 1-4.

Step 2. It can be readily verified, by computing the second-order derivative of L, that for all $\alpha_i > 0$, every stationary point of this function is a saddle point. Therefore, this step does not allow us to draw any conclusions regarding the existence of Pareto solutions.

Step 3. Let $(x_1, x_2) = (\bar{x}_1, \bar{x}_2)$, where $\bar{x}_i \geq 0$ and $\bar{x}_1 + \bar{x}_2 = \frac{7}{4}$, be any point among the candidate Pareto efficient solutions.

Then,

$$\Gamma_1 = \{(x_1, x_2) \mid J_2(x_1, x_2) \geq J_2(\bar{x}_1, \bar{x}_2)\} = \{(x_1, x_2) \mid 2x_2(5-(x_1+x_2))-(x_1+x_2) \geq \frac{13}{2}\bar{x}_2 - \frac{7}{4}\}.$$

Using $\bar{x}_1 + \bar{x}_2 = \frac{7}{4}$, this can be rewritten as

$$\Gamma_1 = \{(x_1, x_2) \mid 2x_2(5 - (x_1 + x_2)) - (x_1 + x_2) \geq -\frac{13}{2}\bar{x}_1 + \frac{77}{8}\}.$$

Using this, for any $(x_1, x_2) \in \Gamma_1$, we obtain:

$$\begin{aligned} J_1(\bar{x}_1, \bar{x}_2) - J_1(x_1, x_2) &= \frac{13}{4}\bar{x}_1 - \frac{7}{4} - x_1(5 - (x_1 + x_2)) + x_1 + x_2 \\ &\geq \frac{1}{2}(x_1 + x_2) - x_2(5 - (x_1 + x_2)) + x_1 + x_2 - \frac{7}{4} \\ &\quad -x_1(5 - (x_1 + x_2)) + x_1 + x_2 \\ &= (x_1 + x_2 - \frac{7}{4})^2 \geq 0. \end{aligned}$$

Hence, all such points are Pareto efficient.

Step 4. The Pareto frontier is given by

$$(J_1, J_2) = (\frac{13}{4}x_1 - \frac{7}{4}, \frac{13}{2}(\frac{7}{4} - x_1) - \frac{7}{4}), \text{ where } 0 \leq x_1 \leq \frac{7}{4}.$$

Equivalently, it can be expressed as

$$J_2 = -2J_1 + 6\frac{1}{8}, \text{ with } J_1 \in [-\frac{7}{4}, \frac{9}{4}].$$

d. The problem we have to address is to determine all Pareto efficient solutions for the set of utility functions $U_i = \beta_i x_i^2$, subject to the budget constraint $x_1 + x_2 + x_3 \leq b$, where $x_i \geq 0$.
Following Algorithm 6.17, we proceed with the following steps.
Step 1. From the corresponding Lagrange function

$$L(x_1, x_2, x_3, \alpha_1, \alpha_2, \alpha_3, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \sum_{i=1}^{3} \alpha_i \beta_i x_i^2 + \sum_{i=1}^{3} \lambda_i x_i - \lambda_4 (x_1 + x_2 + x_3 - b)$$

we obtain the following first-order conditions:

$$\text{1. } 2\alpha_i \beta_i x_i + \lambda_i - \lambda_4 = 0,\ i = 1,2,3;\ \text{2. } \lambda_i x_i = 0,\ i = 1,2,3\ \lambda_4 (x_1 + x_2 + x_3 - b) = 0;$$
$$\text{3. } \alpha_i \geq 0,\ \lambda_i \geq 0,\ [\alpha\ \lambda] \neq 0,\ x_i \geq 0,\ x_1 + x_2 + x_3 \leq b.$$

If the budget is not fully utilized, the utility of one project can be increased without reducing that of the others. Therefore, at any Pareto efficient solution, the budget constraint must bind, i.e., $x_1 + x_2 + x_3 = b$.
Using this, the set of candidate Pareto efficient solutions is

$$\Delta := \{(\bar{x}_1, \bar{x}_2, b - \bar{x}_1 - \bar{x}_2),\ \text{where } \bar{x}_i \geq 0\}.$$

Step 2. Consider, for $\alpha_i > 0$, the maximization of

$$U := \sum_{i=1}^{3} \alpha_i \beta_i x_i^2$$

over the feasible set.
If, for example, $\alpha_1 \beta_1 > \alpha_j \beta_j$, for $j = 2, 3$, then the maximum is attained at $(x_1, x_2, x_3) = (b, 0, 0)$. Similarly, if $\alpha_1 \beta_1 = \alpha_2 \beta_2 > \alpha_3 \beta_3$, it is optimal to set $x_3 = 0$. Finally, also in case $U := \sum_{i=1}^{3} x_i^2$, one can verify that every corner point yields a higher value than any other point of Δ.
Thus, the corner points $(b, 0, 0)$, $(0, b, 0)$ and $(0, 0, b)$ are Pareto efficient solutions.
Step 3. Let $(\bar{x}_1, \bar{x}_2, \bar{x}_3) \in \Delta$. Define

$$\begin{aligned} \Gamma_1 &= \{(x_1, x_2, x_3) \in \Gamma \mid U_2(x) \geq U_2(\bar{x}) \text{ and } U_3(x) \geq U_3(\bar{x})\} \\ &= \{(x_1, x_2, x_3) \in \Gamma \mid x_2 \geq \bar{x}_2 \text{ and } x_3 \geq \bar{x}_3\}. \end{aligned}$$

Then,

$$\arg\max_{x \in \Gamma_1} U_1(x) = \tilde{x}_1 := b - \bar{x}_2 - \bar{x}_3,$$

which implies $U_1(x) \leq U_1(\bar{x})$ for all $x \in \Gamma_1$.
By analogous arguments, the same holds for U_2 and U_3. Hence, every point in Δ is Pareto efficient.

Step 4. The Pareto frontier is given by: $(\beta_1 x_1^2, \beta_2 x_2^2, \beta_3 x_3^2 \mid x_1 + x_2 + x_3 = b,\ x_i \geq 0,\ i = 1, 2, 3\}$.

EXERCISE 6.15 Assume player i wants to maximize $J_i(x_1, x_2)$ w.r.t. x_i, $i = 1, 2$, where in items a)-c) below, $(x_1, x_2) \in \Gamma := \{(x_1, x_2) \mid x_1^2 + x_2^2 \leq 1\}$. Determine the Nash equilibria in the following games.

a. $J_1(x_1, x_2) = x_1$, $J_2(x_1, x_2) = x_2$.

b. $J_1(x_1, x_2) = -x_1^2 + (x_2 - \frac{1}{2})^2$, $J_2(x_1, x_2) = x_1^2 - x_2^2$.

c. $J_1(x_1, x_2) = -x_1^2$, $J_2(x_1, x_2) = x_2^2$.

d. $J_1(x_1, x_2) = -(x_1 - x_2)^2$, $J_2(x_1, x_2) = -(x_1 - x_2 + \frac{1}{2})^2$, where $(x_1, x_2) \in \Gamma := \mathbb{R}^2$.

e. $J_1(x_1, x_2) = -\frac{1}{2}x_1^2 + x_1 x_2$, $J_2(x_1, x_2) = -\frac{1}{4}x_2^4 + x_1 x_2$, where $(x_1, x_2) \in \Gamma := \mathbb{R}^2$.

f. $J_1(x_1, x_2) = 2x_1^{\frac{1}{2}} x_2^{\frac{1}{2}} - x_1$, $J_2(x_1, x_2) = 3x_1^{\frac{1}{3}} x_2^{\frac{2}{3}} - 2x_2$, where $x_i > 0$.

g. $J_1(x_1, x_2) = (\alpha x_1^{\rho_1} + (1 - \alpha)x_2^{\rho_1})^{\frac{v}{\rho_1}} - \beta_1 x_1$, $J_2(x_1, x_2) = x_1^{1-\rho_2} x_2^{\rho_2} - \beta_2 x_2$, with $v \neq \rho_1$, $v \neq 1$, $\rho_1 < 1$, $\alpha \in (0, 1)$, $\rho_2 \in (0, 1)$, $\beta_2 > 0$, and $x_i > 0,\ i = 1, 2$.

h. $J_1(x_1, x_2) = -\frac{1}{3}x_1^3 + \frac{1}{2}(1 + x_2)x_1^2 - x_1 x_2$, $J_2(x_1, x_2) = -x_2^2 + 2\alpha x_1 x_2 + 2\alpha(1 - \alpha)x_2$, where $x_i \geq 0$ and $\alpha > 0$. Hint: first determine the optimal reaction curves of both players. Next, determine their intersection as a function of α.

ANSWER 6.15

a. Assume that $(\hat{x}_1, \hat{x}_2)$ is a Nash efficient solution. Then,

$$\hat{x}_1 = \arg\max_{x_1 \in \Gamma_1} x_1, \text{ where } \Gamma_1 = \{(x_1, \hat{x}_2) \in \Gamma\} = \{x_1 \mid x_1^2 \leq 1 - \hat{x}_2^2\}.$$

It follows that $\hat{x}_1 = \sqrt{1 - \hat{x}_2^2}$.

Similarly,

$$\hat{x}_2 = \arg\max_{x_2 \in \Gamma_2} x_2, \text{ where } \Gamma_2 = \{(\hat{x}_1, x_2) \in \Gamma\} = \{x_2 \mid x_2^2 \leq 1 - \hat{x}_1^2\}.$$

Hence, $\hat{x}_2 = \sqrt{1 - \hat{x}_1^2}$.

These results imply that the only candidates for Nash equilibria are points of the form

$$(x, \sqrt{1 - x^2}), \text{ with } 0 \leq x \leq 1.$$

Now consider an arbitrary point $(x, \sqrt{1 - x^2})$ with $0 \leq x \leq 1$. Then

$$x = \arg\max_{x_1 \in \Gamma_1} x_1, \text{ where } \Gamma_1 = \{(x_1, \sqrt{1 - x^2}) \in \Gamma\} = \{x_1 \mid x_1^2 \leq x^2\}.$$

Similarly,

$$\sqrt{1 - x^2} = \arg\max_{x_2 \in \Gamma_2} x_2, \text{ where } \Gamma_2 = \{(x, x_2) \in \Gamma\} = \{x_2 \mid x_2^2 \leq 1 - x^2\}.$$

Therefore, all points of the form $(x, \sqrt{1 - x^2})$, with $0 \leq x \leq 1$ are Nash (efficient) solutions.

b. It is straightforward to verify that $D_i(0,0) = 0,\ i = 1,2$. Moreover, $D_{11}J_1(x_1,0) \leq 0$ and $D_{22}J_2(0,x_2) \leq 0$.
Therefore, by Theorem 6.28, $(0,0)$ is a Nash efficient solution. This leads to the Nash equilibrium $(\frac{1}{4},0)$.
Finally, it follows directly from the definition that $(\frac{1}{4},0)$ is the unique Nash equilibrium.

c. Assume that $(\hat{x}_1,\hat{x}_2)$ is a Nash efficient solution. Then

$$\hat{x}_1 = \arg\max_{x_1\in\Gamma_1} -x_1^2, \text{ where } \Gamma_1 = \{(x_1,\hat{x}_2)\in\Gamma\} = \{x_1 \mid x_1^2 \leq 1-\hat{x}_2^2\}.$$

It follows that $\hat{x}_1 = 0$.
Next, consider

$$\hat{x}_2 = \arg\max_{x_2\in\Gamma_2} x_2^2, \text{ where } \Gamma_2 = \{(0,x_2)\in\Gamma\} = \{x_2 \mid x_2^2 \leq 1\}.$$

This implies that $\hat{x}_2 = 1$ or $\hat{x}_2 = -1$.
Therefore, the game has two Nash efficient solutions: $(0,1)$ and $(0,-1)$. There is a unique Nash equilibrium: $(0,1)$.

d. Assume that $(\hat{x}_1,\hat{x}_2)$ is a Nash equilibrium. Then

$$\hat{x}_1 = \arg\max_{x_1\in\Gamma_1} -(x_1-\hat{x}_2)^2, \text{ where } \Gamma_1 = \mathbb{R}.$$

It follows that $\hat{x}_1 = \hat{x}_2$. (1)
Next, consider

$$\hat{x}_2 = \arg\max_{x_2\in\Gamma_2} -(\hat{x}_1 - x_2 + \frac{1}{2})^2, \text{ where } \Gamma_2 = \mathbb{R}.$$

This implies that $\hat{x}_2 = \hat{x}_1 + \frac{1}{2}$. (2)
Since there is no pair $(\hat{x}_1,\hat{x}_2)$ that satisfies both (1) and (2) simultaneously, we conclude that this game has no Nash solution.

e. Assume that $(\hat{x}_1,\hat{x}_2)$ is a Nash efficient solution. Then,

$$\hat{x}_1 = \arg\max_{x_1\in\Gamma_1} -\frac{1}{2}x_1^2 - x_1\hat{x}_2, \text{ where } \Gamma_1 = \mathbb{R}.$$

From the firsr-order condition $D_1J_1(x) = 0$, it follows that necessarily $\hat{x}_1 = \hat{x}_2$. (1)
Next, consider

$$\hat{x}_2 = \arg\max_{x_2\in\Gamma_2} -\frac{1}{4}x_2^4 + \hat{x}_1x_2, \text{ where } \Gamma_2 = \mathbb{R}.$$

From $D_2J_2(x) = 0$, it follows that necessarily $\hat{x}_2^3 = \hat{x}_1$. (2)
The only three points that satisfy both (1) and (2) simultaneously are: $(-1,-1);\ (0,0)$ and $(1,1)$. Since $D_{11}J_1(x) < 0$, any point satisfying $D_1J_1(x) = 0$ provides a global maximum for J_1. To determine whether J_2 also achieves a global maximum at any of these candidate Nash efficient locations, we examine $J_2(-1,x_2)$, $J_2(0,x_2)$ and $J_2(1,x_2)$ on the domain $\mathbb{R}$. Notice that for any

fixed x_1,

$$D_{22}J_2(\bar{x}_1, x_2) = -3x_2^2 \leq 0 \text{ for all } x_2 \in \mathbb{R}.$$

Hence $J_2(\bar{x}_1, x_2)$ is concave for any $\bar{x}_1$. Therefore, all three points are Nash efficient.
This yields the two Nash equilibria: $(\frac{1}{2}, \frac{3}{4})$ and $(0, 0)$.

f. Given $J_1(x_1, x_2) = 2x_1^{\frac{1}{2}}x_2^{\frac{1}{2}} - x_1$, we have

$$D_1J_1(x_1, x_2) = x_1^{-\frac{1}{2}}x_2^{\frac{1}{2}} - 1.$$

Thus, $D_1J_1(\hat{x}, \hat{x}) = 0$ for all $\hat{x} > 0$.
Similarly, for $J_2(x_1, x_2)$, we find that $D_2J_2(x_1, x_2) = 0$ iff. $x_1 = x_2$. Therefore, all points of the form $(\hat{x}, \hat{x})$ are potentially Nash efficient locations.
It is straightforward to verify that both $D_{11}J_1(x_1, x_2)$ and $D_{22}J_2(x_1, x_2)$ are negative over their domains. Hence, $J_1(x_1, \hat{x})$ and $J_2(\hat{x}, x_2)$ are concave functions. Consequently, each function attains a global maximum at $\hat{x}$, implying that all points $(\hat{x}, \hat{x})$ are indeed Nash efficient locations.
This leads to the set of Nash solutions: $(\hat{x}, \hat{x})$ with $\hat{x} > 0$. In the (J_1, J_2) plane, this corresponds to all points lying on the line $J_1 = J_2$.

g. With the functions J_i as defined, the derivatives are

$$D_1J_1(x_1, x_2) = v\alpha x_1^{\rho_1 - 1}(\alpha x_1^{\rho_1} + (1-\alpha)x_2^{\rho_1})^{\frac{v}{\rho_1} - 1} - \beta_1$$

and

$$D_2J_2(x_1, x_2) = \rho_2 x_1^{1-\rho_2}x_2^{\rho_2 - 1} - \beta_2.$$

Thus, $D_2J_2(x_1, x_2) = 0$ iff. $x_1 = \mu_1 x_2$, with $\mu_1 := \left(\frac{\beta_2}{\rho_2}\right)^{\frac{1}{1-\rho_2}}$.
Substituting this relationship into $D_1J_1(x_1, x_2) = 0$ shows that there is a unique point satisfying both conditions:

$$(\hat{x}_1, \hat{x}_2) = (\mu_1 s, s), \text{ where } s = \left(\frac{\beta_1}{\alpha v \mu_1^{\rho_1 - 1}(\alpha\mu_1^{\rho_1} + (1-\alpha))^{\frac{v-\rho_1}{\rho_1}}}\right)^{\frac{1}{v-1}}.$$

It is straightforward to verify that both $D_{11}J_1(x_1, x_2) < 0$ and $D_{22}J_2(x_1, x_2) < 0$. Consequently, $J_1(x_1, \hat{x}_2)$ is concave in x_1 and $J_2(\hat{x}_1, x_2)$ is concave in x_2. Therefore, $J_1(x_1, \hat{x}_2)$ attains a global maximum at $\hat{x}_1$ and $J_2(\hat{x}_1, x_2)$ attains a global maximum at $\hat{x}_2$.
Hence, $(\hat{x}_1, \hat{x}_2)$ is a Nash efficient location. The corresponding Nash solution is $(J_1(\mu_1 s, s), J_2(\mu_1 s, s))$.

h. The first partial derivative of J_1 is

$$D_1J_1(x_1, x_2) = -x_1^2 + (1 + x_2)x_1 - x_2 = -(x_1 - 1)(x_1 - x_2).$$

Hence, $D_1J_1(x_1, x_2) = 0$ if either $x_1 = 1$ or $x_1 = x_2$. From the sign scheme of the derivative, we conclude:

1. J_1 has a maximum at $x_1 = x_2$ with $J_1(x_2, x_2) = x_2^2(\frac{1}{6}x_2 - \frac{1}{2})$ if $x_2 \geq 1$.
2. J_1 has a maximum at $x_1 = 1$ with with $J_1(1, x_2) = \frac{1}{6} - \frac{1}{2}x_2$ if $x_2 < 1$.
3. Additionally, J_1 attains a boundary maximum at $x_1 = 0$ with $J_1(0, x_2) = 0$.

Thus, the optimal reaction function of player 1 is: $x_1(x_2) = \begin{cases} 1, & \text{if } x_2 \in [0, \frac{1}{3}] \\ 0, & \text{if } x_2 \in [\frac{1}{3}, 1] \\ 0, & \text{if } x_2 \in [1, 3] \\ x_2, & \text{if } x_2 \geq 3 \end{cases}$

For player 2, the second partial derivative is

$$D_2J_2(x_1, x_2) = -2x_2 + 2\alpha x_1 + 2\alpha(1 - \alpha).$$

Hence, J_2 is maximized at $x_2 = \alpha x_1 + \alpha(1 - \alpha)$, if $x_1 + 1 - \alpha \geq 0$. Otherwise, the maximum occurs at the boundary $x_2 = 0$. Therefore, the optimal reaction function of player 2 is

$$x_2(x_1) = \begin{cases} \alpha x_1 + \alpha(1 - \alpha) & \text{if } \alpha \in [0, 1) \\ 0 & \text{if } \alpha \geq 1 \text{ and } x_1 \in [0, \alpha - 1] \\ \alpha x_1 + \alpha(1 - \alpha) & \text{if } x_1 \geq \alpha - 1 \end{cases}$$

Consequently,

1. If $\alpha \in [0, 1)$, the only candidate for the intersection of the reaction curves is $(x_1, x_2) = (1, 2\alpha - \alpha^2)$, provided $2\alpha - \alpha^2 \in [0, \frac{1}{3}]$. That is, if $\alpha \in [0, 1 - \sqrt{\frac{2}{3}}]$.
2. If $\alpha \in [1, 2]$, the only candidate is $(x_1, x_2) = (1, 0)$.
3. If $\alpha \geq 2$, solving $x_2 = \alpha x_1 + \alpha(1 - \alpha)$ gives three theoretical candidates:
   1. $(x_1, x_2) = (1, 2\alpha - \alpha^2)$ if $2\alpha - \alpha^2 \in [0, \frac{1}{3}]$, which unfortunately never applies;
   2. $(x_1, x_2) = (0, \alpha - \alpha^2)$ if $\alpha - \alpha^2 \in [\frac{1}{3}, 3]$, which also never applies; and
   3. $(x_1, x_2) = (\alpha, \alpha)$ if $\alpha \geq 3$.

In summary, the Nash efficient solutions are:

1. For $\alpha \in [0, 1 - \sqrt{\frac{2}{3}}]$: $(x_1, x_2) = (1, 2\alpha - \alpha^2)$.
2. For $\alpha \in [1, 2]$: $(x_1, x_2) = (1, 0)$.
3. For $\alpha \geq 3$: $(x_1, x_2) = (\alpha, \alpha)$.

For all other values of α, no Nash equilibrium exists.

EXERCISE 6.16 Reconsider Exercise 6.15. Determine the Pareto efficient solutions for items a-f. (Use, e.g., Algorithm 6.17).

ANSWER 6.16

First, notice that the action set Γ is convex. Moreover, it is compact in items a-c.

a. Note that $J_i(x_1, x_2) = x_i$, for $i = 1, 2$, are concave functions. Therefore, by Theorem 6.7, all Pareto efficient solutions can be obtained by solving

$$\max_{(x_1,x_2)\in\Gamma} \alpha x_1 + (1-\alpha)x_2, \text{ where } 0 \leq \alpha \leq 1.$$

Since Γ is compact, this maximization problem admits a solution, which is attained on the boundary of Γ. Using, e.g., the Lagrangian method (Theorem 6.6), one can verify that for every α, the maximum is attainted at

$$(x_1, x_2) = \frac{1}{\sqrt{\alpha^2 + (1-\alpha)^2}}(\alpha, 1-\alpha).$$

By Theorem 6.4 (or Theorem 6.6), it follows that all points of the form

$$(x_1, x_2) = \frac{1}{\sqrt{\alpha^2 + (1-\alpha)^2}}(\alpha, 1-\alpha),\ 0 < \alpha < 1,$$

are Pareto efficient.

For the boundary cases $\alpha = 1$ and $\alpha = 0$, we obtain the points $(1, 0)$ and $(0, 1)$, respectively. From the definition, it is straightforward to verify that both these points are also Pareto efficient.

Hence, the set of Pareto solutions is:

$$\{(x_1, x_2) = \frac{1}{\sqrt{\alpha^2 + (1-\alpha)^2}}(\alpha, 1-\alpha), 0 \leq \alpha \leq 1\} = \{(x_1, x_2) \mid x_1^2 + x_2^2 = 1,\ x_i \geq 0,\ i = 1, 2\}.$$

b. Following Algorithm 6.17, consider the with this problem related Lagrange function:

$$L(x_1, x_2, \alpha_1, \alpha_2, \lambda) := \alpha_1(-x_1^2 + (x_2 - \frac{1}{2})^2) + \alpha_2(x_1^2 - x_2^2) - \lambda(x_1^2 + x_2^2 - 1),\ \alpha_i \geq 0.$$

Following step 1 of the algorithm we consider the first-order conditions:

$$\begin{aligned}
DL_{x_1} = 2x_1(\alpha_2 - \alpha_1 - \lambda) &= 0;\\
DL_{x_2} = 2x_2(\alpha_1 - \alpha_2 - \lambda) - \alpha_1 &= 0;\\
\lambda(x_1^2 + x_2^2 - 1) &= 0;\\
x_1^2 + x_2^2 &\leq 1;\\
(\alpha_1\ \alpha_2\ \lambda) \neq (0\ 0\ 0),\ \lambda &\geq 0.
\end{aligned}$$

Elementary calculations show that the solution set of this system consists of:

1. $(x_1, x_2) = (0, 1)$, with $\alpha_1 = 2(\alpha_2 + \lambda)$, $\alpha_2 \geq 0$, $\lambda \geq 0$, and $(\alpha_2, \lambda) \neq (0, 0)$;
2. $(x_1, x_2) = (0, -1)$, with $\alpha_1 = \frac{2}{3}(\alpha_2 + \lambda)$, $\alpha_2 \geq 0$ $\lambda \geq 0$, and $(\alpha_2, \lambda) \neq (0, 0)$;
3. $(\pm\sqrt{1 - \bar{x}_2^2}, \bar{x}_2)$, where $\bar{x}_2 = -\frac{\alpha_1}{4\lambda}$, $\lambda > 0$, $\alpha_2 = \lambda + \alpha_1$, $0 \leq \alpha_1 \leq 4\lambda$.

Thus, the candidate Pareto efficient solutions are:

$$1.\ (x_1, x_2) = (0, 1);\ 2.\ (x_1, x_2) = (0, -1) \text{ and } 3.\ (\pm\sqrt{1 - \bar{x}_2^2}, \bar{x}_2),\ \text{where } -1 \leq \bar{x}_2 \leq 0.$$

In Step 2 of the algorithm, we determine which of these candidates are Pareto efficient.

1. Observe that $J_1(0, -\frac{1}{2}) > \frac{1}{4} = J_1(0, 1)$ and $J_2(0, -\frac{1}{2}) \geq -1 = J_2(0, 1)$. Hence, $(0, 1)$ is not Pareto efficient.
2. Observe that $J_1(x_1, x_2) < \frac{9}{4} = J_1(0, -1)$ on $\Gamma \backslash \{(0, -1)\}$. Thus, $(0, -1)$ is Pareto efficient.
3. This set consists of all points on the unit circle lying below the x_2 axis.
   Choose $\bar{\alpha}_1 > 0$, $0 < \bar{\lambda} < \frac{1}{4}\bar{\alpha}_1$, and define $\bar{\alpha}_2 := \bar{\lambda} + \bar{\alpha}_1$. Consider the function

   $$L(x_1, x_2, \bar{\alpha}_1, \bar{\alpha}_2, \bar{\lambda}) \text{ on } \Gamma.$$

   Since this function is continuous on the compact set Γ, it attains a global maximum. From the first-order conditions, the only candidates are

   $$(\pm\sqrt{1 - \bar{x}_2^2}, \bar{x}_2),\ \text{where } \bar{x}_2 = -\frac{\bar{\alpha}_1}{4\bar{\lambda}}.$$

   These points yield the same objective values and therefore are global maximizers. By Theorem 6.6 they are Pareto efficient.
   By varying $\bar{\alpha}_1$ and $\bar{\lambda}$, we obtain all points on the unit circle below the x_2 axis, except $(\pm 1, 0)$ and $(0, -1)$. From the second case considered above, we know that $(0, -1)$ is Pareto efficient. Finally, note that $J_2(x_1, x_2) \leq 1 = J_2(\pm 1, 0)$, so the points $(\pm 1, 0)$ are also Pareto efficient. These correspond to the same Pareto solution $(-\frac{3}{4}, 1)$.

c. Following Algorithm 6.17, consider the with this problem related Lagrange function:

$$L(x_1, x_2, \alpha, \lambda) := -\alpha_1 x_1^2 + \alpha_2 x_2^2 - \lambda(x_1^2 + x_2^2 - 1),\ \alpha_i \geq 0.$$

Following step 1 of the algorithm we consider the first-order conditions:

$$\begin{aligned} DL_{x_1} = 2x_1(\alpha_1 + \lambda) &= 0; \\ DL_{x_2} = 2x_2(\alpha_2 - \lambda) &= 0; \\ \lambda(x_1^2 + x_2^2 - 1) &= 0; \\ x_1^2 + x_2^2 &\leq 1; \\ (\alpha_1\ \alpha_2\ \lambda) \neq (0\ 0\ 0),\ \lambda &\geq 0. \end{aligned}$$

Some elementary calculations show that the solution set for above (in)equalities consists of:

1. $(x_1, x_2) = (0, 0)$;
2. $(x_1, x_2) = (0, \bar{x}_2)$, with $-1 \leq \bar{x}_2 \leq 1$, and
3. $(x_1, x_2) = (\bar{x}_1, 0)$, with $-1 \leq \bar{x}_1 \leq 1$.

To determine which solutions are Pareto efficient, observe the following:

1. $J_1(0,1) = 0 = J_1(0,0)$ and $J_2(0,1) > J_2(0,0)$. Hence, $(0,0)$ is not Pareto efficient.
2. $J_1(0,1) = J_1(0,\bar{x}_2)$ and $J_2(0,1) > J_2(0,\bar{x}_2)$ for all $\bar{x}_2 \neq \pm 1$. Therefore, all points $(0,\bar{x}_2)$ with $-1 < \bar{x}_2 < 1$ are not Pareto efficient. Moreover, $J_i(x_1,x_2) \leq J_i(0,\pm 1)$ for $i = 1,2$, so $(0,\pm 1)$ are Pareto efficient.
3. $J_1(\bar{x}_1, \sqrt{1-\bar{x}_1^2}) = J_1(\bar{x}_1,0)$ and $J_2(\bar{x}_1, \sqrt{1-\bar{x}_1^2}) > J_2(\bar{x}_1,0)$ if $\bar{x}_1 \neq \pm 1$. In addition, $J_1(\pm 1,0) < J_1(0,1)$ and $J_2(\pm 1,0) < J_2(0,1)$. Therefore, none of these points are Pareto efficient.

Consequently, there is just one Pareto solution: $(J_1, J_2) = (0,1)$.

d. Note that J_i, for $i = 1,2$, are concave and Γ is convex. Therefore, all Pareto solutions can be obtained by considering the maxima associated with this problem of the function

$$f(x_1,x_2,\alpha_1,\alpha_2) := -\alpha_1(x_1-x_2)^2 - \alpha_2(x_1 - x_2 + \frac{1}{2})^2, \text{ with } (\alpha_1,\alpha_2) \in \mathcal{A}.$$

Observe that f is concave and thus admits a global maximum for every $(\alpha_1,\alpha_2) \in \mathcal{A}$. The Lagrange function associated with this problem is:

$$L(x_1,x_2,\alpha_1,\alpha_2,\lambda) := -\alpha_1(x_1-x_2)^2 - \alpha_2(x_1 - x_2 + \frac{1}{2})^2 - \lambda(x_1^2 + x_2^2 - 1),\ \alpha_i \geq 0.$$

The first-order conditions are:

$$\begin{aligned} DL_{x_1} = \alpha_1(x_1-x_2) + \alpha_2(x_1 - x_2 + \frac{1}{2}) + \lambda x_1 &= 0; \\ DL_{x_2} = \alpha_1(x_1-x_2) + \alpha_2(x_1 - x_2 + \frac{1}{2}) - \lambda x_2 &= 0; \\ \lambda(x_1^2 + x_2^2 - 1) &= 0; \\ x_1^2 + x_2^2 &\leq 1; \\ (\alpha_1\ \alpha_2\ \lambda) \neq (0\ 0\ 0),\ \lambda &\geq 0. \end{aligned}$$

From $DL_{x_1} - DL_{x_2} = 0$, it follows that $\lambda(x_1 + x_2) = 0$.

Case i). $\lambda = 0$. In this case, we obtain

$$(x_1,x_2) = (\bar{x}_1, \bar{x}_1 + \frac{\alpha_2}{2(\alpha_1+\alpha_2)}), \text{ where } x_1^2 + x_2^2 \leq 1 \text{ and } (\alpha_1,\alpha_2) \neq (0,0).$$

Case ii). $\lambda > 0$. Then the solutions are

$$(x_1,x_2) = (\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}}) \text{ and } (x_1,x_2) = (-\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}}).$$

Straightforward calculations show that

$$f(\bar{x}_1,\bar{x}_2) = -\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2} > f(\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}}) > f(\sqrt{-\frac{1}{2}}, \sqrt{\frac{1}{2}}).$$

Hence, all points in Case i) are Pareto efficient. Moreover, the corresponding Pareto frontier is given by

$$(J_1, J_2) = -\frac{1}{4(\alpha_1+\alpha_2)^2}(\alpha_2^2, \alpha_1^2),\ (\alpha_1, \alpha_2) \in \mathcal{A}.$$

Equivalently, for $0 \leq \alpha \leq 1$, this can be written as

$$(J_1, J_2) = -\frac{1}{4}(\alpha^2, (1-\alpha)^2).$$

e. Following Algorithm 6.17, the Lagrange function associated with this problem is:

$$L(x_1, x_2, \alpha_1, \alpha_2) := \alpha_1(-\frac{1}{2}x_1^2 + x_1x_2) + \alpha_2(-\frac{1}{4}x_2^4 + x_1x_2),\ \alpha_i \geq 0.$$

From this the necessary conditions result:

$$\begin{aligned} DL_{x_1} = \alpha_1(-x_1 + x_2) + \alpha_2 x_2 &= 0; \\ DL_{x_2} = \alpha_1 x_1 + \alpha_2(-x_2^3 + x_1) &= 0; \\ (\alpha_1\ \alpha_2) &\neq (0\ 0). \end{aligned}$$

From this, we obtain $\det A = 0$, where $A = \begin{bmatrix} x_2 - x_1 & x_2 \\ x_1 & x_1 - x_2^3 \end{bmatrix}$, together with the conditions $\frac{x_1 - x_2}{x_2} > 0$ and $\frac{x_2^3 - x_1}{x_1} > 0$. That is,

$$\Delta = \{(x_1, x_2) \mid x_1^2 - x_1x_2^3 + x_2^4 = 0, \frac{x_1 - x_2}{x_2} > 0 \text{ and } \frac{x_2^3 - x_1}{x_1} > 0\}.$$

Next, consider the problem

$$\max_{(x_1, x_2) \in \Gamma_1} J_1(x_1, x_2),$$

where

$$\Gamma_1 = \{(x_1, x_2) \mid -\frac{1}{4}x_2^4 + x_1x_2 \geq -\frac{1}{4}\bar{x}_2^4 + \bar{x}_1\bar{x}_2, \text{ where } (\bar{x}_1, \bar{x}_2) \in \Delta\}.$$

It is straightforward to verify that the only candidate interior solution is $(x_1, x_2) = (0, 0)$, which is a saddle-point. Therefore, we turn to the constrained optimization problem on the boundary of Γ_1. The corresponding Lagrange function is:

$$L(x_1, x_2, \lambda) := -\frac{1}{2}x_1^2 + x_1x_2 + \lambda(-\frac{1}{4}x_2^4 + x_1x_2 + \frac{1}{4}\bar{x}_2^4 - \bar{x}_1\bar{x}_2).$$

The first-order conditions are:

$$\begin{aligned} DL_{x_1} = -x_1 + x_2 + \lambda x_2 &= 0; \\ DL_{x_2} = x_1 + \lambda(-x_2^3 + x_1) &= 0; \\ -\frac{1}{4}x_2^4 + x_1x_2 + \frac{1}{4}\bar{x}_2^4 - \bar{x}_1\bar{x}_2 &= 0. \end{aligned}$$

From $DL_{x_1} = 0$, it follows that $\lambda = \frac{x_1 - x_2}{x_2}$. Substituting this into $DL_{x_2} = 0$ shows that $(x_1, x_2) \in \Delta$. Moreover, the Hessian of the Lagrange function is given by

$$HL(x_1, x_2) = \begin{bmatrix} -1 & 1+\lambda \\ 1+\lambda & -3\lambda x_2^2 \end{bmatrix}.$$

One easily verifies that $HL(x_1, x_2)$ is negative definite for every $(x_1, x_2) \in \Delta$. Hence, every solution corresponds to a maximum.
Note that only for a fixed x_2, with $|x_2| \geq 2$, there exists a corresponding x_1 such that $(x_1, x_2) \in \Delta$. This x_1 is given by either

$$x_{11} = \frac{x_2^2}{2}(x_2 + \sqrt{x_2^2 - 4}) \text{ or } x_{12} = \frac{x_2^2}{2}(x_2 - \sqrt{x_2^2 - 4}).$$

One can easily verify that if $x_2 \leq -2$ then $J_1(x_{11}, x_2) < J_1(x_{12}, x_2)$, whereas for $x_2 \geq 2$ the reverse inequality holds. Therefore, J_1 attains a global maximum on Γ_1 at (x_{12}, x_2) when $x_2 \leq -2$, and at (x_{11}, x_2) when $x_2 \geq 2$.
A similar argument applies to the maximization of J_2 over the set

$$\Gamma_2 = \{(x_1, x_2) \mid -\frac{1}{2}x_1^2 + x_1 x_2 \geq -\frac{1}{2}\bar{x}_1^2 + \bar{x}_1\bar{x}_2, \text{ where } (\bar{x}_1, \bar{x}_2) \in \Delta\}.$$

Therefore, the set of Pareto efficient solutions is given by

$$\{(x_1, x_2) \mid x_1 = \frac{x_2^2}{2}(x_2 - \sqrt{x_2^2 - 4}) \text{ if } x_2 \leq 2 \text{ and } x_1 = \frac{x_2^2}{2}(x_2 + \sqrt{x_2^2 - 4}) \text{ if } x_2 \geq 2\}.$$

f. Following Algorithm 6.17, the Lagrange function associated with this problem is:

$$L(x_1, x_2, \alpha_1, \alpha_2) := \alpha_1(2\sqrt{x_1}\sqrt{x_2} - x_1) + \alpha_2(3x_1^{\frac{1}{3}}x_2^{\frac{2}{3}} - 2x_2),\ \alpha_i \geq 0.$$

From this, the necessary conditions are obtained as

$$\begin{aligned} DL_{x_1} = \alpha_1(\frac{\sqrt{x_2}}{\sqrt{x_1}} - 1) + \alpha_2 \frac{x_2^{\frac{2}{3}}}{x_1^{\frac{2}{3}}} &= 0; \\ DL_{x_2} = \alpha_1 \frac{\sqrt{x_1}}{\sqrt{x_2}} + 2\alpha_2(\frac{x_1^{\frac{1}{3}}}{x_2^{\frac{1}{3}}} - 1) &= 0; \\ (\alpha_1\ \alpha_2) &\neq (0\ 0). \end{aligned}$$

This implies that $\det A = 0$, where $A = \begin{bmatrix} \frac{\sqrt{x_2}}{\sqrt{x_1}} - 1 & \frac{x_2^{\frac{2}{3}}}{x_1^{\frac{2}{3}}} \\ \frac{\sqrt{x_1}}{\sqrt{x_2}} & 2(\frac{x_1^{\frac{1}{3}}}{x_2^{\frac{1}{3}}} - 1) \end{bmatrix}$, together with the conditions $x_2 < x_1$, $x_1 < x_2$ and $x_i > 0$.

Clearly, these conditions cannot be satisfied simultaneously. Therefore, $\Delta = \emptyset$, and no Pareto efficient solutions exist.

EXERCISE 6.17 Reconsider Exercise 6.15. Assume player 1 is the leader in all items. Determine the Stackelberg (efficient) solution(s) for items a-f.

ANSWER 6.17

a. For a given $\bar{x}_1$, player 2 aims to maximize x_2 subject to $x_2^2 \leq 1 - \bar{x}_1^2$. This yields

$$x_2^*(\bar{x}_1) = \sqrt{1 - \bar{x}_1^2}, \text{ provided that } \bar{x}_1^2 \leq 1.$$

Player 1 then seeks to maximize x_1 subject to $x_1^2 \leq 1$, which yields $x_1 = 1$ and consequently $x_2 = 0$ as the unique Stackelberg efficient solution.
Thus, the Stackelberg solution is $(1, 0)$.

b. For a given $\bar{x}_1$, player 2 aims to maximize $\bar{x}_1^2 - x_2^2$ subject to $x_2^2 \leq 1 - \bar{x}_1^2$. The maximum is attained at $x_2 = 0$, independently of player 1's action.
Player 1 then seeks to maximize $-x_1^2 + \frac{1}{4}$ subject to $x_1^2 \leq 1$, which yields $x_1 = 0$.
Hence, the unique Stackelberg efficient solution is $(0, 0)$, and the corresponding Stackelberg solution is $(\frac{1}{4}, 0)$.

c. For a given $\bar{x}_1$, player 2 aims to maximize x_2^2 subject to $x_2^2 \leq 1 - \bar{x}1^2$. The maximum is attained at $x_2 = \pm\sqrt{1 - \bar{x}_1^2}$, provided $\bar{x}_1^2 \leq 1$.
Player 1 then seeks to maximize $-x_1^2$ subject to $x_1^2 \leq 1$, which yields $x_1 = 0$.
Thus, the game has two Stackelberg efficient solutions: $(0, 1)$ and $(0, -1)$. There is a unique Stackelberg solution: $(0, 1)$.

d. For a given $\bar{x}_1$, player 2 aims to maximize $-(\bar{x}_1 - x_2 + \frac{1}{2})^2$. The maximum is attained at $x_2 = \bar{x}_1 + \frac{1}{2}$.
Next, player 1 aims to maximize $-\frac{1}{4}$, with $x_1 \in \mathbb{R}$. Hence, all points of the form $(x_1, x_1 + \frac{1}{2})$ provide Stackelberg efficient solutions, leading to the unique Stackelberg solution $(-\frac{1}{4}, 0)$.

e. For a given $\bar{x}_1$, player 2 aims to maximize $-\frac{1}{4}x_2^4 + \bar{x}_1 x_2$.
   1. First, consider the case $\bar{x}_1 \geq 0$. The maximum is attained at $x_2 = \bar{x}_1^{\frac{1}{3}}$.
      Player 1 then seeks to maximize $-\frac{1}{2}x_1^2 + x_1^{\frac{4}{3}}$, which is maximized at $x_1 = \left(\frac{4}{3}\right)^{\frac{3}{4}}$. This yields player 2's optimal response $x_2 = \left(\frac{4}{3}\right)^{\frac{1}{4}}$ with $J_1 = -\frac{1}{2}\left(\frac{4}{3}\right)^{\frac{6}{4}} + \frac{4}{3}$.
   2. If $\bar{x}_1 < 0$, player 2's optimal response is: $x_2 = -|\bar{x}_1|^{\frac{1}{3}}$, and player 1 maximizes at $-\left(\frac{4}{3}\right)^{\frac{3}{4}}$. This yields player 2's optimal reaction $x_2 = -\left(\frac{4}{3}\right)^{\frac{1}{4}}$ with $J_1 = -\frac{1}{2}\left(\frac{4}{3}\right)^{\frac{6}{4}} + \frac{4}{3}$.

   Therefore, there are two Stackelberg efficient solutions:

$$(x_1, x_2) = \left(\left(\frac{4}{3}\right)^{\frac{3}{4}}, \left(\frac{4}{3}\right)^{\frac{1}{4}}\right) \text{ and } (x_1, x_2) = \left(-\left(\frac{4}{3}\right)^{\frac{3}{4}}, -\left(\frac{4}{3}\right)^{\frac{1}{4}}\right).$$

   Both result in the same Stackelberg solution: $(-\frac{1}{2}\left(\frac{4}{3}\right)^{\frac{6}{4}} + \frac{4}{3}, 1)$.

f. For a given $\bar{x}_1$, player 2 seeks to maximize $J_2(x_2) = 3\bar{x}_1^{\frac{1}{3}} x_2^{\frac{2}{3}} - 2x_2$.
Applying the first- and second-order conditions shows that the maximum is attained at $x_2 = 8\bar{x}_1$. Player 1 then aims to maximize $J_1(x_1) = (2\sqrt{8} - 1)x_1$. Since this is a linear function, it has no finite maximum. Therefore, no Stackelberg efficient solutions exist for this game.

EXERCISE 6.18 Reconsider Exercise 6.15, item d. Assume player 1 is the leader in this game who announces to use an affine incentive strategy w.r.t. the actions played by the follower. Determine an incentive strategy within this class of strategies which will maximize the objective of the leader.

ANSWER 6.18

So, player 1 announces to player 2 that (s)he will use the action $x_1 = \alpha x_2 + \beta$. Player 1 chooses the parameters α and β in order to maximize his/her own objective (which is 0 in this case). It is clear that the optimal choice for player 1 is to set α and β such that $x_2 = x_1$, i.e.: $\alpha = 1$ and $\beta = 0$.
Consequently, any action of the form (x, x) with $x \in \mathbb{R}$ is incentive Stackelberg efficient, and this leads to the unique incentive Stackelberg solution $(0, -\frac{1}{4})$.

EXERCISE 6.19 Consider player L who wants to maximize $J_L := -(x_L + x_F + 1)^2(x_L^2 + 1)^2$ w.r.t. x_L and player F who wants to optimize $J_F := -(x_F^2 + 1)^2(x_F + x_L)^2$ w.r.t. x_F. Assume player L is the leader in this game.

a. Determine the Stackelberg solution in this game if the follower wants to maximize his objective.
b. Is it possible to determine a Stackelberg solution in this game if the follower wants to minimize his objective?
c. Next consider the case player F is the leader and player L is the follower in the game and player L wants to maximize his objective. Determine the Stackelberg solution in this game.
d. Compare the answers in item a. and c. Is there a preference for leadership in this game?
e. Show, using Corollary 6.29, that the only Nash efficient candidate solutions are $(x_L, x_F) = (-2, 1)$ and $(x_L, x_F) = (-1, 1)$.
f. Show, using the definition, that both strategies $(x_L, x_F) = (-2, 1)$ and $(x_L, x_F) = (-1, 1)$ are not Nash efficient.

ANSWER 6.19

a. For a fixed value of x_L, consider the maximization of J_F. It is clear that the maximum occurs at $x_F = -x_L$, which gives the follower a value of 0, regardless of the leader's choice of action.
Next, the leader must maximize $J_L = -(x_L^2 + 1)^2$ with respect to x_L. Clearly, choosing $x_L = 0$ yields the maximum value of -1 for the leader.
b. Next, consider the minimization of J_F for a fixed value of x_L. Clearly, by choosing x_L to be arbitrarily large (or small), but different from $-x_F$, the follower can enforce his/her objective to an arbitrarily negative value. So, a Stackelberg solution does not exist.

c. For a fixed value of x_F, consider the maximization of J_L. It is clear that the maximum is achieved at $x_L = -x_F - 1$, which gives $J_L = 0$, independent of the action chosen by player F.
Next, player F must maximize $J_F = -(x_F^2 + 1)^2$ with respect to x_F. Clearly, choosing $x_F = 0$ yields the maximum value of -1 for player F.

d. If L is the leader, the Stackelberg solution is $(-1, 0)$. If F is the leader, the Stackelberg solution is $(0, -1)$. Thus, in this game, it is actually better for a player not the be the leader.

e. Straightforward differentiation of J_L with respect to x_L and J_F with respect to x_F yields the following first-order conditions:

$$\begin{aligned} -2(x_L + x_F + 1)(x_L^2 + 1)[3x_L^2 + 2x_L x_F + 2x_L + 1] &= 0; \\ -2(x_L + x_F)(x_F^2 + 1)[3x_F^2 + 2x_L x_F + 1] &= 0. \end{aligned}$$

Elementary calculations then show that the only solutions satisfying both equations are: $(x_L, x_F) = (-2, 1)$ and $(x_L, x_F) = (-1, 1)$.

f. First, note that $J_i(x_L, x_F) \leq 0$ for all $(x_L, x_F) \in \mathbb{R}^2$ and $i = L, F$. Furthermore,

$$J_L(-2, 1) = J_F(-1, 1) = 0 \text{ and } J_F(-2, 1) = J_L(-1, 1) = -4.$$

Thus, $J_L(x_L, 1) \leq J_L(-2, 1)$ for all $x_L \in \mathbb{R}$. However,

$$J_F(-2, x_F) = -(x_F^2 + 1)(x_F + 1)^2 > J_F(-2, 1) \text{ at } x_F = 2,$$

so, $(-2, 1)$ is not Nash efficient for player F.
Similarly, we have $J_F(-1, x_F) \leq J_F(-1, 1)$ for all $x_F \in \mathbb{R}$. But $J_L(x_L, 1) > J_L(-1, 1)$ at, e.g., $x_L = -2$, so $(-1, 1)$ is not Nash efficient for player L.

Advanced, Deepening and Broadening Exercises

EXERCISE 6.20 Consider a player who wants to maximize simultaneously both $J_1(x) := 2x^2$ and $J_2(x) := -x^2$.

a. Use Algorithm 6.17 to find the Pareto frontier.

b. Consider the general problem setting from Definition 6.3 to find the Pareto frontier. Show that even in case there exist $\alpha_i > 0$ such that $x^* \in \Gamma$ minimizes $\sum_{i=1}^N J_i(x)$, this doesn't exclude x^* to be Pareto efficient.

ANSWER 6.20

a. Following Algorithm 6.17, the steps are as follows:
Step 1. Define

$$L(x, \alpha) := \alpha J_1(x) + (1 - \alpha) J_2(x), \text{ with } 0 \leq \alpha \leq 1.$$

The necessary condition for optimality is $2x(3\alpha - 1) = 0$.

Thus, if $\alpha \neq \frac{1}{3}$, $x = 0$ is a candidate efficient solution. If $\alpha = \frac{1}{3}$, then all $x \in \mathbb{R}$ are candidate efficient solutions.

Step 2. If $0 < \alpha < \frac{1}{3}$, $x = 0$ maximizes $L(x, \alpha)$, so $x = 0$ is Pareto efficient.

If $\frac{1}{3} < \alpha < 1$, $L(x, \alpha)$ has no maximum.

Step 3. For any $0 \neq \bar{x} \in \mathbb{R}$, define

$$\Gamma_1 := \{x \in \mathbb{R} \mid -x^2 \geq -\bar{x}^2\} = [0, \bar{x}] \text{ and } \Gamma_2 := \{x \in \mathbb{R} \mid 2x^2 \geq 2\bar{x}^2\} = [\bar{x}, \infty) \cup (-\infty, -\bar{x}].$$

Notice that $J_1(\bar{x}) = 2\bar{x}^2 \geq J_1(x)$, for all $x \in \Gamma_1$, and $J_2(\bar{x}) = -\bar{x}^2 \geq J_1(x)$ for all $x \in \Gamma_2$.

Hence, every $0 \neq \bar{x} \in \mathbb{R}$ is Pareto efficient.

Step 4. Combining Steps 2 and 3, all $x \in \mathbb{R}$ are Pareto efficient. This gives the Pareto frontier

$$(J_1, J_2) = (2x^2, -x^2),\ x \in \mathbb{R}, \text{ or equivalently, } J_1 = -2J_2.$$

b. Consider the example above with $\alpha \in (\frac{1}{3}, 1)$. Then $x^* = 0$ corresponds to a global minimum, yet, as we showed in item a, it is also Pareto efficient.

EXERCISE 6.21 *(Example where an equilibrium exists that is locally Nash)*

Consider the 2-player non-cooperative game where the two players want to maximize $J_1(x_1, x_2) := -\frac{1}{2}x_1^2 + x_1x_2$ w.r.t. x_1 and $J_2(x_1, x_2) := -\frac{1}{3}x_2^3 + x_1x_2$ w.r.t. x_2, respectively.

a. Show that the only potential Nash equilibria are $(0, 0)$ and $(1, 1)$.

b. Show that both points are not Nash equilibria.

c. Show, there is a neighborhood of $(1, 1)$ where this point qualifies as a Nash equilibrium.

ANSWER 6.21

a. Note that $D_1J_1(x_1, x_2) = -x_1 + x_2$ and $D_{11}J_1(x_1, x_2) = -1 < 0$.

Hence, the reaction curve of player 1, given an action $\bar{x}_2$ by player 2, is: $x_1 = \bar{x}_2$.

Next, observe that $D_2J_2(x_1, x_2) = -x_2^2 + x_1$. Therefore, given an action $\bar{x}_1$ by player 1, a necessary condition for a Nash equilibrium is $x_2^2 = \bar{x}_1$.

The intersection of player 1's reaction curve with the only feasible reaction of player 2 yields the two points: $(0, 0)$ and $(1, 1)$.

b. Note that $J_2(0, x_2) = -\frac{1}{3}x_2^3$, which does not attain a maximum at $x_2 = 0$. Therefore, $(0, 0)$ is not a Nash equilibrium.

Similarly, $J_2(1, x_2) = -\frac{1}{3}x_2^3 + x_2$ does not have a global maximum at $x_2 = 1$. Hence, this point also does not qualify as a Nash equilibrium.

c. From item b., it is clear that if we restrict, for example, x_2 to the interval $(0, \infty)$, then $J_2(1, x_2)$ attains a global maximum on this set. Therefore, if we restrict the game to the set

$$\Gamma := \{(x_1, x_2) \mid x_1 \in \mathbb{R} \text{ and } x_2 > 0\},$$

the game has a Nash equilibrium within Γ.

EXERCISE 6.22 *(Symmetric game with infinite number of asymmetric Nash efficient actions)*
Consider the cost function $J(x_1, x_2) = x_1^2 x_2^2$, where player $i's$ control instrument is $x_i,\ i = 1, 2$. Assume both players want to minimize $J(x_1, x_2)$.

a. Show that this game has an infinite number of asymmetric Nash efficient actions.

b. Show that this game has a unique set of symmetric Nash efficient actions.

c. Show that this game has a unique Nash equilibrium.

ANSWER 6.22

a. Note that $D_1 J(x_1, x_2) = 2x_1 x_2^2$. Hence, $D_1(x_1, x_2) = 0$ at $(0, \bar{x}_2)$ for all $\bar{x}_2 \in \mathbb{R}$, and at $(\bar{x}_1, 0)$ for all $\bar{x}_1 \in \mathbb{R}$.
Similarly, we obtain $D_2 J(x_1, x_2) = 2x_1^2 x_2 = 0$ at these points.
Furthermore, $D_{11} J(x_1, x_2) = 2x_2^2 > 0$ for any $x_2 \neq 0$, and $D_{22} J(x_1, x_2) = 2x_1^2 > 0$ for any $x_1 \neq 0$. Therefore, by Corollary 6.29, the game admits a Nash equilibrium at every point $(0, \bar{x}_2)$, with $\bar{x}_2 \neq 0$, and $(\bar{x}_1, 0)$ with $\bar{x}_1 \neq 0$. Clearly, all these Nash efficient points are not symmetric.

b. According to Corollary 6.29, the only additional point at which a Nash equilibrium may occur is $(0, 0)$. Note that $J(0, 0) \leq J(x_1, 0)$ for all $x_1 \in \mathbb{R}$, and $J(0, 0) \leq J(0, x_2)$ for all $x_2 \in \mathbb{R}$. Therefore, by definition, the game has a Nash equilibrium at $(0, 0)$. Clearly this action profile is symmetric.

c. In items a. and b. we determined all Nash efficient actions. All of them result in the Nash equilibrium $(J_1, J_2) = (0, 0)$.

EXERCISE 6.23 *(Zero-sum game)*
Consider a two-person game where player 1 wants to maximize $J(x_1, x_2)$ w.r.t. $x_1 \in \mathbb{R}^n$ and player two wants to minimize $J(x_1, x_2)$ w.r.t. $x_2 \in \mathbb{R}^m$, where $(x_1, x_2) \in \Gamma \subset \mathbb{R}^{n+m}$.

a. Reformulate this problem into the zero-sum framework where both players want to maximize an objective function $J_i(x_1, x_2)$ w.r.t. $x_i,\ i = 1, 2$, respectively, where $J_1 + J_2 = 0$.

b. Show that $(\hat{x}_1, \hat{x}_2)$ is Nash efficient for the zero-sum game with objective functions $J_i,\ i = 1, 2$, if and only if $(\hat{x}_1, \hat{x}_2)$ is a saddle-point equilibrium for J.

c. Show that if Γ is a compact and convex set, and $J(x_1, x_2)$ is a continuous function which is moreover strictly concave in x_1 (for fixed feasible values of x_2) and strictly convex in x_2 (for fixed feasible values of x_1), the game has a Nash solution.

d. Assume $x_i \in \Gamma_i$, where $\Gamma_i,\ i = 1, 2$, are convex subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Show, using item b., that if $(\hat{x}_1, \hat{x}_2)$ and $(\bar{x}_1, \bar{x}_2)$ are Nash efficient for above mentioned zero-sum game, also $(\hat{x}_1, \bar{x}_2)$ is Nash efficient for this game.

e. Reconsider item c., with $\Gamma = \Gamma_1 \times \Gamma_2$, where $\Gamma_i,\ i = 1, 2$, are compact and convex subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Show that the zero-sum game has a unique Nash solution that is attained at a unique Nash efficient location. Hint: assume there are two different Nash efficient solutions, use next item d. to show that the value of J has to coincide at these points.

f. Consider $\Gamma_i = [0,1],\ i = 1,2$, and $J(x_1,x_2) = (x_1 - x_2)^2$. Show that the zero-sum game, where player 1 wants to maximize J w.r.t. x_1 and player 2 wants to minimize J w.r.t. x_2 has not a Nash equilibrium.

ANSWER 6.23

a. Let $J_1(x_1,x_2) := J(x_1,x_2)$ and $J_2(x_1,x_2) := -J(x_1,x_2)$. Then the problem is equivalent to each player i maximizing $J_i(x_1,x_2)$ w.r.t. x_i, for $i = 1,2$, respectively. Moreover, it is clear that $J_1 + J_2 = 0$.

b. Assume that $(\hat{x}_1,\hat{x}_2)$ is Nash efficient for the zero-sum game with objective functions $J_i,\ i = 1,2$. Then

   i. $J_1(x,\hat{x}_2) \leq J_1(\hat{x}_1,\hat{x}_2), \forall (x,\hat{x}_2) \in \Gamma_1(\hat{x}_2)$. Or, equivalently, $J(x,\hat{x}_2) \leq J(\hat{x}_1,\hat{x}_2), \forall (x,\hat{x}_2) \in \Gamma_1(\hat{x}_2)$.

   ii) $J_2(\hat{x}_1,y) \leq J_2(\hat{x}_1,\hat{x}_2), \forall (\hat{x}_1,y) \in \Gamma_2(\hat{x}_1)$. Or, equivalently, $-J(\hat{x}_1,y) \leq -J(\hat{x}_1,\hat{x}_2), \forall (\hat{x}_1,y) \in \Gamma_2(\hat{x}_1)$. That is: $J(\hat{x}_1,\hat{x}_2) \leq J(\hat{x}_1,y), \forall (\hat{x}_1,y) \in \Gamma_2(\hat{x}_1)$.

   Combining i) and ii) we obtain:

$$J(x,\hat{x}_2) \leq J(\hat{x}_1,\hat{x}_2) \leq J(\hat{x}_1,y), \text{ for all } (x,\hat{x}_2) \in \Gamma_1(\hat{x}_2) \text{ and all } (\hat{x}_1,y) \in \Gamma_2(\hat{x}_1).$$

   The converse implication follows by reversing the above arguments.

c. With J_i as defined in item a., each J_i is continuous and strictly concave with respect to x_i for all $x \in \Gamma_i(x_{-i})$ for $i = 1,2$.

   Thus, all assumptions of Theorem 6.30 are satisfied, and the result follows directly from this theorem.

d. Since $(\hat{x}_1,\hat{x}_2)$ and $(\bar{x}_1,\bar{x}_2)$ are Nash efficient, it follows from item b. that:

$$J(x,\hat{x}_2) \leq J(\hat{x}_1,\hat{x}_2) \leq J(\hat{x}_1,y) \text{ and } J(x,\bar{x}_2) \leq J(\bar{x}_1,\bar{x}_2) \leq J(\bar{x}_1,y),$$

   for all $x \in \Gamma_1$ and all $y \in \Gamma_2$. In particular, this implies:

$$J(\bar{x}_1,\hat{x}_2) \leq J(\hat{x}_1,\hat{x}_2) \leq J(\hat{x}_1,\bar{x}_2) \text{ and } J(\hat{x}_1,\bar{x}_2) \leq J(\bar{x}_1,\hat{x}_2) \leq J(\bar{x}_1,\hat{x}_2).$$

   Combining these inequalities, we obtain:

$$J(\bar{x}_1,\hat{x}_2) \leq J(\hat{x}_1,\hat{x}_2) \leq J(\hat{x}_1,\bar{x}_2) \leq J(\bar{x}_1,\bar{x}_2) \leq J(\bar{x}_1,y)$$

   and

$$J(\bar{x}_1,\hat{x}_2) \geq J(\bar{x}_1,\bar{x}_2) \geq J(\hat{x}_1,\bar{x}_2) \geq J(\hat{x}_1,\hat{x}_2) \geq J(x,\hat{x}_2),$$

   for all $x \in \Gamma_1$ and all $y \in \Gamma_2$. Hence,

$$J(x,\hat{x}_2) \leq J(\bar{x}_1,\hat{x}_2) \leq J(\bar{x}_1,y), \text{ for all} x \in \Gamma_1 \text{ and all } y \in \Gamma_2.$$

That is, $(\bar{x}_1, \hat{x}_2)$ is a saddle-point equilibrium of J. Therefore, by item b., $(\bar{x}_1, \hat{x}_2)$ is Nash efficient for the corresponding zero-sum game.

e. From item c., we know that the game has at least one Nash equilibrium. Now suppose that the game has two Nash efficient locations, $(\hat{x}_1, \hat{x}_2)$ and $(\bar{x}_1, \bar{x}_2)$. Then, by item d., the locations $(\bar{x}_1, \hat{x}_2)$ and $(\hat{x}_1, \bar{x}_2)$ are also Nash efficient. Hence, by item b., we have:

$$J(x_1, \hat{x}_2) \leq J(\bar{x}_1, \hat{x}_2) \leq J(\bar{x}_1, y_1), \text{ for all } x_1 \in \Gamma_1, y_1 \in \Gamma_2$$

and

$$J(x_2, \bar{x}_2) \leq J(\hat{x}_1, \bar{x}_2) \leq J(\hat{x}_1, y_2), \text{ for all } x_2 \in \Gamma_1, y_2 \in \Gamma_2.$$

Now choose $x_1 = \hat{x}_1$, $y_1 = \bar{x}_2$, $x_2 = \bar{x}_1$ and $y_2 = \hat{x}_2$ in the above inequalities. It follows that

$$J(\hat{x}_1, \hat{x}_2) = J(\hat{x}_1, \bar{x}_2) = J(\bar{x}_1, \hat{x}_2) = J(\bar{x}_1, \bar{x}_2).$$

Thus, the Nash equilibrium is unique. Moreover, since J is strictly concave in x_1 and strictly convex in x_2, the Nash efficient locations must coincide as well.

f. For a fixed $x_2 \in [0, \frac{1}{2})$, the best response of player 1 is $x_1 = 1$; for $x_2 \in (\frac{1}{2}, 1]$ the best response of player 1 is $x_1 = 0$; and for $x_2 = \frac{1}{2}$, player 1 is indifferent between $x_1 = 0$ and $x_1 = 1$.
The best response of player 2, for any fixed x_1, is $x_2 = x_1$.
Since the best-response functions do not intersect, this zero-sum game has no Nash equilibrium.

EXERCISE 6.24 *(Symmetric zero-sum game)*
Consider the objective function $J(x_1, x_2)$, with $x_i \in X \subset \mathbb{R}^m$. Assume $J(x_1, x_2) = -J(x_2, x_1)$. Assume player 1 wants to maximize $J(x_1, x_2)$ w.r.t. x_1 and player 2 wants to minimize $J(x_1, x_2)$ w.r.t. x_2.

a. Show that this game is equivalent to a symmetric zero-sum game where player i wants to maximize $J_i(x_1, x_2)$ w.r.t. $x_i,\ i = 1, 2$, where $J_1(x_1, x_2) + J_2(x_1, x_2) = 0$ and $J_1(x_1, x_2) = J_2(x_2, x_1)$.
b. Show that if this game has a set of Nash efficient actions $(\hat{x}_1, \hat{x}_2)$ also $(\hat{x}_2, \hat{x}_1)$ is a set of Nash efficient actions.
c. Show that for any Nash efficient strategy $(\hat{x}_1, \hat{x}_2)$, $J_i(\hat{x}_1, \hat{x}_2) = 0,\ i = 1, 2$.
d. Show that if this game has a set of Nash efficient actions $(\hat{x}_1, \hat{x}_2)$ also $(\hat{x}_1, \hat{x}_1)$ is a set of Nash efficient actions.

ANSWER 6.24

a. Let $J_1(x_1, x_2) := J(x_1, x_2)$ and $J_2(x_1, x_2) := -J(x_1, x_2)$. Then $J_1(x_1, x_2) + J_2(x_1, x_2) = 0$. Moreover, $J_1(x_1, x_2) = J(x_1, x_2) = -J(x_2, x_1) = J_2(x_2, x_1)$.
b. Assume that $(\hat{x}_1, \hat{x}_2)$ is Nash efficient. Then,

$$\text{(i) } J_1(\hat{x}_1, \hat{x}_2) \geq J_1(x, \hat{x}_2), \text{ for all } x \in X \text{ and (ii) } J_2(\hat{x}_1, \hat{x}_2) \geq J_2(\hat{x}_1, x), \text{ for all } x \in X.$$

Hence,

$$J_1(\hat{x}_2, \hat{x}_1) = J_2(\hat{x}_1, \hat{x}_2) \geq J_2(\hat{x}_1, x) = J_1(x, \hat{x}_1), \text{ for all } x \in X$$

and

$$J_2(\hat{x}_2, \hat{x}_1) = J_1(\hat{x}_1, \hat{x}_2) \geq J_1(x, \hat{x}_2) = J_2(\hat{x}_2, x), \forall x \in X.$$

That is, $(\hat{x}_2, \hat{x}_1)$ is a set of Nash efficient actions too.

c. By definition, $J_1(\hat{x}_1, \hat{x}_2) \geq J_1(x, \hat{x}_2)$ for all $x \in X$. (i)
Since $J_1(x, x) + J_2(x, x) = 0$ for all $x \in X$, it follows that $J_1(x, x) = 0$ for all $x \in X$. In particular, taking $x = \hat{x}_2$ in (i) gives $J_1(\hat{x}_1, \hat{x}_2) \geq 0$.
Similarly, from item b.(ii), we obtain $J_2(\hat{x}_1, \hat{x}_2) \geq 0$.
However, since $J_1(\hat{x}_1, \hat{x}_2) + J_2(\hat{x}_1, \hat{x}_2) = 0$, it follows that $J_1(\hat{x}_1, \hat{x}_2) = J_2(\hat{x}_1, \hat{x}_2) = 0$.

d. Assume that $(\hat{x}_1, \hat{x}_2)$ is a Nash efficient set of actions. Then, by item c., $J_i(\hat{x}_1, \hat{x}_2) = 0$ for $i = 1, 2$.
Moreover, as noted in the proof of item c., $J_1(x, x) = 0$ for all $x \in X$.
Now, suppose that $(\hat{x}_1, \hat{x}_1)$ is not Nash efficient. Then there exists some $x \in X$ such that

$$J_1(x, \hat{x}_1) > J_1(\hat{x}_1, \hat{x}_1) = 0.$$

However,

$$J_1(x, \hat{x}_1) = J_2(\hat{x}_1, x) \leq J_2(\hat{x}_1, \hat{x}_2) = 0,$$

which clearly contradits above observation that $J_1(x, \hat{x}_1) > 0$. Therefore, for all $x \in X$, we must have

$$J_1(x, \hat{x}_1) \leq J_1(\hat{x}_1, \hat{x}_1).$$

Similarly, the inequality

$$J_2(\hat{x}_1, x) \leq J_2(\hat{x}_1, \hat{x}_1) \text{ for all } x \in X,$$

follows. This shows that $(\hat{x}_1, \hat{x}_1)$ is indeed Nash efficient.

EXERCISE 6.25 *(Proof Theorem 6.6)*

Consider the notation in Theorem 6.6. Assume there exist $\hat{\alpha}_i \in (0, 1)$, with $\sum_{i=1}^{N} \hat{\alpha}_i = 1$, $\hat{\lambda}_i \geq 0$, $\hat{\mu}_i$ and $\hat{x} \in \Gamma$ such that $\hat{\lambda}_i(g_i(\hat{x}) - b_i) = 0$ and $L(x, \hat{\lambda}, \hat{\mu}) \leq L(\hat{x}, \hat{\lambda}, \hat{\mu}), \forall x \in \Gamma$.
Use Theorem 6.4 to show that $\hat{x}$ is Pareto efficient.

ANSWER 6.25

Since

$$L(x, \hat{\lambda}, \hat{\mu}) \leq L(\hat{x}, \hat{\lambda}, \hat{\mu}), \ \forall x \in \Gamma,$$

it follows from this inequality, together with the other assumptions, that

$$\sum_{i=1}^{N} \hat{\alpha}_i J_i(x) \leq \sum_{i=1}^{N} \hat{\alpha}_i J_i(\hat{x}),\ \forall x \in \Gamma.$$

In other words, $\hat{x}$ maximizes $\sum_{i=1}^{N} \hat{\alpha}_i J_i(x)$ on Γ. Therefore, by Theorem 6.4, $\hat{x}$ is Pareto efficient.

EXERCISE 6.26 *(Proof Theorem 6.7)*
In this exercise you are asked to provide a proof for Theorem 6.7.
So, assume the set of constraint instruments Γ is convex; the objective function $J_i(x)$, player i wants to maximize, is concave, $i = 1, \cdots, N$; and $\hat{x}$ is Pareto efficient.

a. Show that there is no solution x satisfying simultaneously the set of inequalities $J_i(x) - J_i(\hat{x}) \geq 0,\ i = 1, \cdots, N$, where at least one of the inequalities is strict.
b. Use Lemma 5.8 to show there exist $\alpha \in \mathcal{A}$ such that $\sum_{i=1}^{N} \alpha_i J_i(x) \leq \sum_{i=1}^{N} \alpha_i J_i(\hat{x})$.

ANSWER 6.26

a. Since $\hat{x}$ is Pareto effient, there is no solution x such that $J_i(x) \geq J_i(\hat{x})$ for all i, with at least one of the inequalities being strict.
b. Since $J_i(x) - J_i(\hat{x})$ is concave, the set

$$S := \{x \mid J_i(x) - J_i(\hat{x}) > 0,\ i = 1, \cdots, N\}$$

is empty (see item a). Moreover, since $\mathcal{U}$ is convex, it follows from Lemma 5.8 that there exists $\alpha \in \mathcal{A}$ such that

$$\sum_{i=1}^{N} \alpha_i (J_i(x) - J_i(\hat{x})) \leq 0, \text{ for all } x \in \mathcal{U}, \text{ or, equivalently, } \sum_{i=1}^{N} \alpha_i J_i(x) \leq \sum_{i=1}^{N} J_i(\hat{x}).$$

In other words, $\hat{x}$ maximizes a weighted sum of the objective functions, where the weights belong to the set $\mathcal{A}$.

EXERCISE 6.27 *(Proof Corollary 6.13)*
Assume $J_i,\ i = 1, \cdots, N$, are continuous differentiable on an open set $C \subset \mathbb{R}^n$. Let $x^* \in C$ be Pareto efficient for $J_i,\ i = 1, \cdots, N$. Consider

$$L := \sum_{i=1}^{N} \lambda_i J_i(x).$$

a. Use Lemma 6.10 and Theorem 4.34 to show there exist $\lambda_i \geq 0$, and $[\lambda_1, \cdots, \lambda_N] \neq [0, \cdots, 0]$ such that $\sum_{i=1}^{N} \lambda_i D_x J_i(x^*) = 0$.
b. Show there exist $\alpha \in \mathcal{A}$ such that $\sum_{i=1}^{N} \alpha_i D_x J_i(x^*) = 0$.

ANSWER 6.27

a. Note that, following the reasoning of Lemma 6.10, x^* is Pareto efficient only if it maximizes J_1 over the constrained set

$$\{x \in C \mid J_2(x) \geq J_2(x^*), \cdots, J_N(x) \geq J_N(x^*)\}.$$

By Theorem 4.34, this implies the existence of $\lambda_i \geq 0,\ i = 1, \cdots, N$, not all zero, such that the Lagrange function

$$L_1 := \lambda_1 J_1(x) - \sum_{i=2}^{N} \lambda_i (J_i^* - J_i)$$

has a stationary point at x^*.

b. By item a., there exist $\lambda_i \geq 0$, not all zero, such that

$$\sum_{i=1}^{N} \lambda_i DJ_i(x^*) = 0.$$

This implies (noting that $\sum_{j=1}^{N} \lambda_j > 0$) that

$$\sum_{i=1}^{N} \frac{\lambda_i}{\sum_{j=1}^{N} \lambda_j} DJ_i(x^*) = 0.$$

Or, with $\alpha_i := \frac{\lambda_i}{\sum_{j=1}^{N} \lambda_j}$,

$$\sum_{i=1}^{N} \alpha_i DJ_i(x^*) = 0, \text{ with } \alpha \in \mathcal{A}.$$

EXERCISE 6.28 *(Bargaining solutions)*

In this exercise, we will consider three different bargaining solutions that are frequently discussed in the literature. You are asked to calculate all of them in the context of Example 6.8.

In bargaining theory, the central question is which outcome will be realized if the players decide to cooperate. In making this decision, the potential gains relative to the non-cooperative outcome play a key role. This non-cooperative outcome is typically referred to as the threat point or fall-back position.

a. Assume the players use the Nash solution as their fall-back position. Show that this solution is $d = (1, 1)$.

b. In the Egalitarian solution players agree to divide the gains from cooperation equally. Geometrically, this is the intersection point of the 45 degrees line through the fall-back position with the Pareto frontier. Show that this solution is $(\frac{3}{2}, \frac{3}{2})$

c. In the solution proposed by Kalai and Smorodinsky first the ideal position is calculated. The best outcome player i can obtain when all other players play the non-cooperative solution is the i-th coordinate of the ideal position. Show that this ideal position is given by $I := (4\sqrt{\frac{1}{2}} - 1, 4\sqrt{\frac{1}{2}} - 1)$.

Next they propose the intersection point on the Pareto frontier with the line connecting the ideal and fall-back position as the resulting bargaining solution. Show this solution coincides with the Egalitarian solution.

d. The third solution, proposed by Nash, is the solution on the Pareto frontier that maximizes the product of the gains from the fall-back position. That is the point (P_1, P_2) on the Pareto frontier for which $(P_1 - d_1)(P_2 - d_2)$ is maximal. Show this solution also coincides with the Egalitarian solution in this case.

ANSWER 6.28

a. For every y, the function $-(x-1)^2 - y^2 + 2$ attains its maximum at $x = 1$. Similarly, for every x, the function $-x^2 - (y-1)^2 + 2$ attains its maximum at $y = 1$. The intersection of the two reaction curves is the unique point $(d_1, d_2) = (1, 1)$.

b. The Pareto frontier is parameterized by $(2 - 2(1-\alpha)^2, 2 - 2\alpha^2),\ \alpha \in [0, 1]$.
The Egalitarian solution is therefore obtained by solving $(1, 1) + \lambda(1, 1) = (2 - 2(1-\alpha)^2, 2 - 2\alpha^2)$. Elementary calculations show that the solution is attained at $\alpha = \frac{1}{2}$. Hence, the Egalitarian solution is $(\frac{3}{2}, \frac{3}{2})$.

c. The first coordinate of the Ideal point is obtained by solving $1 = 2 - 2\alpha^2$, which yields $\alpha = \sqrt{\frac{1}{2}}$. Hence,

$$I_1 = 2 - 2(1-\alpha)^2 = 2 - 2(1 - \sqrt{\frac{1}{2}})^2 = 4\sqrt{\frac{1}{2}} - 1.$$

Similarly, the second coordinate of the Ideal point is obtained by solving $1 = 2 - 2(1-\alpha)^2$, which yields $\alpha = 1 - \sqrt{\frac{1}{2}}$. Hence,

$$I_2 = 2 - 2\alpha^2 = 4\sqrt{\frac{1}{2}} - 1.$$

Thus, the intersection point is again obtained as the solution of

$$(1, 1) + \lambda(1, 1) = (2 - 2(1-\alpha)^2, 2 - 2\alpha^2).$$

Consequently, the Kalai-Smorodinsky solution coincides with the Egalitarian solution.

d. The bargaining solution proposed by Nash is obtained by maximizing, with respect to $\alpha \in [0, 1]$, the function

$$f(\alpha) := (2 - 2(1-\alpha)^2 - 1)(2 - 2\alpha^2 - 1).$$

Differentiation yields $f'(\alpha) = 1 - 6\alpha^2 + 4\alpha^3$. Since $f'(\frac{1}{2}) = 0$ and f has no no other stationary points, and $f"(\frac{1}{2}) < 0$, it follows that f attains a global maximum at $\alpha = \frac{1}{2}$. Therefore, the solution proposed by Nash coincides with the other two solutions.

EXERCISE 6.29 *(Proof Theorem 6.26)*

In this exercise you are asked to provide a proof for Theorem 6.26 in case there are 2 players.

So, assume the set of constraint instruments $\Gamma \subset \mathbb{R}^2$ is both compact and convex; and player i has instrument x_i at his disposal to maximize $J_i(x),\ i = 1, 2$.

a. Show that for every feasible instrument $(\bar{x}_1, \bar{x}_2) \in \Gamma$, the sets $\Gamma_1(\bar{x}_2) = \{(x_1, \bar{x}_2) \mid (x_1, \bar{x}_2) \in \Gamma\}$ and $\Gamma_2(\bar{x}_1) = \{(\bar{x}_1, x_2) \mid (\bar{x}_1, x_2) \in \Gamma\}$ are convex and compact.

b. Show that if $J_1(x)$ is continuous on Γ, the reaction curve $\hat{x}_1(x_2)$ is well defined in the sense that at all $(x_1, x_2) \in \Gamma$ the reaction curve has a value.

c. Assume both reaction functions $(\hat{x}_i(x_{-i})$ are continuous functions. Consider $f(x_1, x_2) := (\hat{x}_1(x_2), \hat{x}_2(x_1))$ on Γ. Use Brouwer's Fixed Point Theorem, see Exercise 2.57, to show that the game has a Nash equilibrium on Γ.

ANSWER 6.29

a. See item a. Exercise 6.30.

b. Let $\bar{x} = (\bar{x}_1, \bar{x}_2) \in \Gamma$. Then, by item a., $\Gamma_1(\bar{x}_2) = \{(x, \bar{x}_2) \mid (x, \bar{x}_2) \in \Gamma\}$ is a convex and compact set. Since $J_1(x)$ is continuous on Γ, it follows that $J_1(x, \bar{x}_2)$ is also continuous on $\Gamma_1(\bar{x}_2)$. Therefore, by Weierstrass's theorem, $J_1(x, \bar{x}_2)$ attains a global maximum on $\Gamma_1(\bar{x}_2)$, and there exists at least one point in $\Gamma_1(\bar{x}_2)$ at which this maximum is achieved. Hence, $\hat{x}_1(\bar{x}_2)$ is well-defined in the sense that the reaction function at this point $\bar{x}_2$ has at least one value. Note, however, it may have a finite as well as an infinite number of values too.

c. Let $f(x_1, x_2)$ be defined as above. Under the assumption that both reaction functions are continuous, f is a continous function on Γ. Moreover, its values lie in Γ, that is, $f(\Gamma) \subset \Gamma$.
Since Γ is a compact and convex set, Brouwer's Fixed Point Theorem implies that there exists a point $(\bar{x}_1, \bar{x}_2) \in \Gamma$ such that $f(\bar{x}_1, \bar{x}_2) = (\bar{x}_1, \bar{x}_2)$.
In other words, there exists a point $(\bar{x}_1, \bar{x}_2) \in \Gamma$ such that $\hat{x}_i(\bar{x}_{-i}) = \bar{x}_i,\ i = 1, 2$.
Thus, neither player has an incentive to deviate from his action at this point to improve his performance. In other words, the game has a Nash equilibrium at this point.

EXERCISE 6.30 *(Proof Theorem 6.30)*

In this exercise you are asked to provide a proof for Theorem 6.30.

So, assume the set of constraint instruments Γ is both compact and convex; the objective function $J_i(x)$, player i wants to maximize is continuous, and $J_i(x)$ is strictly concave w.r.t. its ith instrument for every feasible instrument in Γ, $i = 1, \cdots, N$,.

a. Show that for every feasible instrument $\bar{x} \in \Gamma$, the set $\Gamma_i(\bar{x}_{-i})$ is convex and compact.

b. Show that for every player i there exists a unique optimal response at every point $\bar{x}_i$ for which $\Gamma_i(\bar{x}_{-i})$ is not empty.

c. Consider, for simplicity, the two-player scalar case. Show by contradiction that the reaction function $\hat{x}_1(x_2)$ is continuous. That is, assume the reaction function is not continuous at some point $\bar{x}_2$. This would imply that there exists a sequence x_2^N converging to $\bar{x}_2$ for which $\hat{x}_1(x_2^N)$ would not converge to $\hat{x}_1(\bar{x}_2)$. Use next the Theorem of Bolzano-Weierstrass which states that every sequence which

values are bounded by some number b, has the property that this sequence has a subsequence s^n that converges, to arrive at a contradiction.

d. Use Brouwer's Fixed Point Theorem, see Exercise 2.57, to show that there exists a Nash equilibrium.

ANSWER 6.30

a. Note that $\bar{x} \in \Gamma_1(\bar{x}_{-1})$, so the set $\Gamma_1(\bar{x}_{-1})$ is not empty. Next, let $P,\ Q \in \Gamma_1(\bar{x}_{-1})$. Then there exist p, q such that $P = (p, \bar{x}_2, \cdots, \bar{x}_N)$ and $Q = (q, \bar{x}_2, \cdots, \bar{x}_N)$ both belong to Γ. Since Γ is convex by assumption, it follows that

$$\alpha(p, \bar{x}_2, \cdots, \bar{x}_N) + (1-\alpha)(q, \bar{x}_2, \cdots, \bar{x}_N) \in \Gamma, \forall \alpha \in [0,1].$$

That is,

$$(\alpha p + (1-\alpha)q, \bar{x}_2, \cdots, \bar{x}_N) \in \Gamma, \forall \alpha \in [0,1].$$

This implies that

$$\alpha P + (1-\alpha)Q \in \Gamma_1(\bar{x}_{-1}), \forall \alpha \in [0,1],$$

and hence $\Gamma_1(\bar{x}_{-1})$ is convex.

By the same reasoning, $\Gamma_i(\bar{x}_{-i})$ is convex for all i.

To show that $\Gamma_i(\bar{x}_{-i})$ is also bounded, note that $\Gamma_i(\bar{x}_{-i}) \subset \Gamma$, where Γ is bounded.

Finally, to prove that $\Gamma_i(\bar{x}_{-i})$ is closed (for $\bar{x} \in \Gamma$), suppose the contrary. Then there exists a sequence

$$y_n := (\bar{x}_1, \cdots, x_{i_n}, \cdots, \bar{x}_N) \in \Gamma_i(\bar{x}_{-i}),\ \forall n,$$

such that $\lim_{n\to\infty} y_n = y \notin \Gamma_i(\bar{x}_{-i})$. However, since Γ is closed and $y_n \in \Gamma$, it follows that $y \in \Gamma$. That is,

$$(\bar{x}_1, \cdots, \lim_{n\to\infty} x_{i_n}, \cdots, \bar{x}_N) \in \Gamma,$$

which implies $y \in \Gamma_i(\bar{x}_{-i})$, a contradiction.

Therefore, for $\bar{x} \in \Gamma$, the set $\Gamma_i(\bar{x}_{-i})$ is closed.

b. Since $\forall \bar{x} \in \Gamma$, $J_i(\bar{x}_{-i})$ is concave and $\Gamma_i(\bar{x}_{-i})$ is compact, it follows that $J_i(\bar{x}_{-i})$ attains a unique maximum $\hat{x}_i$ on $\Gamma_i(\bar{x}_{-i})$.

c. Assume that Γ is not a singleton (since in that case the statement of the theorem is trivially satisfied). Then Γ contains at least one additional point, and by convexity it follows that the line segment connecting these points also lies in Γ.

Suppose that $\hat{x}_1(x_2)$ is not continuous on Γ. Then there exists a sequence x_2^N converging to $\bar{x}_2$ such that $\hat{x}_1(x_2^N)$ does not converge to $\hat{x}_1(\bar{x}_2)$, and $\Gamma_1(x_2^N)$ is nonempty for every x_2^N.

Since $(\hat{x}_1(x_2^N), x_2^N) \in \Gamma$, the sequence $\hat{x}_1(x_2^N)$ is bounded. By the Bolzano-Weierstrass Theorem, it has a subsequence $\hat{x}_1(s_2^N)$ converging to some limit $r \neq \hat{x}_1(\bar{x}_2)$, where s_2^N converges to $\bar{x}_2$.

Hence, we obtain the inequality

$$J_1(x, s_2^N) \leq J_1(\hat{x}_1(s_2^N), s_2^N),\ \forall (x, s_2^N) \in \Gamma. \ (1)$$

Without loss of generality, assume that for some $\bar{N}$ we have $s_2^{\bar{N}} < \bar{x}_2$ and $s_2^{\bar{N}+1} > \bar{x}_2$. Then all points located on the line segment l_1 connecting $(\hat{x}_1(s_2^{\bar{N}}), s_2^{\bar{N}})$ and $(\hat{x}_1(\bar{x}_2), \bar{x}_2)$, and on the line segment l_2 connecting $(\hat{x}_1(s_2^{\bar{N}+1}), s_2^{\bar{N}+1})$ and $(\hat{x}_1(\bar{x}_2), \bar{x}_2)$, belong to Γ.
For sufficiently large N, s_2^N is close to $\bar{x}_2$, and the line $x_2 = s_2^N$ intersects either l_1 or l_2 at a point $(x_1^N, s_2^N) \in \Gamma$.
Thus, from (1), for sufficiently large N,

$$J_1(x_1^N, s_2^N) \leq J_1(\hat{x}_1(s_2^N), s_2^N).$$

Taking limits (note that, by assumption, J_1 is continuous. Therefore, whenever (x^N, y^N) converges to (x, y), $J_1(x^N, y^N)$ converges to $J_1(x, y)$.) yields

$$J_1(\hat{x}_1(\bar{x}_2), \bar{x}_2) \leq J_1(r, \bar{x}_2). \ (2)$$

However, since $J_1(x, \bar{x}_2)$ attains a unique maximum at $\hat{x}_1(\bar{x}_2)$, we must have

$$J_1(r, \bar{x}_2) < J_1(\hat{x}_1(\bar{x}_2), \bar{x}_2),$$

which contradicts (2).
Therefore, the assumption that $\hat{x}_1(x)$ is not continuous must be false.

d. Consider the mapping

$$r(x) = [\hat{x}_1(x_{-1}), \cdots, \hat{x}_n(x_{-n})]^T,$$

which, by item c., is a continuous function of Γ into itself.
Since Γ is compact, Brouwer's Fixed Point Theorem (Theorem 2.57) implies that $r(x)$ has a fixed point: that is, there exists $\bar{x} \in \Gamma$ such that $r(\bar{x}) = \bar{x}$.
In other words, the reaction functions intersect, or equivalently, a Nash solution exists.

EXERCISE 6.31 *(Team problem)*
Consider the problem where all N players share a common objective $J(x_1, \cdots, x_N)$ and player i controls variable $x_i,\ i = 1, \cdots, N$. We call $x^* = (x_1^*, \cdots, x_N^*)$ team optimal if $J(x^*) \geq J(x)$, for all $x \in \Gamma$.

a. Show that if x^* is team optimal, x^* is Nash efficient for the game where all players maximize $J(x)$.
b. Considers the 2-player case with $J(x_1, x_2) = -x_1^2 - x_2^2 - 8x_1x_2 + 10x_1 + 10x_2$. Show that the corresponding game has a Nash equilibrium but not a team optimal solution.
c. Show, using Lemma 6.10 and the definition of Pareto optimality, that x^* is team optimal if and only if x^* is Pareto efficient for the corresponding cooperative game.

d. Consider the 2-player case with $J(x_1, x_2) = -(x_1 - 1)^2(x_2 - 2)^2$. Show that the corresponding team problem has an infinite number of team optimal locations.
e. Consider the 2-player case with $J(x_1, x_2) = -(x_1 - x_2)^4 - \frac{4}{3}(x_1 - x_2)^3 + 4(x_1 - x_2)^2$. Show that the corresponding team problem has an infinite number of team optimal locations and has additional Nash efficient solutions that are not team optimal. Hint: consider the function $f(x) = -x^4 - \frac{4}{3}x^3 + 4x^2$.
f. Consider a team problem that has an optimal solution. Motivate why the introduction of a leader into this problem will not improve its solution.

ANSWER 6.31

a. Since x^* is team optimal, $J(x^*) \geq J(x),\ \forall x \in \Gamma$. Hence, in particular, $x_i^* = \arg\max_{x \in \Gamma_i(x_i^*)} J(x)$.
b. One can easily verify that $D_i J(x) = 0$ for $i = 1, 2$, if and only if $(x_1, x_2) = (1, 1)$. Since $D_{ii} J(x) < 0$ for $i = 1, 2$, it follows that both $J(x_1, 1)$ and $J(1, x_2)$ attain a global maximum at $x_1 = 1$ and $x_2 = 1$, respectively. Hence, $(1, 1)$ is Nash efficient.
   Next, note that, for example, $J(x_1, -x_1) = 6x_1^2$. Therefore, $J(x)$ does not attain a global maximum.
c. When $J_i = J$ for all i, it follows that if x^* is team optimal, then x^* maximizes J_i on Γ, and in particular on the constrained set Γ_i introduced in Lemma 6.10. Therefore, x^* is also Pareto.
   Conversely, if both x^* and y^* are Pareto efficient, the corresponding Pareto solutions are $(J(x^*), \cdots, J(x^*))$ and $(J(y^*), \cdots, J(y^*))$.
   Assume, without loss of generality, that $J(x^*) < J(y^*)$. Then choosing y^* instead of x^* would improve the performance for all players, contradicting the Pareto efficiency of x^*.
   In other words, there is exactly one Pareto solution, and for all $x \in \Gamma$, we have $J(x) \leq J(x^*)$. That is, x^* is team optimal.
d. All points of the form $(1, x_2)$ and $(x_1, 2)$, with $x_i \in \mathbb{R}$, maximize $J(x_1, x_2)$ and are therefore team optimal.
e. The derivative of $f(x) = -x^4 - \frac{4}{3}x^3 + 4x^2$ is

$$f'(x) = -4x^3 - 4x^2 + 8x = -4x(x - 1)(x + 2).$$

   Thus, f has stationary points at $x = -2$, $x = 0$ and $x = 1$. By examining the second-order derivative, we see that f attains a global maximum at $x = -2$, a local minimum at $x = 0$ and a local maximum at $x = 1$.
   This implies that the team problem has a solution at all points of the form $(x_1, x_1 + 2)$ with $x_1 \in \mathbb{R}$, and nowhere else.
   Moreover, it is easy to verify that at all points $(x_1, x_1 - 1)$, we have $J_{ii}(x_1, x_2) < 0$ for $i = 1, 2$. That is, all these points are Nash efficient but do not correspond to team solutions.
f. Introducing a leader into the problem restricts the set of possible instrument allocations Γ. Consequently, the optimization of J must be carried out over a smaller domain, which implies that the maximum of J on this restricted set can be no greater than the maximum value $J(x^*)$.

EXERCISE 6.32 Consider a game with a leader, L, and follower, F, who want to minimize their convex cost function $J_L(x_L, x_F)$ and $J_F(x_L, x_F)$ w.r.t. x_L and x_F, respectively. Assume J_L has a global minimum at $(x_L, x_F) = (a, b)$, J_F has a global minimum at $(x_L, x_F) = (c, d)$, and that both cost functions are differentiable. Assume, moreover, $D_1 J_F(a, b) \neq 0$.

a. Show that the incentive strategy $x_L = \alpha x_F + \beta$, where $\alpha = -\frac{D_2 J_F(a,b)}{D_1 J_F(a,b)}$ and $\beta = a - \alpha b$ leads to a minimum cost for the leader.

b. Use the result from item a) to verify the results of Exercise 6.12 and Example 6.38.

ANSWER 6.32

a. The leader seeks an incentive strategy of the form $x_L = \alpha x_F + \beta$ that minimizes his cost, i.e., $J_L(a, b)$. Thus, we have $a = \alpha b + \beta$ and $x_F = b$.

The follower, in turn, minimizes his cost $J_F(\alpha x_F + \beta, x_F)$ with respect to x_F. Since $J_F(x_L, x_F)$ is convex, every intersection curve of $J_F(x_L, x_F)$ is convex as well. Consequently, any stationary point on such a curve corresponds to a global minimum. Therefore, if there exist α and β such that

$$\alpha D_1 J_F(a, b) + D_2 J_F(a, b) = 0,$$

the follower will indeed minimize his objective.

A straightforward calculation shows that by choosing

$$\alpha := -\frac{D_2 J_F(a, b)}{D_1 J_F(a, b)} \text{ and } \beta := a - \alpha b,$$

we obtain $J_L(\alpha b + \beta, b) = J_L(a, b)$ is minimal, while also satisfying $\alpha D_1 J_F(a, b) + D_2 J_F(a, b) = 0$. This ensures that the follower will minimize his objective under this incentive strategy as well.

b. Note that in Exercise 6.12, we have $D_1 J_F(a, b) = 2(a - c)$ and $D_2 J_F(a, b) = 2(b - d)$, which gives the advertized result.

Furthermore, in Example 6.38, $D_1 J_F(x_L, x_F) = 2x_L - x_F$ and $D_2 J_F(x_L, x_F) = 2x_F - x_L$. Evaluating at $(x_L, x_F) = (0, 5)$ gives $D_1 J_F(0, 5) = -5$ and $D_2 J_F(0, 5) = 10$, which also reproduces the advertized result.

Problems with an Economic Content

EXERCISE 6.33

a. Calculate the Nash equilibrium in case the cost function of firm $i,\ i = 1, 2$, is $C_1(x_1) = 8x_1 + 24$ and $C_2(x) = \frac{1}{2}x_2^2$, respectively. Assume the market price is given by $p = 92 - x_1 - x_2$.

b) Assume welfare, W_i, in country $i,\ i = 1, 2$, can be measured as a function of the amount of pollution produced in both countries, x_i. Let $W_1 = -\frac{1}{3}(2x_1 + x_2 - 10)^3 + 2(2x_1 + x_2 - 10)^2 + 5(2x_1 + x_2 - 10)$ and $W_2 = -\frac{1}{2}(x_1 + 4x_2 - 12)^2 + 4(x_1 + 4x_2 - 12)$. Determine the Nash equilibrium of this game.

c. Consider the next *pursuit-evasion game*. Person 1 is at position x and can run to position $x + u_1$, where $u_1 \in [0, \frac{1}{2}]$. Person 2 is at position y and can run to position $y + u_2$, where $u_2 \in [0, 1]$. Person 1 likes to catch person 2, whereas person 2 likes to avoid person 1. This problem is formalized as that person 1 likes to minimize the squared distance $(x + u_1 - (y + u_2))^2$ w.r.t. u_1, whereas person 2 likes to maximize this squared distance w.r.t. u_2. Calculate the Nash equilibrium of this game if the position of the persons is $(x, y) = (2, 1)$ and $(x, y) = (1, 2)$, respectively. Interpret both outcomes.

ANSWER 6.33

a. Let $\pi_i = px_i - C_i$ for $i = 1, 2$. Then, we have

$$D_1\pi_1 = 84 - 2x_1 - x2 \text{ and } D_2\pi_2 = 92 - x_1 - 3x_2.$$

Setting $D_i\pi_i = 0$ for $i = 1, 2$, we finf that the reaction function of player 1 is

$$x_1 = 42 - \frac{1}{2}x_24$$

and the reaction curve of player 2 is

$$x_2 = \frac{92 - x_1}{3}.$$

The intersection point of these two curves is: $(32, 20)$.

b. We have

$$D_1W_1 = -2(2x_1 + x_2 - 10)^2 + 8(2x_1 + x_2 - 10) + 10 \text{ and } D_2W_2 = -4(x_1 + 4x_2 - 12) + 16.$$

Notice that

$$D_{11}W_1 = -16x_1 - 8x_2 + 96 \text{ and } D_{22}W_2 = -16.$$

Regardless of the value of x_1, the reaction curve of player 2,

$$x_2 = \frac{64 - 4x_1}{16}$$

always maximizes player 2's welfare. For player 1, given a fixed x_2, the reaction curve corresponds to the level of pollution x_1 for which $D_1W_1 = 0$, provided $D_{11}W_1 < 0$.
Substituting $x_2 = \frac{64-4x_1}{16}$ into $D_1W_1 = 0$ yields two candidate points:

$$(x_1, x_2) = (\frac{34}{7}, 2\frac{11}{14}) \text{ and } (x_1, x_2) = (\frac{10}{7}, 3\frac{5}{14}).$$

At the first point, $D_{11}W_1 = -16x_1 - 8x_2 + 96 < 0$, while at the second point, $D_{11}W_1 > 0$. Hence, only $(x_1, x_2) = (\frac{34}{7}, 2\frac{11}{14})$ qualifies as a Nash equilibrium.

c. For fixed $x,\ y,$ and u_2, player 1's minimum distance is achieved at the value of $u_1 \in [0, \frac{1}{2}]$ that minimizes

$$f(u_1) = (x + u_1 - (y + u_2))^2.$$

Notice that $f^{'}(u_1) = 0$ at $u_1 = y + u_2 - x$, where also $f^{"}(u_1) > 0$. Hence, the minimum distance occurs at $u_1 = y + u_2 - x$, provided that $y + u_2 - x \in [0, \frac{1}{2}]$.
If $y + u_2 - x < 0$, the optimal choice is $u_1 = 0$; if $y + u_2 - x > 0$, then $u_1 = \frac{1}{2}$ is optimal.
Similarly, for

$$g(u_2) = (x + u_1 - (y + u_2))^2 \text{ with } u_2 \in [0, 1],$$

the maximum occurs at either $u_2 = 1$ or $u_2 = 0$. Comparing $g(0)$ and $g(1)$ shows that the optimal policy for player 2 is

$$u_2 = 1 \text{ if } x + u_1 - y \leq \frac{1}{2} \text{ and } u_2 = 0 \text{ otherwise.}$$

Applying this to specific cases:

1. if $(x, y) = (2, 1)$, then $x+u_1-y = 1+u_1 \geq \frac{1}{2}$, so player 2 chooses $u_2 = 0$. Since $y+u_2-x = -1$, player 1's optimal choice is $u_1 = 0$.
2. if $(x, y) = (1, 2)$, then $x+u_1-y = -1+u_1 \leq \frac{1}{2}$ so player 2 chooses $u_2 = 1$. Since $y+u_2-x = 2$, player 1's optimal choice is $u_1 = \frac{1}{2}$

These outcomes are intuitive: in the first case $(x, y) = (2, 1)$, player 2 is behind player 1 and cannot improve his position by moving, whereas in the second case $(x, y) = (1, 2)$, player 2 should move as much as possible to avoid being caught by player 1.

EXERCISE 6.34 *(competitive market entry by firms)*
Consider a market of N producers who produce the same good. Assume the price they get is $p = 92-x$, where x is the total amount supplied by all producers together. The cost function of a producer is $C_i(x_i) = 32x_i + 3$.

a. Calculate the Nash equilibrium if $N = 2$.

b. Calculate the Nash equilibrium for an arbitrary N.

c. How many producers will enter this market at most?

ANSWER 6.34

a. Let

$$\pi_i = px_i - C(x) = (92 - x_1 - x_2)x_i - (32x_i + 3) \text{ for } i = 1, 2.$$

Then $D_i\pi_i = 92 - x_1 - x_2 - x_i - 32$ for $i = 1, 2$. From this, the Nash efficient solution can be easily obtained as the solution to the equations $D_i\pi_i = 0,\ i = 1, 2$. Simple calculations yield $(x_1, x_2) = (20, 20)$. Therefore, the Nash equilibrium is: $(\pi_1, \pi_2) = (397, 397)$.

b. Let

$$\pi_i = px_i - C(x) = (92 - \sum_{j=1}^{N} x_j)x_i - (32x_i + 3) \text{ for } i = 1, \cdots, N.$$

Then $D_i\pi_i = 92 - \sum_{j=1}^{N} x_j - x_i - 32,\ i = 1, \cdots, N$. From this, the Nash efficient solution can be easily obtained as the solution to the equations $D_i\pi_i = 0,\ i = 1, \cdots, N$. Simple calculations yield $(x_1, \cdots, x_N) = (\frac{60}{N+1}, \cdots, \frac{60}{N+1})$. Therefore, the Nash equilibrium is: $(\pi_1, \cdots, \pi_N)$, with $\pi_i = (92 - \sum_{j=1}^{N} x_j)x_i - (32x_i + 3) = (92 - \frac{60N}{N+1})\frac{60}{N+1} - (32\frac{60}{N+1} + 3)$.

c. No new producers will enter the market if they cannot earn a profit. Therefore, the number of producers that will enter is determined by the largest value of N for which π_i is nonnegative. By using, e.g., Matlab or simply calculating the roots of the quadratic equation $-3(N+1)^2 - 3600N + 3600(N+1) = 0$, we find that $N = 33$.

EXERCISE 6.35 *(subsidizing production green energy)*

Consider two electricity companies. The first company is a start-up company which intends to produce its electricity using "green" energy, the second company produces electricity using "fossil" energy. The price, p, they get on the market for electricity is: $p = 100 - x$, where x is the total amount of electricity supplied by both companies. The production cost for company 1 are initially $C(x_1) = 2x_1 + 650$ and for company two $C(x_2) = 4x_2 + 10$, respectively.

a. Determine the profit function of both firms and show that, without a subsidy, firm one will not enter the market.

b. Assume government is willing to pay firm one a loan to finance its debts until it reaches the break even point. Calculate how many years the government has to provide a loan. In how many years is company one able to pay the loan back to the government?

c. Calculate the final price and electricity amounts by both companies that will occur in the market.

d. Assume after a couple of years the fixed cost for the first company drop to 15. Additional "green" companies are allowed to enter the electricity market. Assume there are in total n "green" producing electricity companies in the market. Determine the equations the Nash solution has to satisfy. Show (using a computational package) that when 22 "green" producing electricity companies have entered the market the electricity company using "fossil" energy will drop out in the end.

ANSWER 6.35

a. With $\pi_i(x_1, x_2) = p * x_i - C(x_i)$ for $i = 1, 2$,

$$\pi_1(x_1, x_2) = (100 - x_1 - x_2)x_1 - 2x_1 - 650 \text{ and } \pi_2(x_1, x_2) = (100 - x_1 - x_2)x_2 - 4x_2 - 10.$$

If company 1 is not yet active in the market, the maximum profit for firm 2 is achieved where $\pi_2^{'} = 0$ (with $x_1 = 0$). That occurs at $x_2 = 48$. For this level of output, the optimal production for company 1 would be $x_1 = 25$. However, $\pi_1(25, 48) = -25$. Therefore, company one will not enter the market.

b. With a loan of 25, company 1 can survive its first year. Next, we determine the quantity that company 2 will produce in the second year, given the amount produced by company 1 in the first year. This is found by solving $\pi_2^{'} = 0$ with $x_1 = 25$, which yields $x_2 = 35.5$.

The profit of company 1 in the second year is then given by $\pi_1(25, 35.5) = 287.5$. Therefore, company 1 is able to repay its loan back to the government already in its second year of operation.

c. The reaction functions are given by $x_1 = (98 - x_2)/2$ and $x_2 = (96 - x_1)/2$, respectively. The equilibrium is determined by the intersection of these two reaction functions. Simple calculations show $(x_1, x_2) = (33\frac{1}{3}, 31\frac{1}{3})$.

d. If there are n "green" electricity-producing firms, the market price of electricity is given by

$$p = 100 - \sum_{i=1}^{n+1} x_i.$$

Let the first n firms produce "green" electricity. Their profit function is

$$\pi_i(x_1, \cdots, x_{n+1}) = (100 - \sum_{i=1}^{n+1})x_i - (2x_i + 15),\ i = 1, \cdots, n.$$

The profit of the firm using "fossil" energy is

$$\pi_{n+1}(x_1, \cdots, x_{n+1}) = (100 - \sum_{i=1}^{n+1})x_{n+1} - (4x_{n+1} + 10).$$

The Nash efficient solution is attained as the intersection of the following system of equations:

$$2x_i + \sum_{j \neq i}^{n+1} x_j = 98,\ i = 1, \cdots, n, \text{ and } 2x_{n+1} + \sum_{j=1}^{n} x_j = 96.$$

Using, for example, Matlab, it follows that if $n \geq 22$, the profit of the firm using "fossil" energy becomes non-positive. Therefore, it will exit the market then.

EXERCISE 6.36

Reconsider the fiscal policy coordination problem formulated in Exercise 6.4 item c.

a. Calculate as well the Nash efficient equilibrium strategies, the corresponding output as the Nash equilibrium for an arbitrary $\bar{y}$. Show that for $\bar{y} = 1$ these numbers are approximately $(f_1^N, f_2^N) = (0.4948, 0.0493)$; $(y_1^N, y_2^N) = (4.4502, 1.1348)$ and $(J_1^N, J_2^N) = (0.5471, 0.0062)$, respectively.

b. Show numerically that the Pareto solution in Exercise 6.4 item c. for which the distance of $(J_1(\alpha), J_2(\alpha)$ to $(0, 0)$ is as small as possible if $\bar{y} = 1$ is approximately $(J_1^m, J_2^m) = (0.5249, 0.0736)$ with $(f_1^m, f_2^m) = (0.4829, 0.2528)$ and $(y_1^m, y_2^m) = (4.4599, 1.2988)$.

c. Show numerically that the Pareto solution in Exercise 6.4 item c. for which the distance to the Nash solution is as small as possible if $\bar{y} = 1$ is approximately $(J_1^{NB}, J_2^{NB}) = (0.5470, 0.0059)$ with $(f_1^{NB}, f_2^{NB}) = (0.5123, 0.0578)$ and $(y_1^{NB}, y_2^{NB}) = (4.4669, 1.1487)$.

d. Comment on the possibility that fiscal authorities will coordinate their policies in order to achieve their goals.

ANSWER 6.36

a. The reaction function of the fiscal authority in country i for $i = 1, 2$ is

$$f_1 = \frac{1}{3.62}(1.8\bar{y} - 0.18f_2) \text{ and } f_2 = \frac{0.32}{3.28}(\bar{y} - f_1), \text{ respectively.}$$

Thus, the Nash efficient equilibrium strategies are

$$(f_1^N, f_2^N) = (0.4948, 0.0493)\bar{y}.$$

This yields the equilibrium output levels $(y_1^N, y_2^N) = (4.4502, 1.1348)\bar{y}$ and Nash equilibrium $(J_1^N, J_2^N) = (0.5471, 0.0062)\bar{y}^2$.

b. Consider $d(\alpha) := J_1^2(\alpha) + J_2^2(\alpha)$, with $\alpha \in [0, 1]$ and $J_i(\alpha)$ as defined in Exercise 6.4, item c. Then, compute numerically the value of α that minimizes $d(\alpha)$.

c. Consider

$$d_N(\alpha) := (J_1(\alpha) - J_1^N)^2 + (J_2(\alpha) - J_2^N)^2,$$

with $\alpha \in [0, 1]$ and $J_i(\alpha)$ as defined in Exercise 6.4, item c. Then, compute numerically the value of α that minimizes $d_N(\alpha)$, subject to the constraint that $J_i(\alpha) \leq J_i^N$ for all i.

d. We observe that through cooperation both players can improve their objective values by adopting the NB-policies. However, the resulting increase in output is relatively modest, while the effort required from the second fiscal player is about ten times larger than in the noncooperative setting. It is therefore questionable whether both players would agree to this arrangement as the increased effort by player 2 may lead to additional effects that are not captured in the model.
It is even less likely that both players would agree on the m-policies without a side-payment from country 1 to country 2, since country 2's objective value detoriates. This is due to the substantial additional fiscal effort required from country 2. Although its output increases - indeed exceeding its target level - country 1's output rises only marginally. Consequently, it is unclear how country 1 could finance a compensating side-payment to country 2.

EXERCISE 6.37 *(tracking an individual variable within a large population)*

Consider the problem where every individual i likes to track a variable y_i towards some target value $\bar{y}_i$ within a large population. The impact individuals have on each other's variable is assumed to be negligible. But the aggregate impact of the population to every player's variable is supposed to be significant. Below you are asked to analyze the noncooperative Nash outcome of this game under the assumption that the cost of using a control effort of x_i by player i to manipulate the variable y_i is growing quadratically.

We model this problem as follows:

Player i wants to minimize $J_i(x_1, \cdots, x_N) := (y_i - \bar{y}_i)^2 + r_i x_i^2$, where $y_i = ax_i + \frac{1}{N}\sum_{j=1}^{N} x_j + d_i,\ i = 1, \cdots, N$.

Here $r_i > 0$ is some constant which indicates the preference of player i of reaching his target versus

the control efforts he has to use. The constant d_i can be interpreted as the current value of the variable y_i which player i likes to control towards $\bar{y}_i$ using a control effort x_i, and a models the direct impact of control on the target variable.

a. Show that $x_i,\ i = 1, \cdots, N$ are Nash efficient for this game if and only if they satisfy next equations:

$$N(a + \frac{1}{N} + \frac{r_i}{a + \frac{1}{N}})x_i + \sum_{j \neq i} x_j = N(\bar{y}_i - d_i),\ i = 1, \cdots, N.$$

b. Assume throughout $r_i = r$ for all i. Let $s := N(a + \frac{1}{N} + \frac{r}{a+\frac{1}{N}})$. Show that

$$x_i = \frac{N}{s(s+N-2) - N + 1}((s + N - 1)(\bar{y}_i - d_i) - \sum_{j=1}^{N}(\bar{y}_j - d_j)),\ i = 1, \cdots, N,$$

solve the equations advertized in item a.

c. Show that for a large population, the action of player i converges to:
$x_i^c = \frac{a}{(a^2+r)(a^2+a+r)}((a^2 + a + r)(\bar{y}_i - d_i) - ae_{avg})$,
where $e_{avg} = \frac{1}{N}\sum_{i=1}^{N}(\bar{y}_i - d_i)$,
and variable y_i converges to
$y_i^c = \frac{a^2}{a^2+r}\bar{y}_i + \frac{r}{a^2+r}d_i + \frac{1}{N}\sum_{j=1}^{N}\frac{a}{a^2+r}(\bar{y}_j - d_j) - e_{avg}\frac{a^3+a^2}{(a^2+r)(a^2+a+r)}$.

d. Show that if either players do not care about the amount of used control or the direct impact of control on the target variable is high (i.e., either r is almost zero or a is large), the target variable is (almost) realized and the control used by the players converges to $x_i = \frac{1}{a}(\bar{y}_i - d_i) - \frac{1}{a(a+1)}e_{avg}$. This, under the assumption that the target values and current values of the variables are not very high.

e. Assume the current value (initial position) of all players is the same, $d_i = d$, and half of the (large) population is reaching for a high target value of y^h and the other half of the population likes to target the variable y_i at a low value of y^l.
Calculate the approximate value of the instrumens both sets of players use and show that the difference between the approximate output variables reaching for a high target and low target is $\frac{a}{a^2+r}(y^h - y^l)$.
Discuss the impact of control cost on above result and the direct impact of control on the realization of the target variable.

ANSWER 6.37

a. Consider $J_i(x)$ for fixed $x_j,\ j \neq i$. Differentiating $J_i(x)$ with respect to x_i yields the necessary first-order condition:

$$(y_i - \bar{y}_i)(a + \frac{1}{N}) + r_i x_i = 0,\ i = 1, \cdots, N.$$

Moreover,

$$D_{ii}J_i(x) = 2(a + \frac{1}{N})^2 + 2r_i > 0.$$

Hence, for fixed x_j, $j \neq i$, the value x_i indeed minimizes $J_i(x)$.

Therefore, any Nash efficient solution is obtained as the solution to the system of equations above.

Some straightforward algebraic manipulation then yields the equivalent advertized equations.

b. Let

$$s = N(a + \frac{1}{N} + \frac{r}{a + \frac{1}{N}}) \text{ and } e := \sum_{j=1}^{N}(\bar{y}_j - d_j).$$

We need to show that the proposed solution

$$x_i = \frac{N}{s(s+N-2) - N + 1}((s+N-1)(\bar{y}_i - d_i) - e)$$

satisfies

$$sx_i + \sum_{j \neq i} x_j = N(\bar{y}_i - d_i),\ i = 1, \cdots, N. \quad (1)$$

Substituting $x_i,\ i = 1, \cdots, N$, we find that

$$\frac{s(s+N-2) - N + 1}{N}(sx_i + \sum_{j \neq i} x_j - N(\bar{y}_i - d_i))$$

is equal to

$$\begin{aligned} & s(s+N-1)(\bar{y}_i - d_i) - se + (s+N-1)\sum_{j \neq i}(\bar{y}_j - d_j) - (N-1)e - \\ & (s(s+N-2) - N + 1)(\bar{y}_i - d_i) \\ = \ & (s+N-1)((s-1)(\bar{y}_i - d_i) + e) - (s+N-1)e) - (s(s+N-2) - (N+1))(\bar{y}_i - d_i) \\ = \ & 0. \end{aligned}$$

Thus, the proposed x_i indeed satisfies the system of equations (1).

c. Notice that

$$\begin{aligned} x_i &= \frac{N}{s(s+N-2)-N+1}((s+N-1)(\bar{y}_i-d_i)-\sum_{j=1}^{N}(\bar{y}_j-d_j)) \\ &= \frac{N^2}{s(s+N-2)-N+1}(\frac{s+N-1}{N}(\bar{y}_i-d_i)-e_{avg}) \\ &= \frac{1}{\frac{s}{N}(\frac{s}{N}+1-\frac{2}{N})-\frac{1}{N}+\frac{1}{N^2}}((\frac{s}{N}+1-\frac{1}{N})(\bar{y}_i-d_i)-e_{avg}), \end{aligned}$$

from which the conclusion regarding the approximate value of x_i follows immediately. Substituting this approximate value of x_i into

$$y_i = x_i + \frac{1}{N}\sum_{j=1}^{N} x_j + d_i,$$

then gives the corresponding approximate value of y_i.

d. This directly follows from item c. by considering the cases where either $r_i = 0$ (with $\bar{y}_i$ and d_i not being large) or a is much larger than r, $\bar{y}_i$ and d_i.

e. Notice that

$$e_{avg} = \frac{1}{N}(\sum_{i=1}^{\frac{N}{2}}(y^h-d)+\sum_{i=1}^{\frac{N}{2}}(y^l-d)) = \frac{1}{2}(y^h+y^l-2d).$$

Therefore, by item c., for players aiming for a high target value:

$$\begin{aligned} x_i^c &= \frac{a}{(a^2+r)(a^2+a+r)}((a^2+a+r)(\bar{y}_i-d_i)-ae_{avg}) \\ &= \frac{a}{(a^2+r)(a^2+a+r)}((a^2+a+r)(y^h-d)-\frac{1}{2}a(y^h+y^l-2d)) \\ &= \frac{a}{(a^2+r)(a^2+a+r)}((a^2+\frac{1}{2}a+r)(y^h-d)-\frac{1}{2}a(y^l-d)), \end{aligned}$$

and for players aiming for a low target value

$$x_i^c = \frac{a}{(a^2+r)(a^2+a+r)}((a^2+\frac{1}{2}a+r)(y^l-d)-\frac{1}{2}a(y^h-d)).$$

Hence, the difference in target values between the two groups of players is approximately:

$$\begin{aligned} y_h^c - y_l^c &= a(x_h^c - x_l^c) \\ &= \frac{a^2}{(a^2+r)(a^2+a+r)}((a^2+\frac{1}{2}a+r)(y^h-y^l) - \frac{1}{2}a(y^l-y^h)) \\ &= \frac{a^2}{(a^2+r)(a^2+a+r)}(a^2+a+r)(y^h-y^l) \\ &= \frac{a^2}{(a^2+r)}(y^h-y^l). \end{aligned}$$

Notice that when the cost of control is low, both groups of players are able to reach their targets. As control becomes increasingly costly, players exert less control, eventually reaching the point where neither uses any control, meaning both stick to their initial target value d, and the gap between them is zero.

The effect of the parameter a is similar: if a is relatively large compared to r, both groups achieve their target values. Conversely, if a is small relative to r, neither group exerts any control.

EXERCISE 6.38 *(aligning the population average with individual preferences in a large group)*

Consider the problem where every individual i likes to track a variable y towards some target value $\bar{y}_i$ within a large population. The impact individuals have on the target variable is assumed to be negligible. But the aggregate impact of the population is supposed to be significant. Below you are asked to analyze the noncooperative Nash outcome of this game under the assumption that the cost of using a control effort of x_i by player i to manipulate the variable y is growing quadratically.

We model this problem as follows:

Player i wants to minimize $J_i(x_1,\cdots,x_N) := (y-\bar{y}_i)^2 + r_i x_i^2$, where $y = \frac{1}{N}\sum_{i=1}^N x_i + d,\ i = 1,\cdots,N$. Here $r_i > 0$ is some constant which indicates the preference of player i of reaching his target versus the control efforts he has to use. The constant d can be interpreted as the current value of the variable y which player i likes to control towards $\bar{y}_i$ using a control effort x_i.

a. Show that $x_i,\ i = 1,\cdots,N$, are Nash efficient for this game if and only if they satisfy next equations:

$$(1+N^2 r_i)x_i + \sum_{j\neq i} x_j = N(\bar{y}_i - d),\ i = 1,\cdots,N.$$

b. Show that for every individual player the gap between the target and realization of the target variable, $\bar{y}_i - y$, equals $N r_i x_i$.

c. Next assume all players have the same aversion to using control, i.e., $r_i = r,\ \forall i$. Let $s := 1 + N^2 r$. Show that

$$x_i = \frac{N}{s(s+N-2)-N+1}((s+N-1)(\bar{y}_i - d) - \sum_{j=1}^N (\bar{y}_j - d)),\ i = 1,\cdots,N,$$

solve the equations advertized in item a.

d. Use items b. and c. to show that $\bar{y}_i - y = \bar{y}_i - d - \frac{1}{N(Nr+1)} \sum_{j=1}^{N} (\bar{y}_j - d)$. Assume $r \neq 0$ and the average value $\frac{1}{N} \sum_{j=1}^{N} (\bar{y}_j - d)$ converges. Show that the larger the population becomes, the target variable is almost not influenced by the control anymore. Or, stated differently, for large populations no tracking occurs.

e. Assume half of the population is reaching for a high target value of y^h, with $r_i = r_h$, and the other half of the population likes to target the variable y_i at a low value of y^l, with $r_i = r_l$. Moreover, assume both groups combine "forces" to reach their target value. That is, within both groups all control instruments are used collectively (but symmetrically) to reach individual's targets.
Show that at the Nash solution of this "collective" game $y = \frac{1}{4r_h r_l + r_h + r_l} (r_l(\bar{y}_h - d) + r_h(\bar{y}_l - d)) + d$.

f. Show that in item e., the fraction $\frac{r_h}{r_l}$ determines to what extent a population group will succeed in the realization of its goal.

ANSWER 6.38

a. Consider $J_i(x)$ with x_j fixed for $j \neq i$. Differentiating $J_i(x)$ with respect to x_i gives the necessary first-order condition:

$$(y - \bar{y}_i)\frac{1}{N} + r_i x_i = 0, \ i = 1, \cdots, N.$$

Moreover,

$$D_{ii} J_i(x) = 2(\frac{1}{N^2} + r_i) > 0,$$

so x_i indeed minimizes $J_i(x)$ for fixed x_j, $j \neq i$.
Therefore, any Nash efficient solution is obtained by solving the above equations. With some straightforward algebraic manipulation, these lead to the advertized equations.

b. Notice that

$$\bar{y}_i - y = \bar{y}_i - d - \frac{1}{N} \sum_{i=1}^{N} x_i.$$

From the equations in item a, we have

$$\sum_{i=1}^{N} x_i = N(\bar{y}_i - d) - N^2 r_i x_i.$$

Substituting this into the previous expression then yields the stated result.

c. Let $s := 1 + N^2 r$ and $e := \sum_{j=1}^{N} (\bar{y}_j - d)$. We need to show that the proposed

$$x_i = \frac{N}{s(s+N-2) - N + 1} ((s + N - 1)(\bar{y}_i - d) - e)$$

satisfy

$$s x_i + \sum_{j \neq i} x_j = N(\bar{y}_i - d) \text{ or, equivalently, } s x_i + \sum_{j \neq i} x_j - N(\bar{y}_i - d) = 0, \text{ for } i = 1, \cdots, N.$$

Substitution of $x_i,\ i = 1, \cdots, N$, into the above expression shows:
$\frac{s(s+N-2)-N+1}{N}(sx_i + \sum_{j\neq i} x_j - N(\bar{y}_i - d))$ equals:

$$\begin{aligned} & s(s+N-1)(\bar{y}_i - d) - se + (s+N-1)\sum_{j\neq i}(\bar{y}_j - d) - (N-1)e - \\ & (s(s+N-2) - N + 1)(\bar{y}_i - d) \\ = \ & (s+N-1)((s-1)(\bar{y}_i - d_i) + e) - (s+N-1)e) - (s(s+N-2) - (N+1))(\bar{y}_i - d) \\ = \ & 0. \end{aligned}$$

d. By item b., we have $\bar{y}_i - y = Nrx_i$. Substituting x_i from item c. gives

$$\bar{y}_i - y = \frac{N^2 r}{s(s+N-2) - N + 1}((s+N-1)(\bar{y}_i - d) - \sum_{j=1}^{N}(\bar{y}_j - d)).$$

Replacing s with $N^2 r + 1$ and performing some straightforward algebra then yields the advertized result.

Now, assume that the average gap

$$\frac{1}{N}\sum_{j=1}^{N}(\bar{y}_j - d)$$

converges to Δ. Then, for fixed r and large N,

$$\frac{1}{N(Nr+1)}\sum_{j=1}^{N}(\bar{y}_j - d) \approx \frac{1}{Nr+1}\Delta,$$

which tends to zero as N becomes large.

Consequently, $\bar{y}_i - y$ converges to $\bar{y}_i - d$ for large N, which means that y is almost equal to d. In other words, under these assumptions, the target variable changes very little from its initial value.

e. By assumption, half of the population uses $x_i = x_h$ and the other half uses $x_i = x_l$, with the values of these two instruments determined collectively within each group by minimizing the objective function

$$J_i = (y - \bar{y}_i)^2 + r_i x_i^2,\ i = h, l.$$

With

$$y = \frac{1}{2N}(Nx_h + Nx_l) + d = \frac{1}{2}(x_h + x_l) + d,$$

the first-order conditions are

$$y - \bar{y}_i + 2r_i x_i = 0,\ i = h, l.$$

Solving these equations gives

$$(x_h, x_l) = \frac{1}{4r_h r_l + r_h + r_l}(\frac{1}{2}(\bar{y}_h - \bar{y}_l) + 2r_l(\bar{y}_h - d), \frac{1}{2}(\bar{y}_l - \bar{y}_h) + 2r_h(\bar{y}_l - d)),$$

so that $y = \frac{1}{4r_h r_l + r_h + r_l}(r_l(\bar{y}_h - d) + r_h(\bar{y}_l - d)) + d$.

f. Let $\alpha := \frac{r_h}{r_l}$. It then follows from item c that

$$y = \frac{1}{4\alpha r_l + \alpha + 1}((\bar{y}_h - d) + \alpha(\bar{y}_l - d)) + d.$$

Thus, if α is close to zero, $y \approx \bar{y}_h$, whereas for large α, $y \approx \bar{y}_l$.

EXERCISE 6.39 *(taxing import industry by government)*

Consider two countries. In both countries there is a firm exporting goods to the other country. Assume firm 2 produces in country 2 flowers that it exports to country 1, whereas firm 1 produces oil in country 1 which it exports to country 2. Assume the demand for flowers in country 1, q_1, is given by $q_1 = \bar{q}_1 - \alpha_1(p_1 + t_1) + \beta\bar{y}_1$ and the demand for oil in country 2, $\bar{q}_2$ by $q_2 = \bar{q}_2 - \alpha_2(p_2 + t_2) + \beta\bar{y}_2$. Here, $\bar{y}_i$ denotes national income and t_i the tax raised by country $i,\ i = 1, 2$. Furthermore, p_1 is the price of flowers determined in country 2 by firm 2 and p_2 is the price of oil determined by firm 1 in country 1. The $\bar{q}_i$ may be interpreted as representing exogenous demand for the product i in both countries.

Assume firm 1 wants to maximize its profit $\pi_1 := p_2 q_2 - \gamma_1 q_2$. So, firm 1's cost of oil production are assumed to be linearly dependent on its demand. Firm 2 wants to maximize its profit $\pi_2 := p_1 q_1 - \gamma_2(p_2 + t_2)q_1$. The production of flowers is taking place in green houses that are heated using oil. Therefore its cost are linearly related to the oil price times the quantity sold.

Finally it is assumed that governments in both countries want to maximize on the one hand tax income and on the other hand sustain their exporting industry. Which is formalized as they determine tax that maximize a weighted average of both objectives, i.e. to maximize $\rho_i\pi_i + t_i q_i,\ i = 1, 2$, respectively.

a. Show that the Nash equilibrium of this game is given by:

$$\begin{bmatrix} p_1 \\ t_1 \\ p_2 \\ t_2 \end{bmatrix} = \frac{1}{d}\begin{bmatrix} \alpha_2(\alpha_1 s_1 \rho_2 \gamma_2^2 - 3\alpha_2 s_1 - 4\alpha_1 s_2 \gamma_2 - 2\alpha_1\alpha_2\gamma_1\gamma_2) \\ \alpha_2(2\alpha_1 s_2\gamma_2 + \alpha_1\alpha_2\gamma_1\gamma_2 - 3\alpha_2 s_1) \\ \alpha_1(\alpha_1\alpha_2\rho_2\gamma_1\gamma_2^2 + \alpha_1 s_2\rho_2\gamma_2^2 - 3\alpha_2 s_2 - \alpha_2 s_1\rho_2\gamma_2 - 6\alpha_2^2\gamma_1) \\ \alpha_1(3\alpha_2^2\gamma_1 - \alpha_1\alpha_2\rho_2\gamma_1\gamma_2^2 + 2\alpha_2 s_1\rho_2\gamma_2 - 3\alpha_2 s_2 - \alpha_1 s_2\rho_2\gamma_2^2) \end{bmatrix},$$

where $s_i := \bar{q}_i + \beta_i\bar{y}_i,\ i = 1, 2$, and $d = \alpha_1\alpha_2(\alpha_1\rho_2\gamma_2^2 - 9\alpha_2)$.

b. Assume government 1 is the leader in the game and the other 3 players play Nash.

Show that in principle government 1 has the incentive to raise taxes ad infinitum. Show that the best actions where the economy still survives are given by:

$$\begin{bmatrix} p_1 \\ p_2 \\ t_2 \end{bmatrix} = \frac{1}{3\alpha_1\alpha_2}\begin{bmatrix} \alpha_2(3s_1 - 6\alpha_1 t_1) \\ \alpha_1(2\alpha_2\gamma_1 + s_2 + \alpha_1\rho_2\gamma_2 t_1) \\ \alpha_1(s_2 - \alpha_2\gamma_1 - 2\alpha_1\rho_2\gamma_2 t_1) \end{bmatrix},$$

where $t_1 = \frac{2\gamma_2 s_2 + \alpha_2\gamma_1\gamma_2 - 3\alpha_2 s_1}{\alpha_1\rho_2\gamma_2^2 - 6\alpha_1\alpha_2}$ (assuming parameters are such that $t_1 > 0$).

c. Next assume Government 2 announces that it will levy the same taxes as country 1.
Calculate the Nash efficient solutions for the resulting 3-player game.

d. Assume that, under the conditions of item c., Government 1 is the leader in the 3-player game. Show that a unique efficient solution exists, assuming the two firms play Nash if $\alpha_1(1+\frac{\gamma_2}{2}) > \frac{\alpha_2\rho_1}{2}$. Under which additional conditions will this solution be realized?

ANSWER 6.39

a. Firm 1 aims to maximize its profit, π_1, with respect to p_2. The derivative is

$$D_{p_2}\pi_1 = q_2 - (p_2 - \gamma_1)\alpha_2.$$

Note that

$$D_{p_2p_2}\pi_1 = -2\alpha_2 < 0,$$

so π_1 attains a global maximum at its stationary point.
Similarly, firm 2 seeks to maximize its profit, π_2, with respect to p_1. The derivative is

$$D_{p_1}\pi_2 = q_1 - (p_1 - \gamma_2(p_2 + t_2))\alpha_1.$$

Note that

$$D_{p_1p_1}\pi_2 = -2\alpha_1 < 0,$$

so π_2 also attains a global maximum at its stationary point.
Government 1 chooses its tax t_1 to maximize its utility

$$U_1 := \rho_1\pi_1 + t_1q_1.$$

The derivative is $D_{t_1}U_1 = q_1 - \alpha_1t_1$, and $D_{t_1t_1}U_1 = -2\alpha_1 < 0$, so U_1 has a global maximum at its stationary point.
Government 2 chooses its tax t_2 to maximize its utility

$$U_2 := \rho_2\pi_2 + t_2q_2.$$

The derivative is $D_{t_2}U_2 = q_2 - \rho_2\gamma_2q_1 - \alpha_2t_2$, and $D_{t_2t_2}U_2 = -2\alpha_2 < 0$, so U_2 likewise achieves a global maximum at its stationary point.
Hence, the Nash efficient solution is given by the solution to:

$$\begin{bmatrix} 0 & 0 & -2\alpha_2 & -\alpha_2 \\ -2\alpha_1 & -\alpha_1 & \alpha_1\gamma_2 & \alpha_1\gamma_2 \\ -\alpha_1 & -2\alpha_1 & 0 & 0 \\ \alpha_1\rho_2\gamma_2 & \alpha_1\rho_2\gamma_2 & -\alpha_2 & -2\alpha_2 \end{bmatrix} \begin{bmatrix} p_1 \\ t_1 \\ p_2 \\ t_2 \end{bmatrix} = \begin{bmatrix} -\alpha_2\gamma_1 - s_2 \\ -s_1 \\ -s_1 \\ \rho_2\gamma_2s_1 - s_2 \end{bmatrix}.$$

Using Matlab, for instance, we obtain the advertized result from this.

b. For both firms and the government, the first-order conditions are the same when assuming that the taxes of country 1 are given. This leads to the system of equations:

$$\begin{bmatrix} 0 & -2\alpha_2 & -\alpha_2 \\ -\alpha_1 & 0 & 0 \\ \alpha_1\rho_2\gamma_2 & -\alpha_2 & -2\alpha_2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ t_2 \end{bmatrix} = \begin{bmatrix} -\alpha_2\gamma_1 - s_2 \\ 2\alpha_1 t_1 - s_1 \\ \rho_2\gamma_2 s_1 - \alpha_1\rho_2\gamma_2 t_1 - s_2 \end{bmatrix}.$$

From this, for a fixed t_1, the Nash efficient solution is:

$$\begin{bmatrix} p_1 \\ p_2 \\ t_2 \end{bmatrix} = \frac{1}{\alpha_1\alpha_2} \begin{bmatrix} \alpha_2(s_1 - 2\alpha_1 t_1) \\ \frac{\alpha_1}{3}(s_2 + 2\alpha_2\gamma_1 + \alpha_1\rho_2\gamma_2 t_1) \\ \frac{\alpha_1}{3}(s_2 - \alpha_2\gamma_1 - 2\alpha_1\rho_2\gamma_2 t_1) \end{bmatrix}.$$

Government 1, anticipating this reaction, chooses t_1 to maximize its utility:

$$U_1(t_1) := \rho_1\pi_1(t_1) + t_1 q_1(t_1) = \rho_1(p_2(t_1) - \gamma_1)q_2(t_1) + t_1 q_1(t_1).$$

The derivative with respect to t_1 is computed as:

$$\begin{aligned} D_{t_1}U_1 &= \rho_1 q_2 D_{t_1}p_2 + q_1 + \rho_1(p_2 - \gamma_1)D_{t_1}q_2 + t_1 D_{t_1}q_1 \\ &= \rho_1 q_2 D_{t_1}p_2 + q_1 - \alpha_2\rho_1(p_2 - \gamma_1)(D_{t_1}p_2 + D_{t_1}t_2) - \alpha_1 t_1(D_{t_1}p_1 + 1) \\ &= (\rho_1 q_2 - \alpha_2\rho_1(p_2 - \gamma_1))D_{t_1}p_2 + q_1 - \alpha_2\rho_1(p_2 - \gamma_1)D_{t_1}t_2 - \alpha_1 t_1(D_{t_1}p_1 + 1) \\ &= (\rho_1 q_2 - \alpha_2\rho_1(p_2 - \gamma_1))\frac{\alpha_1\rho_2\gamma_2}{3\alpha_2} + s_1 - \alpha_1(p_1 + t_1) + \rho_1(p_2 - \gamma_1)\frac{2\alpha_1\rho_2\gamma_2}{3} + \alpha_1 t_1 \\ &= \frac{\alpha_1\rho_1\rho_2\gamma_2}{3\alpha_2}s_2 - \frac{\alpha_1\rho_1\rho_2\gamma_2}{3}(t_2 + \gamma_1) + s_1 - \alpha_1 p_1 \\ &= \frac{\alpha_1\rho_1\rho_2\gamma_2}{3\alpha_2}(s_2 - \alpha_2\gamma_1) - \frac{\alpha_1\rho_1\rho_2\gamma_2}{3}\frac{1}{3\alpha_2}(s_2 - \alpha_2\gamma_1 - 2\alpha_1\rho_2\gamma_2 t_1) + 2\alpha_1 t_1 \\ &= \frac{2\alpha_1\rho_1\rho_2\gamma_2}{9\alpha_2}(s_2 - \alpha_2\gamma_1) + (\frac{2\alpha_1\rho_1\rho_2\gamma_2}{9\alpha_2}\alpha_1\rho_2\gamma_2 + 2\alpha_1)t_1. \end{aligned}$$

From this, we see that $D_{t_1t_1}U_1 > 0$, so the stationary point corresponds to a global minimum. Therefore, the maximum for Government 1 would, in principle, be achieved by raising the tax t_1 without bound.

However, increasing t_1 reduces firm 1's profit, which will become non-positive beyond a certain level. Specifically:

$$\begin{aligned} \pi_2(t_1) &= (p_1(t_1) - \gamma_2(p_2(t_1) + t_2(t_1))q_1(t_1) \\ &= \frac{1}{3\alpha_1\alpha_2}(3\alpha_2(s_1 - 2\alpha_1 t_1) - \gamma_2(\alpha_2\gamma_1 + 2s_2 - \alpha_1\rho_2\gamma_2 t_1))(s_1 - \alpha_1(\frac{1}{\alpha_1}(s_1 - 2\alpha_1 t_1) + t_1)) \\ &= \frac{1}{3\alpha_1\alpha_2}(3\alpha_2 s_1 - 2\gamma_2 s_2 - \alpha_2\gamma_1\gamma_2 + (\alpha_1\rho_2\gamma_2^2 - 6\alpha_1\alpha_2)t_1)\alpha_1 t_1. \end{aligned}$$

Thus, if

$$t_1 > \bar{t}_1 := \frac{2\gamma_2 s_2 + \alpha_2\gamma_1\gamma_2 - 3\alpha_2 s_1}{\alpha_1\rho_2\gamma_2^2 - 6\alpha_1\alpha_2},$$

firm 1 will stop exporting flowers. Hence, the maximum tax Government 1 can impose without causing the collapse of the economy is $\bar{t}_1$.

c. Setting $t_2 = t_1$, we follow the approach of item a. to derive the first-order conditions for the Nash game played by Government 1, firm 1 and firm 2.
Firm 1 seeks to maximize π_1 with respect to p_2. Notice that

$$D_{p_2}\pi_1 = q_2 - (p_2 - \gamma_1)\alpha_2, \text{ and } D_{p_2p_2}\pi_1 = -2\alpha_2 < 0.$$

Hence, π_1 attains a global maximum at its stationary point.
Similarly, firm 2 aims to maximize π_2 with respect to p_1. Simple calculations yield

$$D_{p_1}\pi_2 = q_1 - (p_1 - \gamma_2(p_2 + t_1))\alpha_1, \text{ and } D_{p_1p_1}\pi_2 = -2\alpha_1 < 0.$$

Thus, π_2 also reaches a global maximum at its stationary point.
Government 1 maximizes its utility

$$U_1 := \rho_1\pi_1 + t_1 q_1 \text{ with respect to } t_1.$$

Straightforward differentiation gives

$$D_{t_1}U_1 = \rho_1 D_{t_1}\pi_1 + q_1 - \alpha_1 t_1 = -\alpha_2\rho_1(p_2 - \gamma_1) + s_1 - \alpha_1 p_1 - 2\alpha_1 t_1, \text{ and } D_{t_1t_1}U_1 = -2\alpha_1 < 0.$$

So, U_1 attains a global maximum at its stationary point.
Therefore, assuming the matrix below is invertible, there exists a unique Nash efficient solution, which is given by the solution of:

$$\begin{bmatrix} -2\alpha_1 & \alpha_1\gamma_2 - \alpha_1 & \alpha_1\gamma_2 \\ -\alpha_1 & -2\alpha_1 & -\alpha_2\rho_1 \\ 0 & -\alpha_2 & -2\alpha_2 \end{bmatrix} \begin{bmatrix} p_1 \\ t_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} -s_1 \\ -s_1 - \alpha_2\rho_1\gamma_1 \\ -s_2 - \alpha_2\gamma_1 \end{bmatrix}.$$

Using, for example, Matlab, we can solve this system. Defining

$$d := 6\alpha_1 + \alpha_1\gamma_2 - 2\alpha_2\rho_1,$$

the solution is:

$$\begin{bmatrix} p_1 \\ t_1 \\ p_2 \end{bmatrix} = \frac{1}{d}\begin{bmatrix} 2\alpha_1\gamma_1\gamma_2 - \alpha_2\rho_1\gamma_1 + (2 + \gamma_2 - \frac{\alpha_2\rho_1}{\alpha_1})s_1 + (\rho_1 - \rho_1\gamma_2 + \frac{2\alpha_1\gamma_2}{\alpha_2})s_2 \\ 2\alpha_2\rho_1\gamma_1 - \alpha_1\gamma_1\gamma_2 + 2s_1 - (2\rho_1 + \frac{\alpha_1\gamma_2}{\alpha_2})s_2 \\ \gamma_1(3\alpha_1 + \alpha_1\gamma_2 - 2\alpha_2\rho_1) - s_1 + \frac{\alpha_1(3+\gamma_2)}{\alpha_2}s_2 \end{bmatrix}.$$

d. From item c., for a fixed t_1, we have

$$D_{p_2}\pi_1 = q_2 - (p_2 - \gamma_1)\alpha_2, \text{ and } D_{p_1}\pi_2 = q_1 - (p_1 - \gamma_2(p_2 + t_1))\alpha_1.$$

Furthermore, as argued previously, any stationary point corresponds to a global minimum. Hence, the Nash solution for the two firms is obtained as the solution of:

$$\begin{bmatrix} -2\alpha_1 & \alpha_1\gamma_2 \\ 0 & -2\alpha_2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} -s_1 - (\alpha_1\gamma_2 - \alpha_1)t_1 \\ \alpha_2 t_1 - s_2 - \alpha_2\gamma_1 \end{bmatrix}.$$

Solving this system yields the unique Nash efficient prices:

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} \frac{\gamma_1\gamma_2}{4} + (\frac{\gamma_2}{4} - \frac{1}{2})t_1 + \frac{1}{2\alpha_1}s_1 + \frac{\gamma_2}{4\alpha_2}s_2 \\ \frac{\gamma_1}{2} - \frac{1}{2}t_1 + \frac{1}{2\alpha_2}s_2 \end{bmatrix}.$$

Government 1's utility function is then

$$U_1(t_1) = \rho_1\pi_1(t_1) + t_1 q_1(t_1).$$

Its derivative with respect to t_1 is

$$\begin{aligned} D_{t_1}U_1(t_1) &= \rho_1 D_{t_1}\pi_1(t_1) + q_1 + t_1 D_{t_1}q_1(t_1) \\ &= \rho_1(q_2 D_{t_1}p_2(t_1) + (p_2(t_1) - \gamma_1)D_{t_1}q_2(t_1)) + s_1 - \alpha_1(p_1(t_1) + t_1) - \\ &\quad \alpha_1 t_1(1 + D_{t_1}p_1(t_1)) \\ &= \rho_1(-\frac{1}{2}q_2 - \frac{\alpha_2}{2}(p_2(t_1) - \gamma_1)) + s_1 - \alpha_1(p_1(t_1) + t_1) - \alpha_1 t_1(\frac{1}{2} + \frac{\gamma_2}{4}) \\ &= -\frac{1}{2}\rho_1(s_2 - \alpha_2 t_1 - \alpha_2\gamma_1) + s_1 - \alpha_1(p_1(t_1) + t_1) - \alpha_1 t_1(\frac{1}{2} + \frac{\gamma_2}{4}) \\ &= -\frac{1}{2}\rho_1(s_2 - \alpha_2\gamma_1 - \alpha_2 t_1) + s_1 - \alpha_1(\frac{\gamma_1\gamma_2}{4} + (1 + \frac{\gamma_2}{2})t_1 + \frac{1}{2\alpha_1}s_1 + \frac{\gamma_2}{4\alpha_2}s_2) \\ &= (\frac{\alpha_2\rho_1}{2} - \alpha_1(1 + \frac{\gamma_2}{2}))t_1 + \frac{1}{2}s_1 - (\frac{\alpha_1\gamma_2}{4\alpha_2} + \frac{\rho_1}{2})s_2 + \frac{\alpha_2\rho_1\gamma_1}{2} - \frac{\alpha_1\gamma_1\gamma_2}{4}. \end{aligned}$$

Therefore, if

$$\alpha_1(1 + \frac{\gamma_2}{2}) > \frac{\alpha_2\rho_1}{2},$$

Government 1 maximizes its utility by choosing

$$t_1 = \frac{s_1 - (\frac{\alpha_1\gamma_2}{2\alpha_2} + \rho_1)s_2 + \alpha_2\rho_1\gamma_1 - \frac{\alpha_1\gamma_1\gamma_2}{2}}{\alpha_1(2 + \gamma_2) - \alpha_2\rho_1}.$$

An equilibrium is achieved at this efficient location, provided that the profits of both firms remain non-negative.

Bibliography

[1] Axelrod, R., 2006. *The Evolution of Cooperation* (Revised ed.), Perseus Books Group, New York.

[2] Başar, T., and Olsder, G.J., 1999. *Dynamic Noncooperative Game Theory*, SIAM, Philadelphia.

[3] Bertsekas, D.P., 2015. *Convex Optimization Algorithms*. Athena Scientific, Belmont MA.

[4] Binmore, K.G., 1981. *Mathematical Analysis: a straightforward approach*, second edition. Cambridge University Press, Cambridge.

[5] Boyd, S., and Vandenberghe, L., 2004. *Convex Optimization*. Cambridge University Press, Cambridge.

[6] Chiang, A.C., 1992. *Elements of Dynamic Optimization*. McGraw-Hill, New York

[7] Engwerda, J.C., 2005. *LQ Dynamic Optimization and Differential Games*. John Wiley, Chichester.

[8] Fan, K., Glicksberg, I., and Hoffman, A.J., 1957. Systems of inequalities involving convex functions. *American Mathematical Society Proceedings*, Vol.8, 617-622.

[9] Farkas, J., 1902. Theorie der einfachen Ungleichungen. *Journal für die Reine und Angewandte Mathematik*, Vol.124, 1-27. doi:10.1515/crll.1902.124.1.

[10] Fliege, J., and Svaiter, B.F., 2000. Steepest descent methods for multicriteria optimization. *Mathematical Methods of Operations Research*, Vol.51, 479-494.

[11] Hestenes, M.R., 1966. *Calculs of Variations and Optimal Control Theory.* John Wiley, London.

[12] John, F., 1948. Extremum problems with inequalities as subsidiary conditions, in: Friedrichs, K.O., et al. (eds.), *Studies and Essays, Courant Anniversary Volume*, Wiley/Interscience.

[13] Lay, D.C., Lay, S.R., McDonald, J.J., 2016. *Linear Algebra and its Applications*. Pearson, New York.

[14] Markowitz, H.M., 1952. Portfolio Selection. *The Journal of Finance*, Vol.7, 77-91.

[15] Minkowski, H., 1896. *Geometrie der Zahlen*. I, Teubner, Leipzig, 1896; II. Teubner, Leipzig, 1910.

[16] Neumann, J. von, and Morgenstern, O., 1947. *Theory of Games and Economic Behaviour.* Princeton University Press, Princeton NJ.

[17] Nesterov, Y., 2004. *Introductory Lectures on Convex Optimization.* Springer-Verlag, Berlin.

[18] Nocedal, J., and Wright, S.J., 1999. *Numercal Optimization.* Springer Verlag, Berlin.

[19] Olsder, G.J., 2009. Phenomena in inverse Stackelberg games, part 1: static problems. *Journal of Optimization Theory and its Applications*, Vol.143, 589-600.

[20] Osborne, M.J., 2009. *An Introduction to Game Theory*, Oxford University Press.

[21] Roos, C, Terlaky, T., and Vial, J-P., 2006. *Interior Point Methods for Linear Optimization, Second Edition*. Springer-Verlag, Berlin.

[22] Schrijver, A., 1998. *Theory of Linear and Integer Programming*. Wiley, Chichester.

[23] Sierksma, G, and Zwols, Y., 2015. *Linear and Integer Optimization Theory and Practice, Third Edition*. Chapman & Hall, CRC Press, London.

[24] Simon, C.P., and Blume L., 1994. *Mathematics for Economists*, W.W. Norton & Company, New York.

[25] Soltan, V., 2021. Support and separation properties of convex sets in finite dimension. *Extracta Mathematicae*, Vol.36, 241-278.

[26] Stackelberg, H. von, 1934. *Marktform und Gleichgewicht*, Springer-Verlag, Berlin.

[27] Sydsaeter, K., and Hammond, P., 2012. *Essential Mathematics for Economic Analysis*. Pearson, London.

[28] Takayama, A., 1974. *Mathematical Economics*. The Dryden Press, Hinsdale, Illinois.

[29] Thomson, W., 1994. Cooperative models of bargaining, in: Aumann, R., and Hart, S., (eds.), *Handbook of Game Theory, Vol.2*, Elsevier Science, 1238-1277.

[30] Watson, J., 2013. *Strategy, An Introduction to Game Theory*. W.W. Norton & Company, London.

Index

Symbols

$O(x)$, *77*

B

Bolzano-Weierstrass Theorem, *321*
Brouwer's Fixed Point Theorem, *25*
    multi-variable case, *85*

C

Cauchy-Schwartz inequality, *67*
Chebyshev distance problem, *205*
Concave function
    first-order characterization maximum, *256*
    global extremal point, *255*
constraint qualifications, *222*
    Slater's, *252*
continuous function
    combination property, *70*
    image continuity property, *71*
    no partial derivatives, *84*
Contraction Mapping Theorem, *26*
Convex function
    global extremal point, *28*
convex function
    combination property, *74*
convex programming problem, *137*

D

differentiable function
    continuity, *72*
    definition extended, *85*
    no continuous partial derivatives, *84*

E

economic illustrations
    competitive market entry, *327*
consumer choice problem, *222*
fiscal coordination, *282*, *286*, *329*
green farming, *227*
homogeneous production function, *96*
marginal output rule, *36*
optimal consumption of products, *266*
optimal fertilization, *224*
optimal production planning, *281*
pollution
    optimization, *96*
portfolio selection
    convex formulation, *268*
    Markowitz, *141*
price discrimination, *98*
production function
    CES, *151*
    optimal use production factors, *151*
production optimization, *143*
production under delivery constraints, *271*
profit maximization
    firm, *110*
profit maximization by a firm, *99*
pursuit-evasion game, *326*
sharing limited production resources, *281*
sharing the cakes, *282*
stock acquisition, *147*
subsidizing production green energy, *328*
taxing import industry by government, *337*
tracking average variable within a large population, *334*
tracking individual variable within a large population, *330*
utility optimization

derivation demand function, *148*
Engel curve, *150*
income consumption curve, *150*
worst-case growth optimization, *294*
epigraph, *248*
Equality Constraint Optimization Problem
Inequality constraint problem
reformulation strict inequality as equality constraint, *136*
Least-squares optimization, *140*
linear constraints, *132*
Equivalent optimization problem
change of variables, *207*
extremal points
convex characterization minimum, *34*
sufficient characterization, *33*

F

Farkas'lemma, *221*
fixed point, *26*
fixed point theorems
Contraction Mapping Theorem, *86*
function
invertible, *19*
functions
homogeneous, *96*

G

Geometric programming, *208*

I

Inequality constraint problem
reformulation inequality as equality constraint, *201*
reformulations into framework, *200*
instantaneous rate of growth
per capita, *40*
Interior-Point Method
Barrier Method, *215*
Inverse function theorem
derivative inverse function, *32*

L

Limit
calculation rules, *69*
coordinate-criterion, *68*
definition, *19*, *68*
line segment, *73*
linear complementarity problem, *265*
equivalent quadratic programming problem, *265*
Linear Programming Problem
Lagrange dual function, *263*
Linear-fractional programming, *210*
logistic growth curve, *224*

M

Max-Min inequality, *261*
Mean Value Theorem, *22*
multi-variable case, *73*
minimum function, *80*
minimum of a function
convex characterization, *34*
monomial, *208*

N

Newton-Raphson
multi-variable case, *77*
scalar case, *29*
Nonlinear Programming Problem
linear complementarity problem, *265*
quadratic case, *265*

O

one-to-one mapping, *207*

P

Pareto
solution
social outcome, *282*
partial derivative

non-differentiable, *82*
polyhedron, *206*
posynomial, *209*
pursuit-evasion game, *326*

Q

quadratic programming problem, *265*
equality constraints, *264*

R

Rolle's Theorem, *21*

S

Seperating Hyperplane Theorem
Minkowski separation theorem, *219*
slack variable, *201*

T

Taylor's Theorem, *24*
Taylor's theorem
multi-variable case, *74*
vector-valued case, *77*
triangle inequality, *67*

Z

zero-sum game, *314*
zero-sum game, symmetric, *316*

This book reviews a range of standard mathematical tools used to solve optimization problems.

It discusses the conditions under which function optima exist, considering both single-variable and multi-variable functions. Because resources are often limited, special attention is given to constrained optimization problems, where the domain of the function is restricted. Convex and linear programming problems are examined in greater detail due to their favorable numerical properties.

Finally, the book addresses a number of concepts related to optimization problems involving multiple agents or multiple objectives.

Each chapter ends with a collection of exercises divided into three categories. The first category contains exercises designed to help students assess their understanding of the material presented. The second category promotes a deeper engagement with mathematical rigor, including proofs of theorems discussed in the main text as well as related topics. The third category demonstrates how the theoretical tools can be applied in economics to enhance understanding of real-world phenomena.

In this accompanying volume, we provide solutions to all exercises included in the book.



Jacob Engwerda was born in Nieuweschans, the Netherlands, in 1958. He obtained his master's degree from the University of Groningen in 1982 and earned his Ph.D. in mathematics, with a specialization in systems and control theory, from the Eindhoven University of Technology in 1988.

From 1988 to 2019, he served as an associate professor in the Department of Econometrics and Operations Research at Tilburg University. He is the author of *LQ Dynamic Optimization and Differential Games* and a coauthor of *Dynamic Modeling of Monetary and Fiscal Cooperation Among Nations and The Interval Market Model in Mathematical Finance*.

His research centers on the development of system-theoretic methods for control problems in economics, with a particular emphasis on game theory. He has published more than 250 scientific papers. Since retiring in 2019, he continues to conduct some research in his spare time.

OPEN PRESS TiU