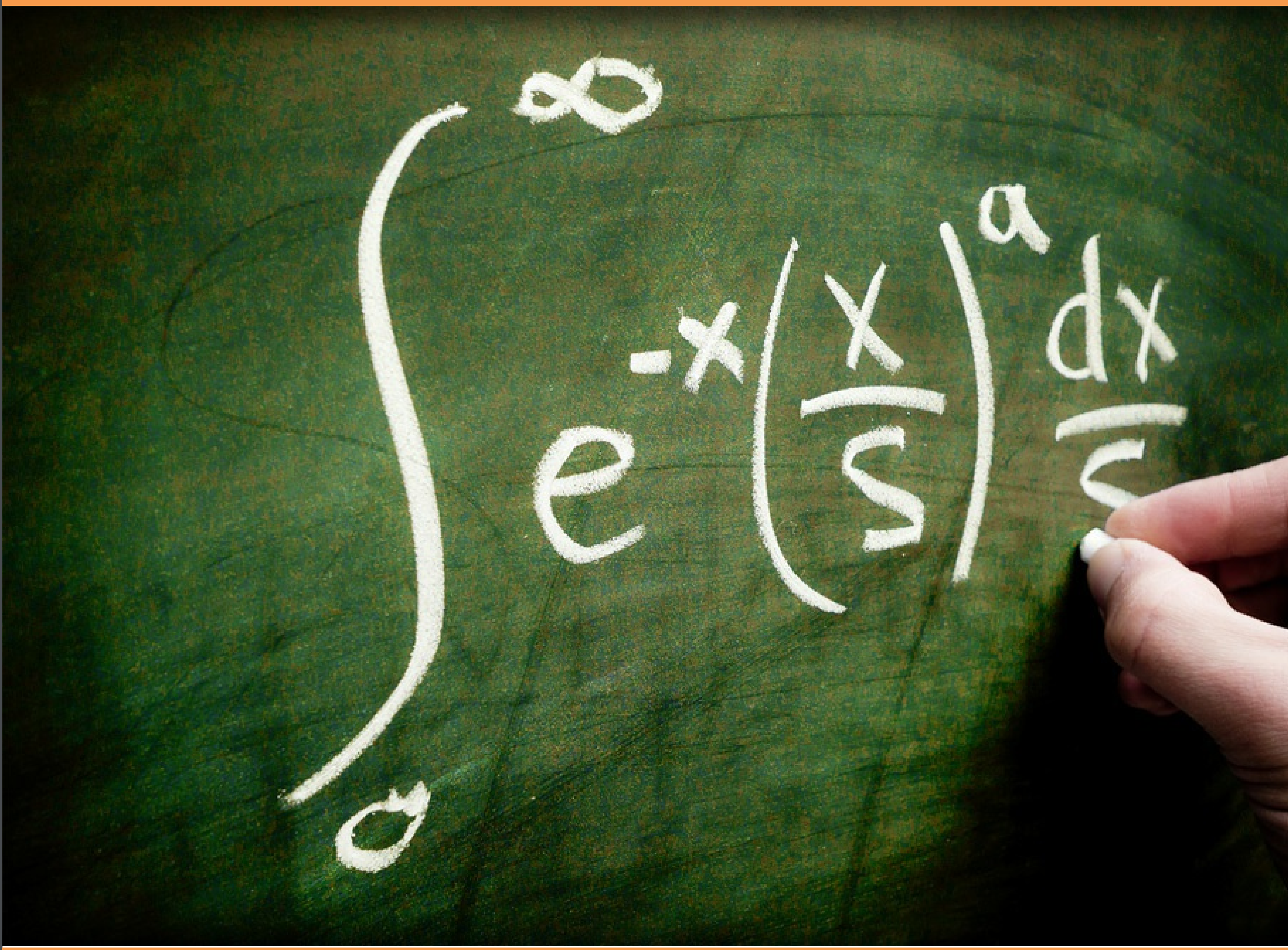
bookboon.com

Linear and Convex Optimization

Convexity and Optimization – Part II

Lars-Åke Lindahl



Download free books at
bookboon.com

LARS-ÅKE LINDAHL

LINEAR AND CONVEX OPTIMIZATION

CONVEXITY AND OPTIMIZATION – PART II

Download free eBooks at bookboon.com

Linear and Convex Optimization: Convexity and Optimization – Part II
1st edition
© 2016 Lars-Åke Lindahl & bookboon.com
ISBN 978-87-403-1383-3

CONTENTS

Download free eBooks at bookboon.com

Click on the ad to read more

Download free eBooks at bookboon.com

Download free eBooks at bookboon.com

Download free eBooks at bookboon.com

Download free eBooks at bookboon.com

Preface

Part II in this series of three on convexity and optimization is about linear and convex optimization.

We start by studying some equivalent ways to formulate a given optimization problem and then present some classical model examples.

Duality is an important principle in many areas of mathematics, so also in optimization theory. To each minimization problem we can associate a dual maximization problem by means of the so-called Lagrange function, and the two problems have the same optimal value provided certain conditions are fulfilled. We devote two chapters to the study of duality for general convex optimization problems and then treat the special case of linear programming in a separate chapter.

The simplex algorithm, which until the mid 1980's was the only practical algorithm for solving large linear optimization problems, is studied in the last chapter.

Uppsala, April 2016
Lars-Åke Lindahl

Download free eBooks at bookboon.com

List of symbols

$\operatorname{con} X$	conic hull of X, see Part I
$\operatorname{cvx} X$	convex hull of X, see Part I
$\operatorname{dom} f$	the effective domain of the function f, i.e. $\{x \mid -\infty < f(x) < \infty\}$
$\operatorname{ext} X$	set of extreme points of X, see Part I
$\operatorname{recc} X$	recession cone of X, see Part I
f'	derivate or gradient of f, see Part I
$v_{\max}$, $v_{\min}$	optimal values, p. 2
$B(a;r)$	open ball centered at a with radius r
$\overline{B}(a;r)$	closed ball centered at a with radius r
$I(x)$	set of active constraints at x, p. 42
$L(x,\lambda)$	Lagrange function, p. 32
$M_{\hat{r}}[x]$	object obtained by replacing the element in M at location r by x, p. 93
$\mathbf{R}_+$, $\mathbf{R}_{++}$	$\{x \in \mathbf{R} \mid x \geq 0\}$, $\{x \in \mathbf{R} \mid x > 0\}$
$\mathbf{R}_-$	$\{x \in \mathbf{R} \mid x \leq 0\}$
$\overline{\mathbf{R}}$, $\underline{\mathbf{R}}$, $\underline{\overline{\mathbf{R}}}$	$\mathbf{R} \cup \{\infty\}$, $\mathbf{R} \cup \{-\infty\}$, $\mathbf{R} \cup \{\infty, -\infty\}$
X^+	dual cone of X, see Part I
$\mathbf{1}$	the vector $(1, 1, \ldots, 1)$
$\phi(\lambda)$	dual function $\inf_x L(x,\lambda)$, p. 33
∇f	gradient of f
$[x,y]$	line segment between x and y
$]x,y[$	open line segment between x and y
$\\|\cdot\\|_1$, $\\|\cdot\\|_2$, $\\|\cdot\\|_\infty$	ℓ^1-norm, Euclidean norm, maximum norm, see Part I

Download free eBooks at bookboon.com

Chapter 9

Optimization

The Latin word *optimum* means 'the best'. The optimal alternative among a number of different alternatives is the one that is the best in some way. Optimization is therefore, in a broad sense, the art of determining the best.

Optimization problems occur not only in different areas of human planning, but also many phenomena in nature can be explained by simple optimization principles. Examples are light propagation and refraction in different media, thermal conductivity and chemical equilibrium.

In everyday optimization problems, it is often difficult, if not impossible, to compare and evaluate different alternatives in a meaningful manner. We shall leave this difficulty aside, for it can not be solved by mathematical methods. Our starting point is that the alternatives are ranked by means of a function, for example a profit or cost function, and that the option that gives the maximum or minimum function value is the best one.

The problems we will address are thus purely mathematical − to minimize or maximize given functions over sets that are given by a number of constraints.

9.1 Optimization problems

Basic notions

We begin by recalling the following notation from Part I:

$$\begin{aligned}\overline{\mathbf{R}} &= \mathbf{R}\cup\{+\infty\}\\ \underline{\mathbf{R}} &= \mathbf{R}\cup\{-\infty\}\\ \underline{\overline{\mathbf{R}}} &= \mathbf{R}\cup\{-\infty,+\infty\}.\end{aligned}$$

Download free eBooks at bookboon.com

For the problem of minimizing a function $f\colon \Omega \to \overline{\mathbf{R}}$ over a subset X of the domain Ω of the function, we use the notation

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in X. \end{array}$$

Here, s.t. is an abbreviation for the phrase *subject to the condition.*

The elements of the set X are called the *feasible points* or *feasible solutions* of the optimization problem. The function f is the *objective function.*

Observe that vi allow $+\infty$ as a function value of the objective function in a minimization problem.

The *(optimal) value* $v_{\min}$ of the minimization problem is by definition

$$v_{\min} = \begin{cases} \inf\{f(x) \mid x \in X\} & \text{if } X \neq \emptyset, \\ +\infty & \text{if } X = \emptyset. \end{cases}$$

The optimal value is thus a real number if the objective function is bounded below and not identically equal to $+\infty$ on the set X, the value is $-\infty$ if the function is not bounded below on X, and the value is $+\infty$ if the objective function is identically equal to $+\infty$ on X or if $X = \emptyset$.

Of course, we will also study maximization problems, and the problem of maximizing a function $f\colon \Omega \to \underline{\mathbf{R}}$ over X will be written

$$\begin{array}{ll} \max & f(x) \\ \text{s.t.} & x \in X. \end{array}$$

The *(optimal) value* $v_{\max}$ of the maximization problem is defined by

$$v_{\max} = \begin{cases} \sup\{f(x) \mid x \in X\} & \text{if } X \neq \emptyset, \\ -\infty & \text{if } X = \emptyset. \end{cases}$$

The optimal value of a minimization or maximization problem is in this way always defined as a real number, $-\infty$ or $+\infty$, i.e. as an element of the extended real line $\overline{\underline{\mathbf{R}}}$. If the value is a real number, we say that the optimization problem has a *finite* value.

A feasible point x_0 for an optimization problem with objective function f is called an *optimal point* or *optimal solution* if the value of the problem is finite and equal to $f(x_0)$. An optimal solution of a minimization problem is, in other words, the same as a global minimum point (with a finite value). Of course, problems with finite optimal values need not necessarily have any optimal solutions.

Download free eBooks at bookboon.com

From a mathematical point of view, there is no difference in principle between maximization problems and minimization problems, since the optimal values $v_{\max}$ and $v_{\min}$ of the problems

$$\begin{array}{ll} \max & f(x) \\ \text{s.t.} & x \in X \end{array} \qquad \text{and} \qquad \begin{array}{ll} \min & -f(x), \\ \text{s.t.} & x \in X \end{array}$$

respectively, are connected by the simple relation $v_{\max} = -v_{\min}$, and x_0 is a maximum point of f if and only if x_0 is a minimum point of $-f$. For this reason, we usually only formulate results for minimization problems.

Finally, a comment as to why we allow $+\infty$ and $-\infty$ as function values of the objective functions as this seems to complicate matters. The most important reason is that sometimes we have to consider functions that are defined as pointwise suprema of an infinite family of functions, and the supremum function may assume infinite values even if all functions in the family assume only finite values. The alternative to allowing functions with values in the extended real line would be to restrict the domain of these supremum functions, and this is neither simpler nor more elegant.

Deloitte.
Discover the truth at www.deloitte.ca/careers
© Deloitte & Touche LLP and affiliated entities.
Download free eBooks at bookboon.com
Click on the ad to read more

General comments

There are some general and perhaps completely obvious comments that are relevant for many optimization problems.

Existence of feasible points

This point may seem trivial, for if a problem has no feasible points then there is not much more to be said. It should however be remembered that the set of feasible points is seldom given explicitly. Instead it is often defined by a system of equalities and inequalities, which may not be consistent.

If the problem comes from the "real world", simplifications and defects in the mathematical model may lead to a mathematical problem that lacks feasible points.

Existence of optimal solutions

Needless to say, a prerequisite for the determination of the optimal solution of a problem is that there is one. Many theoretical results are of the form 'If x_0 is an optimal solution, then x_0 satisfies these conditions.' Although this usually restricts the number of potential candidates for optimal points, it does not prove the existence of such points.

From a practical point of view, however, the existence of an optimal solution − and its exact value, if such a solution exists − may not be that important. In many applications one is often satisfied with a feasible solutions that is good enough.

Uniqueness

Is the optimal solution, if such a solution exists, unique? The answer is probably of little interest for somebody looking for the solution of a practical problem − he or she should be satisfied by having found a best solution even if there are other solutions that are just as good. And if he or she would consider one of the optimal solutions better than the others, then we can only conclude that the optimization problem is not properly set from the start, because the objective function apparently does not include everything that is required to sort out the best solution.

However, uniqueness of an optimal solution may sometimes lead to interesting properties that can be of use when looking for the solution.

Download free eBooks at bookboon.com

Dependence on parameters and sensitivity

Sometimes, and in particular in problems that come directly from ”reality”, objective functions and constraints contain parameters, which are only given with a certain accuracy and, in the worst case, are more or less coarse estimates. In such cases, it is not sufficient to determine the optimal solution, but it is at least as important to know how the solution changes when parameters are changed. If a small perturbation of one parameter alters the optimal solution very much, there is reason to consider the solution with great skepticism.

Qualitative aspects

Of course, it is only for a small class of optimization problems that one can specify the optimum solution in exact form, or where the solution can be described by an algorithm that terminates after finitely many iterations. The mathematical solution to an optimization problem often consists of a number of necessary and/or sufficient conditions that the optimal solution must meet. At best, these can be the basis for useful numerical algorithms, and in other cases, they can perhaps only be used for qualitative statements about the optimal solutions, which however in many situations can be just as interesting.

Algorithms

There is of course no numerical algorithm that solves all optimization problems, even if we restrict ourselves to problems where the constraint set is defined by a a finite number of inequalities and equalities. However, there are very efficient numerical algorithms for certain subclasses of optimization problems, and many important applied optimization problems happen to belong to these classes. We shall study some algorithms of this type in the last chapter of this Part II and in Part III.

The development of good algorithms has been just as important as the computer development for the possibility of solving big optimization problems, and much of the algorithm development has occurred in recent decades.

9.2 Classification of optimization problems

To be able to say anything sensible about the minimization problem

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in X \end{array}$$

Download free eBooks at bookboon.com

we must make various assumptions about the objective function $f\colon \Omega \to \overline{\mathbf{R}}$ and about the set X of feasible points.

We will always assume that Ω is a subset of $\mathbf{R}^n$ and that the set X can be expressed as the solution set of a number of inequalities and equalities, i.e. that

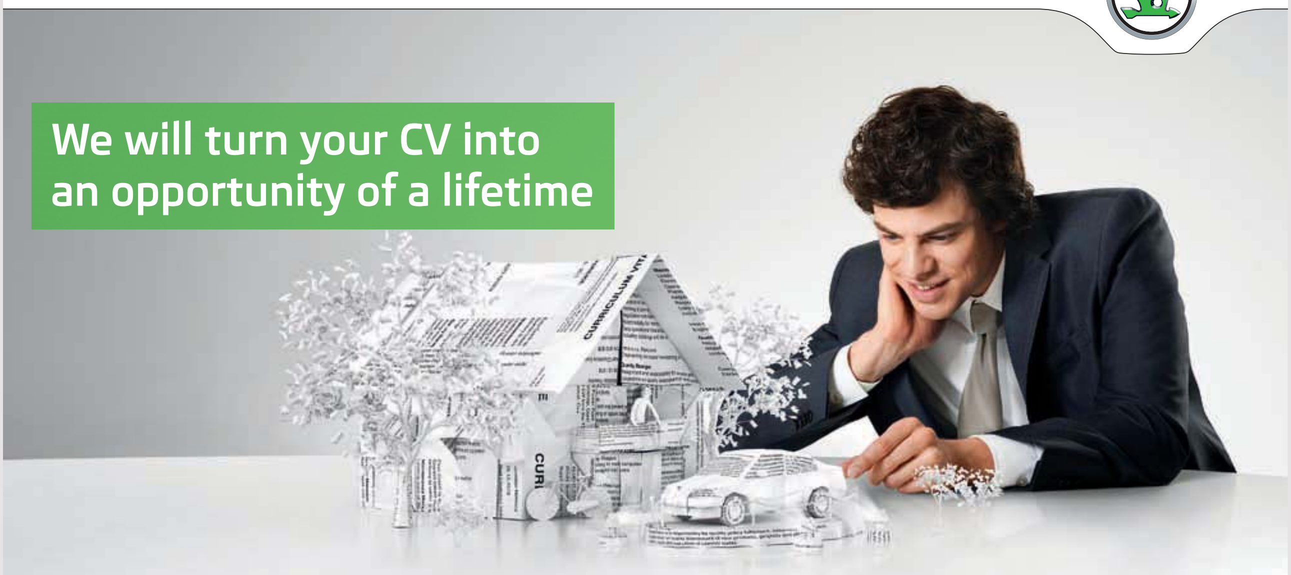
$$X = \{x \in \Omega \mid g_1(x) \le 0,\ \dots,\ g_p(x) \le 0,\ g_{p+1}(x) = 0,\ \dots,\ g_m(x) = 0\}$$

where $g_1, g_2, \dots, g_m$ are real valued functions defined on Ω.

We do not exclude the possibility that all constraints are equalities, i.e. that $p = 0$, or that all constraints are inequalities, i.e. that $p = m$, or that there are no constraints at all, i.e. that $m = 0$.

Since the equality $h(x) = 0$ can be replaced by the two inequalities $\pm h(x) \le 0$, we could without loss of generality assume that all constraints are inequalities, but it is convenient to formulate results for optimization problems with equalities among the constraints without first having to make such rewritings.

If $\hat{x}$ is a feasible point and $g_i(\hat{x}) = 0$, we say that the i:th constraint is *active* at the point $\hat{x}$. All constraints in the form of equalities are, of course, active at all feasible points.

SIMPLY CLEVER
ŠKODA


Do you like cars? Would you like to be a part of a successful brand?
We will appreciate and reward both your enthusiasm and talent.
Send us your CV. You will be surprised where it can take you.
Send us your CV on
www.employerforlife.com
Download free eBooks at bookboon.com
Click on the ad to read more

The condition $x \in \Omega$ is (in the case $\Omega \neq \mathbf{R}^n$) of course also a kind of constraint, but it plays a different role than the other constraints. We will sometimes call it the *implicit* constraint in order to distinguish it from the other *explicit* constraints. If Ω is given as the solution set of a number of inequalities of type $h_i(x) \leq 0$ and the functions h_i, the objective function and the explicit constraint functions are defined on the entire space $\mathbf{R}^n$, we can of course include the inequalities $h_i(x) \leq 0$ among the explicit conditions and omit the implicit constraint.

The domain Ω will often be clear from the context, and it is in these cases not mentioned explicitly in the formulation of the optimization problem. The minimization problem (P) will therefore often be given in the following form

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & \begin{cases} g_i(x) \leq 0, & i = 1, 2, \dots, p \\ g_i(x) = 0, & i = p+1, \dots, m. \end{cases} \end{array}$$

Linear programming

The problem of maximizing or minimizing a linear form over a polyhedron, which is given in the form of an intersection of closed halvspaces in $\mathbf{R}^n$, is called *linear programming*, abbreviated LP. The problem (P) is, in other words, an LP problem if the objective function f is linear and X is the set of solutions to a finite number of linear equalities and inequalities.

We will study LP problems in detail in Chapter 12.

Convex optimization

The minimization problem

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & \begin{cases} g_i(x) \leq 0, & i = 1, 2, \dots, p \\ g_i(x) = 0, & i = p+1, \dots, m \end{cases} \end{array}$$

with implicit constraint $x \in \Omega$ is called *convex*, if the set Ω is convex, the objective function $f \colon \Omega \to \overline{\mathbf{R}}$ is convex, and the constraint functions g_i are convex for $i = 1, 2, \dots, p$ and affine for $i = p+1, \dots, m$.

The affine conditions $g_{p+1}(x) = 0, \dots, g_m(x) = 0$ in a convex problem can of course be summarized as $Ax = b$, where A is an $(m-p) \times n$-matrix.

The set X of feasible points is convex in a convex minimization problem, for

$$X = \bigcap_{i=1}^{p} \{x \in \Omega \mid g_i(x) \leq 0\} \cap \bigcap_{i=p+1}^{m} \{x \mid g_i(x) = 0\},$$

Download free eBooks at bookboon.com

and this expresses X as an intersection of sublevel sets of convex functions and hyperplanes.

A maximization problem

$$\begin{array}{ll} \max & f(x) \\ \text{s.t.} & x \in X \end{array}$$

is called convex if the corresponding equivalent minimization problem

$$\begin{array}{ll} \min & -f(x) \\ \text{s.t.} & x \in X \end{array}$$

is convex, which means that the objective function f has to be concave.

LP problems are of course convex optimization problems. General convex optimization problems are studied in Chapter 11.

Convex quadratic programming

We get a special case of convex optimization if X is a polyhedron and the objective function f is a sum of a linear form and a positive semidefinite quadratic form, i.e. has the form $f(x) = \langle c, x\rangle + \langle x, Qx\rangle$, where Q is a positive semidefinite matrix. The problem (P) is then called *convex quadratic programming.* LP problems constitute a subclass of the convex quadratic problems, of course.

Non-linear optimization

Non-linear optimization is about optimization problems that are not supposed to be LP problems. Since non-linear optimization includes almost everything, there is of course no general theory that can be applied to an arbitrary non-linear optimization problem.

If f is a differentiable function and X is a "decent" set in $\mathbf{R}^n$, one can of course use differential calculus to attack the minimization problem (P). We recall in this context the Lagrange theorem, which gives a necessary condition for the minimum (and maximum) when

$$X = \{x \in \mathbf{R}^n \mid g_1(x) = g_2(x) = \cdots = g_m(x) = 0\}.$$

A counterpart of Lagrange's theorem for optimization problems with constraints in the form of inequalities is given in Chapter 10.

Integer programming

An *integer programming* problem is a mathematical optimization problem in which some or all of the variables are restricted to be integers. In particular,

Download free eBooks at bookboon.com

a *linear integer problem* is a problem of the form

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & x \in X \cap (\mathbf{Z}^m \times \mathbf{R}^{n-m}) \end{array}$$

where $\langle c, x\rangle$ is a linear form and X is a polyhedron in $\mathbf{R}^n$.

Many problems dealing with flows in networks, e.g. commodity distribution problems and maximum flow problems, are linear integer problems that can be solved using special algorithms.

Simultaneous optimization

The title refers to a type of problems that are not really optimization problems in the previous sense. There are many situations, where an individual may affect the outcome through his actions without having full control over the situation. Some variables may be in the hands of other individuals with completely different desires about the outcome, while other variables may be of a completely random nature. The problem to in some sense optimize the outcome could then be called *simultaneous optimization*.

Download free eBooks at bookboon.com
Click on the ad to read more

Simultaneous optimization is the topic of *game theory*, which deals with the behavior of the various agents in conflict situations. Game theoretical concepts and results have proved to be very useful in various contexts, e.g. in economics.

9.3 Equivalent problem formulations

Let us informally call two optimization problems *equivalent* if it is possible to determine in an automatical way an optimal solution to one of the problems, given an optimal solution to the other, and vice versa.

A trivial example of equivalent problems are, as already mentioned, the problems

$$\begin{array}{ll} \max & f(x) \\ \text{s.t.} & x \in X \end{array} \qquad \text{and} \qquad \begin{array}{ll} \min & -f(x). \\ \text{s.t.} & x \in X \end{array}$$

We now describe some useful transformations that lead to equivalent optimization problem

Elimination of equalities

Consider the problem

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & \begin{cases} g_i(x) \le 0, & i = 1, 2, \dots, p \\ g_i(x) = 0, & i = p+1, \dots, m. \end{cases} \end{array}$$

If it is possible to solve the subsystem of equalities and express the solution in the form $x = h(y)$ with a parameter y running over some subset of $\mathbf{R}^d$, then we can eliminate the equalities and rewrite problem (P) as

$$\text{(P}'\text{)} \qquad \begin{array}{ll} \min & f(h(y)) \\ \text{s.t.} & g_i(h(y)) \le 0, \quad i = 1, 2, \dots, p \end{array}$$

If $\hat{y}$ is an optimal solution to (P$'$), then $h(\hat{y})$ is of course an optimal solution to (P). Conversely, if $\hat{x}$ is an optimal solution to (P), then $\hat{x} = h(\hat{y})$ for some value $\hat{y}$ of the parameter, and this value is an optimal solution to (P$'$).

The elimination is always possible (by a simple algorithm) if all constraint equalities are affine, i.e. if the system can be written in the form $Ax = b$ for some $(m-p) \times n$-matrix A. Assuming that the system is consistent, the solution set is an affine subspace of dimension $d = n - \operatorname{rank} A$, and there exists an $n \times d$-matrix C of rank d and a particular solution x_0 to the system

Download free eBooks at bookboon.com

such that $Ax = b$ if and only if $x = Cy + x_0$ for some $y \in \mathbf{R}^d$. The problem (P) is thus in this case equivalent to the problem

$$\begin{array}{ll} \min & f(Cy + x_0) \\ \text{s.t.} & g_i(Cy + x_0) \le 0, \qquad i = 1, 2, \dots, p \end{array}$$

(with implicit constraint $Cy + x_0 \in \Omega$).

In convex optimization problems, and especially in LP problems, we can thus, in principle, eliminate the equalities from the constraints and in this way replace the problem by an equivalent optimization problem without any equality constraints.

Slack variables

The inequality $g(x) \le 0$ holds if and only if there is a number $s \ge 0$ such that $g(x) + s = 0$. By thus replacing all inequalities in the problem

(P)
$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & \begin{cases} g_i(x) \le 0, & i = 1, 2, \dots, p \\ g_i(x) = 0, & i = p+1, \dots, m \end{cases} \end{array}$$

with equalities, we obtain the following equivalent problem

(P′)
$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & \begin{cases} g_i(x) + s_i = 0, & i = 1, 2, \dots, p \\ g_i(x) = 0, & i = p+1, \dots, m \\ s_i \ge 0, & i = 1, 2, \dots, p \end{cases} \end{array}$$

with $n + p$ variables, m equality constraints and p simple inequality constraints. The new variables s_i are called *slack variables*.

If $\hat{x}$ is an optimal solution to (P), we get an optimal solution $(\hat{x}, \hat{s})$ to (P′) by setting $\hat{s}_i = -g_i(\hat{x})$. Conversely, if $(\hat{x}, \hat{s})$ is an optimal solution to the last mentioned problem, then $\hat{x}$ is of course an optimal solution to the original problem.

If the original constraints are affine, then so are all new constraints. The transformation thus transforms LP problems to LP problems.

Inequalities of the form $g(x) \ge 0$ can of course similarly be written as equalities $g(x) - s = 0$ with nonnegative variables s. These new variables are usually called *surplus variables*.

Download free eBooks at bookboon.com

Nonnegative variables

Every real number can be written as a difference between two nonnegative numbers. In an optimization problem, we can thus replace an unrestricted variable x_i, i.e. a variable that a priori may assume any real value, with two nonnegative variables x_i' and x_i'' by setting

$$x_i = x_i' - x_i'', \quad x_i' \geq 0, \; x_i'' \geq 0.$$

The number of variables increases with one and the number of inequalities increases with two for each unrestricted variable that is replaced, but the transformation leads apparently to an equivalent problem. Moreover, convex problems are transfered to convex problems and LP problems are transformed to LP problems.

Example 9.3.1. The LP problem

$$\begin{array}{ll} \min & x_1 + 2x_2 \\ \text{s.t.} & \begin{cases} x_1 + x_2 \geq 2 \\ 2x_1 - x_2 \leq 3 \\ x_1 \geq 0 \end{cases} \end{array}$$

Download free eBooks at bookboon.com

Click on the ad to read more

is transformed, using two slack/surplus variables and by replacing the unrestricted variable x_2 with a difference of two nonnegative variables, to the following equivalent LP problem in which all variables are nonnegative and all remaining constraints are equalities.

$$\begin{array}{ll} \min & x_1 + 2x_2' - 2x_2'' + 0s_1 + 0s_2 \\ \text{s.t.} & \begin{cases} x_1 + x_2' - x_2'' - s_1 \qquad = 2 \\ 2x_1 - x_2' + x_2'' \qquad + s_2 = 3 \\ \qquad x_1, x_2', x_2'', s_1, s_2 \geq 0. \end{cases} \end{array} \qquad \square$$

Epigraph form

Every optimization problem can be replaced by an equivalent problem with a linear objective function, and the trick to accomplish this is to utilize the epigraph of the original objective function. The two problems

$$\text{(P)} \quad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in X \end{array} \qquad \text{and} \qquad \text{(P}'\text{)} \quad \begin{array}{ll} \min & t \\ \text{s.t.} & \begin{cases} f(x) \leq t \\ x \in X \end{cases} \end{array}$$

are namely equivalent, and the objective function in (P′) is linear. If $\hat{x}$ is an optimal solution to (P), then $(\hat{x}, f(\hat{x}))$ is an optimal solution to (P′), and if $(\hat{x}, \hat{t})$ is an optimal solution to (P′), then $\hat{x}$ is an optimal solution to (P).

If problem (P) is convex, i.e. has the form

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & \begin{cases} g_i(x) \leq 0, & i = 1, 2, \dots, p \\ g_i(x) = 0, & i = p+1, \dots, m \end{cases} \end{array}$$

with convex functions f and g_i for $1 \leq i \leq p$, and affine functions g_i for $i \geq p+1$, then the epigraph variant

$$\begin{array}{ll} \min & t \\ \text{s.t.} & \begin{cases} f(x) - t \leq 0, & \\ g_i(x) \leq 0, & i = 1, 2, \dots, p \\ g_i(x) = 0, & i = p+1, \dots, m \end{cases} \end{array}$$

is also a convex problem.

So there is no restriction to assume that the objective function of a convex program is linear when we are looking for general properties of such programs.

Download free eBooks at bookboon.com

Piecewise affine objective functions

Suppose that X is a polyhedron (given as an intersection of closed halfspaces) and consider the convex optimization problem

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in X \end{array}$$

where the objective function $f(x)$ is piecewise affine and given as

$$f(x) = \max\{\langle c_i, x\rangle + b_i \mid i = 1, 2, \dots, m\}.$$

The epigraph transformation results in the equivalent convex problem

$$\begin{array}{ll} \min & t \\ \text{s.t.} & \begin{cases} \max\limits_{1\le i\le m} (\langle c_i, x\rangle + b_i) \le t \\ \hfill x \in X, \end{cases} \end{array}$$

and since $\max_{1\le i\le m} \alpha_i \le t$ if and only if $\alpha_i \le t$ for all i, this problem is in turn equivalent to the LP problem

$$\text{(P}'\text{)} \qquad \begin{array}{ll} \min & t \\ \text{s.t.} & \begin{cases} \langle c_i, x\rangle - t + b_i \le 0, & i = 1, 2, \dots, m \\ \qquad x \in X. \end{cases} \end{array}$$

The constraint set of this LP problem is a polyhedron in $\mathbf{R}^n \times \mathbf{R}$.

If instead the objective function in problem (P) is a sum

$$f(x) = f_1(x) + f_2(x) + \cdots + f_k(x)$$

of piecewise affine functions f_i, then problem (P) is equivalent to the convex problem

$$\begin{array}{ll} \min & t_1 + t_2 + \cdots + t_k \\ \text{s.t.} & \begin{cases} f_i(x) \le t_i & i = 1, 2, \dots, k \\ \quad x \in X \end{cases} \end{array}$$

and this problem becomes an LP problem if every inequality $f_i(x) \le t_i$ is expressed as a system of linear inequalities in a similar way as above.

Download free eBooks at bookboon.com

9.4 Some model examples

Diet problem

Let us start with a classical LP problem that was formulated and studied during the childhood of linear programming. The goal of the diet problem is to select a set of foods that will satisfy a set of daily nutritional requirements at minimum cost. There are n foods $L_1, L_2, \dots, L_n$ available at a cost of $c_1, c_2, \dots, c_n$ dollars per unit. The foods contain various nutrients $N_1, N_2, \dots, N_m$ (proteins, carbohydrates, fats, vitamins, etc.). The number of units of nutrients per unit of food is shown by the following table:

	L_1	L_2	$\dots$	L_n
N_1	a_{11}	a_{12}	$\dots$	a_{1n}
N_2	a_{21}	a_{22}	$\dots$	a_{2n}
$\vdots$				
N_m	a_{m1}	a_{m2}	$\dots$	a_{mn}

Buying $x_1, x_2, \dots, x_n$ units of the foods, one thus obtains

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n$$

Download free eBooks at bookboon.com

Click on the ad to read more

units of nutrient N_i at a cost of

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n.$$

Suppose that the daily requirement of the different nutrients is $b_1, b_2, \dots, b_m$ and that it is not harmful to have too much of any substance. The problem to meet the daily requirement at the lowest possible cost is called the *diet problem*. Mathematically, it is of the form

$$\begin{array}{ll} \min & c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ \text{s.t.} & \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \geq b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \geq b_2 \\ \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \geq b_m \\ x_1, x_2, \dots, x_n \geq 0. \end{cases} \end{array}$$

The diet problem is thus an LP problem. In addition to determining the optimal diet and the cost of this, it would be of interest to answer the following questions:

1. How does a price change of one or more of the foods affect the optimal diet and the cost?
2. How is the optimal diet affected by a change of the daily requirement of one or more nutrients?
3. Suppose that pure nutrients are available on the market. At what price would it be profitable to buy these and satisfy the nutritional needs by eating them instead of the optimal diet? Hardly a tasty option for a gourmet but perhaps possible in animal feeding.

Assume that the cost of the optimal diet is z, and that its cost changes to $z + \Delta z$ when the need for nutrient N_1 is changed from b_1 to $b_1 + \Delta b_1$, ceteris paribus. It is obvious that the cost can not be reduced when demand increases, so therefore $\Delta b_1 > 0$ entails $\Delta z \geq 0$. If it is possible to buy the nutrient N_1 in completely pure form to the price p_1, then it is economically advantageous to meet the increased need by taking the nutrient in pure form, provided that $p_1 \Delta b_1 \leq \Delta z$. The maximum price of N_1 which makes nutrient in pure form an economical alternative is therefore $\Delta z / \Delta b_1$, and the limit as $\Delta b_1 \to 0$, i.e. the partial derivative $\frac{\partial z}{\partial b_1}$, is called the *dual price* or the *shadow price* in economic literature.

It is possible to calculate the nutrient shadow prices by solving an LP problem closely related to the diet problem. Assume again that the market provides nutrients in pure form and that their prices are $y_1, y_2, \dots, y_m$. Since one unit of food L_i contains $a_{1i}, a_{2i}, \dots, a_m$ units of each nutrient, we can

Download free eBooks at bookboon.com

"manufacture" one unit of food L_i by buying just this set of nutrients, and hence it is economically advantageous to replace all foods by pure nutrients if

$$a_{1i}y_1 + a_{2i}y_2 + \cdots + a_m y_m \le c_i$$

for $i = 1, 2, \ldots, n$. Under these conditions the cost of the required daily ration $b_1, b_2, \ldots, b_m$ is at most equal to the maximum value of the LP problem

$$\begin{array}{ll}
\max & b_1y_1 + b_2y_2 + \cdots + b_my_m \\
\text{s.t.} & \begin{cases} a_{11}y_1 + a_{21}y_2 + \ldots + a_{m1}y_m \le c_1 \\ a_{12}y_1 + a_{22}y_2 + \ldots + a_{m2}y_m \le c_2 \\ \qquad\vdots \\ a_{1n}y_1 + a_{2n}y_2 + \ldots + a_{mn}y_m \le c_n \\ y_1,\, y_2, \ldots, y_m \ge 0. \end{cases}
\end{array}$$

We will show that this so called *dual problem* has the same optimal value as the original diet problem and that the optimal solution is given by the shadow prices.

Production planning

Many problems related to production planning can be formulated as LP problems, and a pioneer in the field was the Russian mathematician and economist Leonid Kantorovich, who studied and solved such problems in the late 1930s. Here is a typical such problem.

A factory can manufacture various goods $V_1, V_2, \ldots, V_n$. This requires various inputs (raw materials and semi-finished goods) and different types of labor, something which we collectively call production factors $P_1, P_2, \ldots, P_m$. These are available in limited quantities $b_1, b_2, \ldots, b_m$. In order to manufacture, market and sell one unit of the respective goods, production factors are needed to an extent given by the following table:

	V_1	V_2	$\ldots$	V_n
P_1	a_{11}	a_{12}	$\ldots$	a_{1n}
P_2	a_{21}	a_{22}	$\ldots$	a_{2n}
$\vdots$				
P_m	a_{m1}	a_{m2}	$\ldots$	a_{mn}

Every manufactured product V_j can be sold at a profit which is c_j dollars per unit, and the goal now is to plan the production $x_1, x_2, \ldots, x_n$ of the various products so that the profit is maximized.

Download free eBooks at bookboon.com

Manufacturing $x_1, x_2, \dots, x_n$ units of the goods consumes $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$ units of production factor P_i and results in a profit equal to $c_1x_1 + c_2x_2 + \cdots + c_nx_n$. The optimization problem that we need to solve is thus the LP problem

$$
\begin{array}{ll}
\max & c_1x_1 + c_2x_2 + \cdots + c_nx_n \\
\text{s.t.} & \begin{cases}
a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \le b_1 \\
a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \le b_2 \\
\qquad\qquad\qquad\qquad\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \le b_m \\
\qquad\qquad\quad x_1,\, x_2, \dots, x_n \ge 0.
\end{cases}
\end{array}
$$

Here it is reasonable to ask similar questions as for the diet problem, i.e. how is the optimal solution and the optimal profit affected by

1. altered pricing $c_1, c_2, \dots, c_n$;
2. changes in the resource allocation.

If we increase a resource P_i that is already fully utilized, so does (normally) the profit. What will the price of this resource be for the expansion to pay off? The critical price is called the shadow price, and it can be interpreted as a partial derivative, and as the solution to a dual problem.

Transportation problem

The transportation problem is another classical LP problem that was formulated and solved before the invention of the simplex algorithm

A commodity (e.g. gasoline) is stored at m places $S_1, S_2, \dots, S_m$ and demanded at n other locations $D_1, D_2, \dots, D_n$. The quantity of the commodity available at S_i is a_i units, while b_j units are demanded at D_j. To ship 1 unit from storage place S_i to demand center D_j costs c_{ij} dollars.

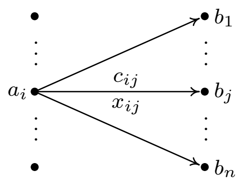


Figure 9.1. The transportation problem

The total supply, i.e. $\sum_{i=1}^m a_i$, is assumed for simplicity to be equal to the total demand $\sum_{j=1}^n b_j$, so it is possible to meet the demand by distributing x_{ij} units from S_i to D_j. To do this at the lowest transportation cost gives rise to the LP problem

$$\begin{array}{ll} \min & \displaystyle\sum_{i=1}^m \sum_{j=1}^n c_{ij}x_{ij} \\ \text{s.t.} & \begin{cases} \sum_{j=1}^n x_{ij} = a_i, & i = 1, 2, \dots, m \\ \sum_{i=1}^m x_{ij} = b_j, & j = 1, 2, \dots, n \\ x_{ij} \geq 0, & \text{all } i, j. \end{cases} \end{array}$$

An investment problem

An investor has 1 million dollars, which he intends to invest in various projects, and he has found m interesting candidates $P_1, P_2, \dots, P_m$ for this. The return will depend on the projects and the upcoming economic cycle. He thinks he can identify n different economic situations $E_1, E_2, \dots, E_n$, but it is impossible for him to accurately predict what the economy will look like in the coming year, after which he intends to collect the return. However, one can accurately assess the return of each project during the various economic cycles; each invested million dollars in project P_i will yield a return

Download free eBooks at bookboon.com

of a_{ij} million dollars during business cycle E_j. We have, in other words, the following table of return for various projects and business cycles:

	E_1	E_2	$\dots$	E_n
P_1	a_{11}	a_{12}	$\dots$	a_{1n}
P_2	a_{21}	a_{22}	$\dots$	a_{2n}
$\vdots$				
P_m	a_{m1}	a_{m2}	$\dots$	a_{mn}

Our investor intends to invest $x_1, x_2, \dots, x_m$ million dollars in the various projects, and this will give him the return

$$a_{1j}x_1 + a_{2j}x_2 + \dots + a_{mj}x_m$$

million dollars, assuming that the economy will be in state E_j. Since our investor is a very cautious person, he wants to guard against the worst possible outcome, and the worst possible outcome for the investment $x_1, x_2, \dots, x_m$ is

$$\min_{1\le j\le n} \sum_{i=1}^m a_{ij}x_i.$$

He therefore wishes to maximize this outcome, which he does by solving the problem

$$\begin{array}{ll} \max & \displaystyle\min_{1\le j\le n} \sum_{i=1}^m a_{ij}x_i \\ \text{s.t.} & x \in X \end{array}$$

where X is the set $\{(x_1, x_2, \dots, x_m) \in \mathbf{R}^m_+ \mid \sum_{i=1}^m x_i = 1\}$ of all possible ways to distribute one million on the various projects.

In this formulation, the problem is a convex maximization problem with a piecewise affine concave objective function. However, we can transform it into an equivalent LP problem by making use of a hypograph formulation. Utilizing the techniques of the previous section, we see that the investor's problem is equivalent to the LP problem

$$\begin{array}{ll} \max & v \\ \text{s.t.} & \begin{cases} a_{11}x_1 + a_{21}x_2 + \ldots + a_{m1}x_m \ge v \\ a_{12}x_1 + a_{22}x_2 + \ldots + a_{m2}x_m \ge v \\ \qquad\qquad\qquad\qquad\qquad\quad \vdots \\ a_{1n}x_1 + a_{2n}x_2 + \ldots + a_{mn}x_m \ge v \\ \quad x_1 + \quad\ x_2 + \ldots + \qquad x_m = 1 \\ \qquad\qquad\qquad x_1,\ x_2, \dots, x_m \ge 0. \end{cases} \end{array}$$

Download free eBooks at bookboon.com

Two-person zero-sum game

Two persons, row player Rick and column player Charlie, each choose, independently of each other, an integer. Rick chooses a number i in the range $1 \leq i \leq m$ and Charlie a number j in the range $1 \leq j \leq n$. If they choose the pair (i, j), Rick wins a_{ij} dollars of Charlie, and to win a negative amount is of course the same as to loose the corresponding positive amount.

The numbers m, n and a_{ij} are supposed to be known by both players, and the objective of each player is to win as much as possible (or equivalently, to loose as little as possible). There is generally no best choice for any of the players, but they could try to maximize their expected winnings by selecting their numbers at random with a certain probability distribution.

Suppose Rick chooses the number i with probability x_i, and Charlie chooses the number j with probability y_j. All probabilities are of course nonnegative numbers, and $\sum_{i=1}^m x_i = \sum_{j=1}^n y_j = 1$. Let

$$X = \{x \in \mathbf{R}^m_+ \mid \sum_{i=1}^m x_i = 1\} \quad \text{and} \quad Y = \{y \in \mathbf{R}^n_+ \mid \sum_{j=1}^n y_j = 1\}.$$

The elements in X are called the row player's *mixed strategies*, and the elements in Y are the column player's mixed strategies.

Download free eBooks at bookboon.com
Click on the ad to read more

Since the players choose their numbers independently of each other, the outcome (i, j) will occur with probability $x_i y_j$. Rick's pay-off is therefore a random variable with expected value

$$f(x, y) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} x_i y_j.$$

Row player Rick can now conceivably argue like this: "The worst that can happen to me, if I choose the probability distribution x, is that my opponent Charlie happens to choose a probability distribution y that minimizes my expected profit $f(x, y)$". In this case, Rick will obtain the amount

$$g(x) = \min_{y\in Y} f(x, y) = \min_{y\in Y} \sum_{j=1}^{n} y_j \Big(\sum_{i=1}^{m} a_{ij} x_i\Big).$$

The sum $\sum_{j=1}^{n} y_j \big(\sum_{i=1}^{m} a_{ij} x_i\big)$ is a weighted arithmetic mean of the n numbers $\sum_{i=1}^{m} a_{ij} x_i$, $j = 1, 2, \dots, n$, with the weights $y_1, y_2, \dots, y_n$, and such a mean is greater than or equal to the smallest of the n numbers, and equality is obtained by putting all weight on this smallest number. Hence,

$$g(x) = \min_{1\le j\le n} \sum_{i=1}^{m} a_{ij} x_i.$$

Rick, who wants to maximize his outcome, should therefore choose to maximize $g(x)$, i.e. Rick's problem becomes

$$\begin{array}{ll} \max & g(x) \\ \text{s.t.} & x \in X. \end{array}$$

This is exactly the same problem as the investor's problem. Hence, Rick's optimal strategy, i.e. optimal choice of probabilities, coincides with the optimal solution to the LP problem

$$\begin{array}{ll} \max & v \\ \text{s.t.} & \begin{cases} a_{11}x_1 + a_{21}x_2 + \ldots + a_{m1}x_m \ge v \\ a_{12}x_1 + a_{22}x_2 + \ldots + a_{m2}x_m \ge v \\ \qquad\qquad\qquad\qquad\qquad\qquad\quad \vdots \\ a_{1n}x_1 + a_{2n}x_2 + \ldots + a_{mn}x_m \ge v \\ \quad x_1 + \quad x_2 + \ldots + \quad x_m = 1 \\ \qquad\qquad\quad x_1,\, x_2, \dots, x_m \ge 0. \end{cases} \end{array}$$

Download free eBooks at bookboon.com

The column player's problem is analogous, but he will of course minimize the maximum expected outcome $f(x, y)$. Charlie must therefore solve the problem

$$\begin{array}{ll} \min & \displaystyle\max_{1\le i\le m} \sum_{j=1}^{n} a_{ij}y_j \\ \text{s.t.} & y \in Y \end{array}$$

to find his optimal strategy, and this problem is equivalent to the LP problem

$$\begin{array}{ll} \min & u \\ \text{s.t.} & \begin{cases} a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n \le u \\ a_{21}y_1 + a_{22}y_2 + \dots + a_{2n}y_n \le u \\ \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ a_{m1}y_1 + a_{m2}y_2 + \dots + a_{mn}y_n \le u \\ y_1 + y_2 + \dots + y_n = 1 \\ y_1,\, y_2, \dots, y_n \ge 0. \end{cases} \end{array}$$

The two players' problems are examples of dual problems, and it follows from results that will appear in Chapter 12 that they have the same optimal value.

Consumer Theory

The behavior of consumers is studied in a branch of economics known as microeconomics. Assume that there are n commodities $V_1, V_2, \dots, V_n$ on the market and that the price of these goods is given by the price vector $p = (p_1, p_2, \dots, p_n)$. A basket x consisting of $x_1, x_2, \dots, x_n$ units of the goods thus costs $\langle p, x\rangle = p_1x_1 + p_2x_2 + \dots + p_nx_n$.

A consumer values her benefit of the commodity bundle x by using a subjective utility function f, where $f(x) > f(y)$ means that she prefers x to y. A reasonable assumption about the utility function is that every convex combination $\lambda x + (1 - \lambda)y$ of two commodity bundles should be valued as being at least as good as the worst of the two bundles x and y, i.e. that $f(\lambda x + (1 - \lambda)y) \ge \min\bigl(f(x), f(y)\bigr)$. The utility function f is assumed, in other words, to be quasiconcave, and a stronger assumption, which is often made in the economic literature and that we are making here, is that f is concave.

Suppose now that our consumer's income is I, that the entire income is disposable for consumption, and that she wants to maximize her utility.

Download free eBooks at bookboon.com

Then, the problem that she needs to solve is the convex optimization problem

$$\begin{array}{ll} \max & f(x) \\ \text{s.t.} & \begin{cases} \langle p, x\rangle \leq I \\ \quad x \geq 0. \end{cases} \end{array}$$

To determine empirically a consumer's utility function is of course almost impossible, so microtheory is hardly useful for quantitative calculations. However, one can make qualitative analyzes and answer questions of the type: How does an increase in income change the consumer behavior? and How does changes in the prices of the goods affect the purchasing behavior?

Portfolio optimization

A person intends to buy shares in n different companies $C_1, C_2, \dots, C_n$ for S dollars. One dollar invested in the company C_j gives a return of R_j dollars, where R_j is a random variable with known expected value

$$\mu_j = \mathrm{E}[R_j].$$

The covariances

$$\sigma_{ij} = \mathrm{E}[(R_i - \mu_i)(R_j - \mu_j)]$$

are also assumed to be known.

Download free eBooks at bookboon.com
Click on the ad to read more

The expected total return $e(x)$ from investing $x = (x_1, x_2, \dots, x_n)$ dollars in the companies $C_1, C_2, \dots, C_n$ is given by

$$e(x) = \mathrm{E}\Big[\sum_{j=1}^{n} x_j R_j\Big] = \sum_{j=1}^{n} \mu_j x_j,$$

and the variance of the total return is

$$v(x) = \mathrm{Var}\,\Big[\sum_{j=1}^{n} x_j R_j\Big] = \sum_{i,j=1}^{n} \sigma_{ij} x_i x_j.$$

Note that $v(x)$ is a positive semi-definite quadratic form.

It is not possible for our person to maximize the total return, because the return is a random variable, i.e. depends on chance. However, he can maximize the expected total return under appropriate risk conditions, i.e. requirements for the variance. Alternatively, he can minimize the risk with the investment given certain requirements on the expected return. Thus there are several possible strategies, and we will formulate three such.

(i) The strategy to maximize the expected total return, given an upper bound B on the variance, leads to the convex optimization problem

$$\begin{array}{ll} \max & e(x) \\ \text{s.t.} & \begin{cases} \phantom{x_1 + x_2 + \cdots +{}} v(x) \le B \\ x_1 + x_2 + \cdots + x_n = S \\ \phantom{x_1 + x_2 + \cdots + x_n{}} x \ge 0. \end{cases} \end{array}$$

(ii) The strategy to minimize the variance of the total return, given a lower bound b on the expected return, gives rise to the convex quadratic programming problem

$$\begin{array}{ll} \min & v(x) \\ \text{s.t.} & \begin{cases} \phantom{x_1 + x_2 + \cdots +{}} e(x) \ge b \\ x_1 + x_2 + \cdots + x_n = S \\ \phantom{x_1 + x_2 + \cdots + x_n{}} x \ge 0. \end{cases} \end{array}$$

(iii) The two strategies can be considered together in the following way. Let $\epsilon \ge 0$ be a (subjective) parameter, and consider the convex quadratic problem

$$\begin{array}{ll} \min & \epsilon v(x) - e(x) \\ \text{s.t.} & \begin{cases} x_1 + x_2 + \cdots + x_n = S \\ \phantom{x_1 + x_2 + \cdots + x_n{}} x \ge 0 \end{cases} \end{array}$$

with optimal solution $x(\epsilon)$. We leave as an exercise to show that

$$v(x(\epsilon_1)) \ge v(x(\epsilon_2)) \qquad \text{and} \qquad e(x(\epsilon_1)) \ge e(x(\epsilon_2))$$

Download free eBooks at bookboon.com

if $0 \leq \epsilon_1 \leq \epsilon_2$. The parameter ϵ is thus a measure of the person's attitude towards risk; the smaller the ϵ, the greater the risk (= variance) but also the greater expected return.

Snell's law of refraction

We will study the path of a light beam which passes through n parallel transparent layers. The j:th slice S_j is assumed to be a_j units wide and to consist of a homogeneous medium in which the speed of light is v_j. We choose a coordinate system as in figure 9.2 and consider a light beam on its path from the origin on the surface of the first slice to a point with y-coordinate b on the outer surface of the last slice.

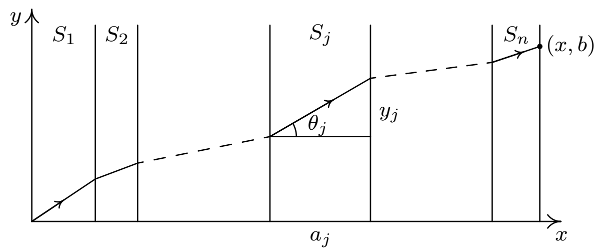


Figure 9.2. The path of a light beam through layers with different refraction indices.

According to Fermat's principle, the light chooses the fastest route. The path of the beam is therefore determined by the optimal solution to the convex optimization problem

$$\begin{aligned} \min \quad & \sum_{j=1}^{n} v_j^{-1} \sqrt{y_j^2 + a_j^2} \\ \text{s.t.} \quad & \sum_{j=1}^{n} y_j = b, \end{aligned}$$

and we obtain *Snell's law of refraction*

$$\frac{\sin \theta_i}{\sin \theta_j} = \frac{v_i}{v_j}$$

by solving the problem.

Download free eBooks at bookboon.com

Overdetermined systems

If a system of linear equations $Ax = b$ with n unknowns and m equations is inconsistent, i.e. has no solutions, you might want to still determine the best approximate solution, i.e. the n-tuple $x = (x_1, x_2, \dots, x_n)$ that makes the error as small as possible. The error is by definition the difference $Ax - b$ between the left and the right hand side of the equation, and as a measure of the size of the error we use $\|Ax - b\|$ for some suitably chosen norm.

The function $x \mapsto \|Ax - b\|$ is convex, so the problem of minimizing $\|Ax - b\|$ over all $x \in \mathbf{R}^n$ is a convex problem regardless of which norm is used, but the solution depends on the norm, of course. Let as usual a_{ij} denote the element at location i, j in the matrix A, and let $b = (b_1, b_2, \dots, b_m)$.

1. The so-called *least square solution* is obtained by using the Euclidean norm $\|\cdot\|_2$. Since $\|Ax - b\|_2^2 = \sum_{i=1}^m (a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n - b_i)^2$, we get the least square solution as the solution of the convex quadratic problem

$$\text{minimize} \sum_{i=1}^{m} (a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n - b_i)^2.$$


In the past four years we have drilled
89,000 km
That's more than **twice** around the world.
Who are we?
We are the world's largest oilfield services company[1]. Working globally—often in remote and challenging locations—we invent, design, engineer, and apply technology to help our customers find and produce oil and gas safely.
Who are we looking for?
Every year, we need thousands of graduates to begin dynamic careers in the following domains:
- Engineering, Research and Operations
- Geoscience and Petrotechnical
- Commercial and Business
What will you be?
Schlumberger
careers.slb.com
[1]Based on Fortune 500 ranking 2011. Copyright © 2015 Schlumberger. All rights reserved.
Download free eBooks at bookboon.com
Click on the ad to read more

The gradient of the objective function is equal to zero at the optimal point, which means that the optimal solution is obtained as the solution to the linear system

$$A^TAx = A^Tb.$$

2. By instead using the $\|\cdot\|_\infty$ norm, one obtains the solution that gives the smallest maximum deviation between the left and the right hand side of the linear system $Ax = b$. Since

$$\|Ax - b\|_\infty = \max_{1\le i\le m} |a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n - b_i|,$$

the objective function is now piecewise affine, and the problem is therefore equivalent to the LP problem

$$\begin{array}{ll} \min & t \\ \text{s.t.} & \begin{cases} \pm(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - b_1) \le t \\ \qquad\vdots \\ \pm(a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n - b_m) \le t. \end{cases} \end{array}$$

3. Instead of minimizing the sum of squares of the differences between left and right sides, we can of course minimize the sum of the absolute value of the differences, i.e. use the $\|\cdot\|_1$-norm. Since the objective function

$$\|Ax - b\|_1 = \sum_{i=1}^{m} |a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n - b_i|$$

is a sum of convex piecewise affine functions, our convex minimization problem is in this case equivalent to the LP problem

$$\begin{array}{ll} \min & t_1 + t_2 + \cdots + t_m \\ \text{s.t.} & \begin{cases} \pm(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - b_1) \le t_1 \\ \qquad\vdots \\ \pm(a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n - b_m) \le t_m. \end{cases} \end{array}$$

Largest inscribed ball

A convex set X with nonempty interior is given in $\mathbf{R}^n$, and we want to determine a ball $B(x, r)$ in X (with respect to a given norm) with the largest possible radius r. We assume that X can be described as the solution set to a system of inequalities, i.e. that

$$X = \{x \in \mathbf{R}^n \mid g_i(x) \le 0,\ i = 1, 2, \ldots, m\},$$

Download free eBooks at bookboon.com

with convex functions g_i.

The ball $B(x,r)$ lies in X if and only if $g_i(x+ry) \le 0$ for all y with $\|y\| \le 1$ and $i = 1, 2, \dots, m$, which makes it natural to consider the functions

$$h_i(x,r) = \sup_{\|y\|\le 1} g_i(x+ry), \quad i = 1, 2, \dots, m.$$

The functions h_i are convex since they are defined as suprema of convex functions in the variables x and r.

The problem of determining the ball with the largest possible radius has now been transformed into the convex optimization problem

$$\begin{aligned} &\max\ r \\ &\text{s.t.} \quad h_i(x,r) \le 0, \quad i = 1, 2, \dots, m. \end{aligned}$$

For general convex sets X, it is of course impossible to determine the functions h_i explicitly, but if X is a polyhedron, $g_i(x) = \langle c_i, x\rangle - b_i$, and the norm in question is the ℓ^p-norm, then it follows from Hölder's inequality that

$$h_i(x,r) = \sup_{\|y\|_p\le 1} (\langle c_i, x\rangle + r\langle c_i, y\rangle - b_i) = \langle c_i, x\rangle + r\|c_i\|_q - b_i$$

for $r \ge 0$, where $\|\cdot\|_q$ denotes the dual norm.

The problem of determining the center x and the radius r of the largest ball that is included in the polyhedron

$$X = \{x \in \mathbf{R}^n \mid \langle c_i, x\rangle \le b_i,\ i = 1, 2, \dots, m\}$$

has now been reduced to the LP problem

$$\begin{aligned} &\max\ r \\ &\text{s.t.} \quad \langle c_i, x\rangle + r\|c_i\|_q \le b_i, \quad i = 1, 2, \dots, m. \end{aligned}$$

Exercises

9.1 In a chemical plant one can use four different processes P_1, P_2, P_3, and P_4 to manufacture the products V_1, V_2, and V_3. Produced quantities of the various products, measured in tons per hour, for the various processes are shown in the following table:

	P_1	P_2	P_3	P_4
V_1	-1	2	2	1
V_2	4	1	0	2
V_3	3	1	2	1

Download free eBooks at bookboon.com

(Process P_1 thus consumes 1 ton of V_1 per hour!) Running processes P_1, P_2, P_3, and P_4 costs 5 000, 4 000, 3 000, and 4 000 dollars per per hour, respectively. The plant intends to produce 16, 40, and 24 tons of products V_1, V_2, and V_3 at the lowest possible cost. Formulate the problem of determining an optimal production schedule.

9.2 Bob has problems with the weather. The weather occurs in the three states pouring rain, drizzle and sunshine. Bob owns a raincoat and an umbrella, and he is somewhat careful with his suit. The raincoat is difficult to carry, and the same applies – though to a lesser degree – to the umbrella; the latter, however, is not fully satisfactory in case of pouring rain. The following table reveals how happy Bob considers himself in the various situations that can arise (the numbers are related to his blood pressure, with 0 corresponding to his normal state).

	Pouring rain	Drizzle	Sunshine
Raincoat	2	1	−2
Umbrella	1	2	−1
Only suit	−4	−2	2

In the morning, when Bob goes to work, he does not know what the weather will be like when he has to go home, and he would therefore choose the clothes that optimize his mind during the walk home. Formulate Bob's problem as an LP problem.

9.3 Consider the following two-person game in which each player has three alternatives and where the payment to the row player is given by the following payoff matrix.

	1	2	3
1	1	0	5
2	3	3	4
3	2	4	0

In this case, it is obvious which alternatives both players must choose. How will they play?

9.4 Charlie and Rick have three cards each. Both have the ace of diamonds and the ace of spades. Charlie also has the two of diamonds, and Rick has the two of spades. The players play simultaneously one card each. Charlie wins if both these cards are of the same color and loses in the opposite case. The winner will receive as payment the value of his winning card from the opponent, with ace counting as 1. Write down the payoff matrix for this two-person game, and formulate column player Charlie's problem to optimize his expected profit as an LP problem.

Download free eBooks at bookboon.com

9.5 The overdetermined system

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 - x_2 = 0 \\ 3x_1 + 2x_2 = 4 \end{cases}$$

has no solution.

a) Determine the least square solution.

b) Formulate the problem of determining the solution that minimizes the maximum difference between the left and the right hand sides of the system.

c) Formulate the problem of determining the solution that minimizes the sum of the absolut values of the differences between the left and the right hand sides.

9.6 Formulate the problem of determining

a) the largest circular disc,

b) the largest square with sides parallel to the coordinate axes,

that is contained in the triangle bounded by the lines $x_1 - x_2 = 0$, $x_1 - 2x_2 = 0$ and $x_1 + x_2 = 1$.

Click on the ad to read more

Chapter 10

The Lagrange function

10.1 The Lagrange function and the dual problem

The Lagrange function

To the minimization problem

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & \begin{cases} g_i(x) \le 0, & i = 1, 2, \dots, p \\ g_i(x) = 0, & i = p+1, \dots, m \end{cases} \end{array}$$

with $x \in \Omega$ as implicit condition and m explicit constraints, the first p of which in the form of inequalities, we shall associate a dual maximization problem, and the tool to accomplish this is the Lagrange function defined below. To avoid trivial matters we assume that $\operatorname{dom} f \neq \emptyset$, i.e. that the objective function $f\colon \Omega \to \overline{\mathbf{R}}$ is not identically equal to ∞ on Ω.

X denotes as before the set of feasible points in the problem (P), i.e.

$$X = \{x \in \Omega \mid g_1(x) \le 0, \, \dots, \, g_p(x) \le 0, \, g_{p+1}(x) = 0, \, \dots, \, g_m(x) = 0\},$$

and $v_{\min}(P)$ is the optimal value of the problem.

Definition. Let

$$\Lambda = \mathbf{R}^p_+ \times \mathbf{R}^{m-p}.$$

The function $L\colon \Omega \times \Lambda \to \overline{\mathbf{R}}$, defined by

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x),$$

is called the *Lagrange function* of the minimization problem (P), and the variables $\lambda_1, \lambda_2, \dots, \lambda_m$ are called *Lagrange multipliers*.

Download free eBooks at bookboon.com

For each $x \in \operatorname{dom} f$, the expression $L(x, \lambda)$ is the sum of a real number and a linear form in $\lambda_1, \lambda_2, \dots, \lambda_m$. Hence, the function $\lambda \mapsto L(x, \lambda)$ is affine (or rather, the restriction to Λ of an affine function on $\mathbf{R}^m$). The Lagrange function is thus especially *concave in the variable* λ for each fixed $x \in \operatorname{dom} f$.

If $x \in \Omega \setminus \operatorname{dom} f$, then obviously $L(x, \lambda) = \infty$ for all $\lambda \in \Lambda$. Hence,

$$\inf_{x \in \Omega} L(x, \lambda) = \inf_{x \in \operatorname{dom} f} L(x, \lambda) < \infty$$

for all $\lambda \in \Lambda$.

Definition. For $\lambda \in \Lambda$, we define

$$\phi(\lambda) = \inf_{x \in \Omega} L(x, \lambda)$$

and call the function $\phi \colon \Lambda \to \underline{\mathbf{R}}$ the *dual function* associated to the minimization problem (P).

It may of course happen that the domain

$$\operatorname{dom} \phi = \{\lambda \in \Lambda \mid \phi(\lambda > -\infty\}$$

of the dual function is empty; this occurs if the functions $x \mapsto L(x, \lambda)$ are unbounded below on Ω for all $\lambda \in \Lambda$.

Theorem 10.1.1. *The dual function ϕ of the minimization problem (P) is concave and*

$$\phi(\lambda) \le v_{\min}(P)$$

for all $\lambda \in \Lambda$.

Hence, $\operatorname{dom} \phi = \emptyset$ if the objective function f in the original problem (P) is unbounded below on the constraint set, i.e. if $v_{\min}(P) = -\infty$.

Proof. The functions $\lambda \to L(x, \lambda)$ are concave for $x \in \operatorname{dom} f$, which means that the function ϕ is the infimum of a family of concave functions. It therefore follows from Theorem 6.2.4 in Part I that ϕ is concave.

Suppose $\lambda \in \Lambda$ and $x \in X$; then $\lambda_i g_i(x) \le 0$ for $i \le p$ and $\lambda_i g_i(x) = 0$ for $i > p$, and it follows that

$$L(x, \lambda) = f(x) + \sum_{i=1}^{n} \lambda_i g_i(x) \le f(x),$$

and that consequently

$$\phi(\lambda) = \inf_{x \in \Omega} L(x, \lambda) \le \inf_{x \in X} L(x, \lambda) \le \inf_{x \in X} f(x) = v_{\min}(P). \qquad \square$$

Download free eBooks at bookboon.com

The following optimality criterion is now an immediate consequence of the preceding theorem.

Theorem 10.1.2 (Optimality criterion). *Suppose $\hat{x}$ is a feasible point for the minimization problem (P) and that there is a point $\hat{\lambda} \in \Lambda$ such that*

$$\phi(\hat{\lambda}) = f(\hat{x}).$$

Then $\hat{x}$ is an optimal solution.

Proof. The common value $f(\hat{x})$ belongs to the intersection $\overline{\mathbf{R}} \cap \underline{\mathbf{R}} = \mathbf{R}$ of the codomains of f and ϕ, and it is thus a real number, and by Theorem 10.1.1, $f(\hat{x}) \leq v_{\min}(P)$. Hence, $f(\hat{x}) = v_{\min}(P)$. $\square$

EXAMPLE 10.1.1. Let us consider the simple minimization problem

$$\begin{array}{ll} \min & f(x) = x_1^2 - x_2^2 \\ \text{s.t.} & x_1^2 + x_2^2 \leq 1. \end{array}$$

The Lagrange function is

$$\begin{aligned} L(x_1, x_2, \lambda) &= x_1^2 - x_2^2 + \lambda(x_1^2 + x_2^2 - 1) \\ &= (\lambda + 1)x_1^2 + (\lambda - 1)x_2^2 - \lambda \end{aligned}$$

with $(x_1, x_2) \in \mathbf{R}^2$ and $\lambda \in \mathbf{R}_+$.

The Lagrange function is unbounded below when $0 \leq \lambda < 1$, and it attains the minimum value $-\lambda$ for $x_1 = x_2 = 0$ when $\lambda \geq 1$, so the dual function ϕ is given by

$$\phi(\lambda) = \begin{cases} -\infty\,, & \text{if } 0 \leq \lambda < 1 \\ -\lambda\,, & \text{if } \lambda \geq 1. \end{cases}$$

We finally note that the optimality condition $\phi(\hat{\lambda}) = f(\hat{x})$ is satisfied by the point $\hat{x} = (0, 1)$ and the Lagrange multiplier $\hat{\lambda} = 1$. Hence, $(0, 1)$ is an optimal solution. $\square$

The optimality criterion gives a sufficient condition for optimality, but it is not necessary, as the following trivial example shows.

EXAMPLE 10.1.2. Consider the problem

$$\begin{array}{ll} \min & f(x) = x \\ \text{s.t.} & x^2 \leq 0. \end{array}$$

Download free eBooks at bookboon.com

There is only one feasible point, $\hat{x} = 0$, which is therefore the optimal solution. The Lagrange function $L(x, \lambda) = x + \lambda x^2$ is bounded below for $\lambda > 0$ and

$$\phi(\lambda) = \inf_{x \in \mathbf{R}} (x + \lambda x^2) = \begin{cases} -1/4\lambda, & \text{if } \lambda > 0 \\ -\infty, & \text{if } \lambda = 0. \end{cases}$$

But $\phi(\lambda) < 0 = f(\hat{x})$ for all $\lambda \in \Lambda = \mathbf{R}_+$, so the optimality criterion in Theorem 10.1.2 is not satisfied by the optimal point. $\square$

For the converse of Theorem 10.1.2 to hold, some extra condition is thus needed, and we describe such a condition in Chapter 11.1.

The dual problem

In order to obtain the best possible lower estimate of the optimal value of the minimization problem (P), we should, in the light of Theorem 10.1.1, maximize the dual function. This leads to the following definition.

Download free eBooks at bookboon.com
Click on the ad to read more

Definition. The optimization problem

$$\text{(D)} \qquad \begin{array}{ll} \max & \phi(\lambda) \\ \text{s.t.} & \lambda \in \Lambda \end{array}$$

is called the *dual problem* of the minimization problem (P).

The dual problem is a convex problem, irrespective of whether the problem (P) is convex or not, because the dual function is concave. The value of the dual problem will be denoted by $v_{\max}(D)$ with the usual conventions for $\pm\infty$-values.

Our next result is now an immediate corollary of Theorem 10.1.1.

Theorem 10.1.3 (Weak duality). *The following inequality holds between the optimal values of the problem(P) and its dual problem (D):*

$$v_{\max}(D) \le v_{\min}(P).$$

The inequality in the above theorem is called *weak duality.* If the two optimal values are equal, i.e. if

$$v_{\max}(D) = v_{\min}(P)$$

then we say that *strong duality* holds for problem (P).

Weak duality thus holds for all problems while strong duality only holds for special types of problems. Of course, strong duality prevails if the optimality criterion in Theorem 10.1.2 is satisfied.

Example 10.1.3. Consider the minimization problem

$$\begin{array}{ll} \min & x_1^3 + 2x_2 \\ \text{s.t.} & x_1^2 + x_2^2 \le 1. \end{array}$$

It is easily verified that the minimum is attained for $x = (0, -1)$ and that the optimal value is $v_{\min}(P) = -2$. The Lagrange function

$$L(x_1, x_2, \lambda) = x_1^3 + 2x_2 + \lambda(x_1^2 + x_2^2 - 1) = x_1^3 + \lambda x_1^2 + 2x_2 + \lambda x_2^2 - \lambda$$

tends, for each fixed $\lambda \ge 0$, to $-\infty$ as $x_2 = 0$ and $x_1 \to -\infty$. The Lagrange function is in other words unbounded below on $\mathbf{R}^2$ for each λ, and hence $\phi(\lambda) = -\infty$ for all $\lambda \in \Lambda$. The value of the dual problem is therefore $v_{\max}(D) = -\infty$, so strong duality does not hold in this problem. $\square$

The Lagrange function, the dual function and the dual problem of a minimization problem of the type (P) are defined in terms of the constraint functions of the problem. Therefore, it may be worth emphasizing that

Download free eBooks at bookboon.com

problems that are equivalent in the sense that they have the same objective function f and the same set X of feasible points do not necessarily have equivalent dual problems. Thus, strong duality may hold for one way of framing a problem but fail to hold for other ways. See exercise 10.2.

Example 10.1.4. Let us find the dual problem of the LP problem

$$\text{(LP-P)} \qquad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & \begin{cases} Ax \geq b \\ x \geq 0. \end{cases} \end{array}$$

Here A is an $m \times n$-matrix, c is a vector in $\mathbf{R}^n$ and b a vector in $\mathbf{R}^m$. Let us rewrite the problem in the form

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & \begin{cases} b - Ax \leq 0 \\ x \in \mathbf{R}^n_+ \end{cases} \end{array}$$

with $x \in \mathbf{R}^n_+$ as an implicit constraint. The matrix inequality $b - Ax \leq 0$ consists of m linear inequalities, and the Lagrangefunction is therefore defined on the product set $\mathbf{R}^n_+ \times \mathbf{R}^m_+$, and it is given by

$$L(x, \lambda) = \langle c, x\rangle + \langle \lambda, b - Ax\rangle = \langle c - A^T\lambda, x\rangle + \langle b, \lambda\rangle.$$

For fixed λ, $L(x, \lambda)$ is bounded below on the set $\mathbf{R}^n_+$ if and only if $c - A^T\lambda \geq 0$, with minimum value equal to $\langle b, \lambda\rangle$ attained at $x = 0$. The dual function $\phi\colon \mathbf{R}^m_+ \to \underline{\mathbf{R}}$ is thus given by

$$\phi(\lambda) = \begin{cases} \langle b, \lambda\rangle\,, & \text{if } A^T\lambda \leq c \\ -\infty\,, & \text{otherwise.} \end{cases}$$

The dual problem to the LP problem (LP-P) is therefore also an LP problem, namely (after renaming the parameter λ to y) the LP problem

$$\text{(LP-D)} \qquad \begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & \begin{cases} A^Ty \leq c \\ y \geq 0. \end{cases} \end{array}$$

Note the beautiful symmetry between the two problems.

By weak duality, we know for sure that the optimal value of the maximization problem is less than or equal to the optimal value of the minimization problem. As we shall see later, strong duality holds for LP problems, i.e. the two problems above have the same optimal value, provided at least one of the problems has feasible points. □

Download free eBooks at bookboon.com

We now return to the general minimization problem

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & \begin{cases} g_i(x) \le 0, & i = 1, 2, \dots, p \\ g_i(x) = 0, & i = p+1, \dots, m \end{cases} \end{array}$$

with X as the set of feasible points, Lagrange function $L\colon \Omega \times \Lambda \to \overline{\mathbf{R}}$, and dual function ϕ. Our next theorem shows that the optimality criterion in Theorem 10.1.2 can be formulated as a saddle point condition on the Lagrange function.

Theorem 10.1.4. *Suppose* $(\hat{x}, \hat{\lambda}) \in \Omega \times \Lambda$. *The following three conditions are equivalent for the optimization problem (P):*

(i) $\hat{x} \in X$ *and* $f(\hat{x}) = \phi(\hat{\lambda})$, *i.e. the optimality criterion is satisfied.*

(ii) *For all* $(x, \lambda) \in \Omega \times \Lambda$,

$$L(\hat{x}, \lambda) \le L(\hat{x}, \hat{\lambda}) \le L(x, \hat{\lambda}),$$

i.e. $(\hat{x}, \hat{\lambda})$ *is a* saddle point *for the Lagrange function.*

(iii) $\hat{x} \in X$, $\hat{x}$ *minimizes the function* $x \mapsto L(x, \hat{\lambda})$ *when* x *runs through* Ω, *and*

$$\hat{\lambda}_i g_i(\hat{x}) = 0$$

for $i = 1, 2, \dots, p$.

Thus, $\hat{x}$ *is an optimal solution to the problem (P) if any of the equivalent conditions (i)–(iii) is satisfied.*

The condition in (iii) that $\hat{\lambda}_i g_i(\hat{x}) = 0$ for $i = 1, 2, \dots, p$ is called *complementarity.* An equivalent way to express this, which explains the name, is

$$\hat{\lambda}_i = 0 \quad \text{or} \quad g_i(\hat{x}) = 0.$$

A constraint with a positive Lagrange multiplier is thus necessarily active at the point $\hat{x}$.

Proof. (i) $\Rightarrow$ (ii): For $\hat{x} \in X$ and arbitrary $\lambda \in \Lambda$ $(= \mathbf{R}^p_+ \times \mathbf{R}^{n-p})$ we have

$$L(\hat{x}, \lambda) = f(\hat{x}) + \sum_{i=1}^{m} \lambda_i g_i(\hat{x}) = f(\hat{x}) + \sum_{i=1}^{p} \lambda_i g_i(\hat{x}) \le f(\hat{x}),$$

since $\lambda_i \ge 0$ and $g_i(\hat{x}) \le 0$ for $i = 1, 2, \dots, p$. Moreover,

$$\phi(\hat{\lambda}) = \inf_{z \in \Omega} L(z, \hat{\lambda}) \le L(x, \hat{\lambda}) \qquad \text{for all } x \in \Omega.$$

Download free eBooks at bookboon.com

If $f(\hat{x}) = \phi(\hat{\lambda})$, then consequently

$$L(\hat{x}, \lambda) \leq f(\hat{x}) = \phi(\hat{\lambda}) \leq L(x, \hat{\lambda})$$

for all $(x, \lambda) \in \Omega \times \Lambda$, and by the particular choice of $x = \hat{x}$, $\lambda = \hat{\lambda}$ in this inequality, we see that $f(\hat{x}) = L(\hat{x}, \hat{\lambda})$. This proves the saddle point inequality in (ii) with $L(\hat{x}, \hat{\lambda}) = f(\hat{x})$.

(ii) $\Rightarrow$ (iii): It is obvious that $\hat{x}$ minimizes the function $L(\,\cdot\,, \hat{\lambda})$ if and only if the right part of the saddle point inequality holds. The minimum value is moreover finite (due to our tacit assumption $\operatorname{dom} f \neq \emptyset$), and hence $f(\hat{x})$ is a finite number.

The left part of the saddlepoint inequality means that

$$f(\hat{x}) + \sum_{i=1}^{m} \lambda_i g_i(\hat{x}) \leq f(\hat{x}) + \sum_{i=1}^{m} \hat{\lambda}_i g_i(\hat{x})$$

for all $\lambda \in \Lambda$, or equivalently that

$$\sum_{i=1}^{m} (\lambda_i - \hat{\lambda}_i) g_i(\hat{x}) \leq 0$$

for all $\lambda \in \Lambda$.

Now fix the index k and choose in the above inequality the number λ so that $\lambda_i = \hat{\lambda}_i$ for all i except $i = k$. It follows that

$$(\lambda_k - \hat{\lambda}_k)g_k(\hat{x}) \leq 0 \tag{10.1}$$

for all such λ.

If $k > p$, we choose $\lambda_k = \hat{\lambda}_k \pm 1$ with the conclusion that $\pm g_k(\hat{x}) \leq 0$, i.e. that $g_k(\hat{x}) = 0$. For $k \leq p$ we instead choose $\lambda_k = \hat{\lambda}_k + 1$, with the conclusion that $g_k(\hat{x}) \leq 0$. Thus, $\hat{x}$ satisfies all the constraints, i.e. $\hat{x} \in X$.

For $k \leq p$ we finally choose $\lambda_k = 0$ and $\lambda_k = 2\hat{\lambda}_k$, respectively, in the inequality (10.1) with $\pm\hat{\lambda}_k g_k(\hat{x}) \leq 0$ as result. This means that $\hat{\lambda}_k g_k(\hat{x}) = 0$ for $k \leq p$, and the implication (ii) $\Rightarrow$ (iii) is now proved.

(iii) $\Rightarrow$ (i): From (iii) follows at once

$$\phi(\hat{\lambda}) = \inf_{x\in\Omega} L(x, \hat{\lambda}) = L(\hat{x}, \hat{\lambda}) = f(\hat{x}) + \sum_{i=1}^{m} \hat{\lambda}_i g_i(\hat{x}) = f(\hat{x}),$$

which is condition (i). □

If the objective and constraint functions f and $g_1, g_2, \dots, g_m$ are differentiable, so is the Lagrange function $L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x)$, and we use $L'_x(x_0, \lambda)$ as the notation for the value of the derivative of the function $x \mapsto L(x, \lambda)$ at the point x_0, i.e.

$$L'_x(x_0, \lambda) = f'(x_0) + \sum_{i=1}^{m} \lambda_i g'_i(x_0).$$

If the differentiable function $x \mapsto L(x, \lambda)$ has a minimum at an interior point x_0 in Ω, then $L'_x(x_0, \lambda) = 0$. The following corollary is thus an immediate consequence of the implication (i) $\Rightarrow$ (iii) in Theorem 10.1.4.

Corollary 10.1.5. *Suppose that $\hat{x}$ is an optimal solution to the minimization problem (P), that $\hat{x}$ is an interior point of the domain Ω, that the objective and constraint functions are differentiable at $\hat{x}$, and that the optimality criterion $f(\hat{x}) = \phi(\hat{\lambda})$ is satisfied by some Lagrange multiplier $\hat{\lambda} \in \Lambda$. Then*

$$\begin{cases} L'_x(\hat{x}, \hat{\lambda}) = 0 & \textit{and} \\ \hat{\lambda}_i g_i(\hat{x}) = 0 & \textit{for } i = 1, 2, \dots, p. \end{cases} \tag{KKT}$$

The system (KKT) is called the *Karush–Kuhn–Tucker condition.*

Download free eBooks at bookboon.com

The equality $L'_x(\hat{x}, \hat{\lambda}) = 0$ means that

$$f'(\hat{x}) + \sum_{i=1}^{m} \hat{\lambda}_i g'_i(\hat{x}) = 0,$$

which written out in more detail becomes

$$\begin{cases} \dfrac{\partial f}{\partial x_1}(\hat{x}) + \displaystyle\sum_{i=1}^{m} \hat{\lambda}_i \dfrac{\partial g_i}{\partial x_1}(\hat{x}) = 0 \\ \qquad\qquad \vdots \\ \dfrac{\partial f}{\partial x_n}(\hat{x}) + \displaystyle\sum_{i=1}^{m} \hat{\lambda}_i \dfrac{\partial g_i}{\partial x_n}(\hat{x}) = 0. \end{cases}$$

Example 10.1.5. In Example 10.1.1 we found that $\hat{x} = (0, 1)$ is an optimal solution to the minimization problem

$$\begin{array}{ll} \min & x_1^2 - x_2^2 \\ \text{s.t.} & x_1^2 + x_2^2 \le 1 \end{array}$$

and that the optimality criterion is satisfied with $\hat{\lambda} = 1$. The Lagrange function is $L(x, \lambda) = x_1^2 - x_2^2 + \lambda(x_1^2 + x_2^2 - 1)$, and indeed, $x = (0, 1)$ and $\lambda = 1$ satisfy the KKT-system

$$\begin{cases} \dfrac{\partial L(x,\lambda)}{\partial x_1} = 2(\lambda + 1)x_1 = 0 \\ \dfrac{\partial L(x,\lambda)}{\partial x_1} = 2(\lambda - 1)x_2 = 0 \\ \qquad \lambda(x_1^2 + x_2^2 - 1) = 0. \end{cases}$$

□

10.2 John's theorem

Conditions which guarantee that the KKT condition is satisfied at an optimal point, are usually called *constraint qualification conditions*, and in the next chapter we will describe such a condition for convex problems. In this section we will study a different qualifying condition, John's condition, for general optimization problems with constraints in the form of inequalities.

Let us therefore consider a problem of the form

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \le 0, \qquad i = 1, 2, \ldots, m \end{array}$$

Download free eBooks at bookboon.com

with implicit constraint set Ω, i.e. domain for the objective and the constraint functions.

Whether a constraint is active or not at an optimal point plays a major role, and affine constraints are thereby easier to handle than other constraints. Therefore, we introduce the following notations:

$$\begin{aligned} I_{\text{aff}}(x) &= \{i \mid \text{the function } g_i \text{ is affine and } g_i(x) = 0\}, \\ I_{\text{oth}}(x) &= \{i \mid \text{the function } g_i \text{ is not affine and } g_i(x) = 0\}, \\ I(x) &= I_{\text{aff}}(x) \cup I_{\text{oth}}(x). \end{aligned}$$

So $I_{\text{aff}}(x)$ consists of the indices of all active affine constraints at the point x, $I_{\text{oth}}(x)$ consists of the indices of all other active constraints at the point, and $I(x)$ consists of the indices of all active constraints at the point.

Theorem 10.2.1 (John's theorem)**.** *Suppose $\hat{x}$ is a local minimum point for the problem (P), that $\hat{x}$ is an interior point in Ω, and that the functions f and $g_1, g_2, \dots, g_m$ are differentiable at the point $\hat{x}$. If there exists a vector $z \in \mathbf{R}^n$ such that*

$$\text{(J)} \qquad \begin{cases} \langle g_i'(\hat{x}), z\rangle \geq 0 & \text{for all } i \in I_{\text{aff}}(\hat{x}) \\ \langle g_i'(\hat{x}), z\rangle > 0 & \text{for all } i \in I_{\text{oth}}(\hat{x}), \end{cases}$$

Download free eBooks at bookboon.com
Click on the ad to read more

then there exist Lagrange parameters $\hat{\lambda} \in \mathbf{R}^m_+$ *such that*

$$\text{(KKT)} \qquad \begin{cases} L'_x(\hat{x}, \hat{\lambda}) = 0 \\ \hat{\lambda}_i g_i(\hat{x}) = 0 \qquad \textit{for } i = 1, 2, \dots, m. \end{cases}$$

Remark 1. According to Theorem 3.3.5 in Part I, the system (J) is solvable if and only if

$$(\mathrm{J}') \qquad \begin{cases} \displaystyle\sum_{i \in I(\hat{x})} u_i g'_i(\hat{x}) = 0 \\ \qquad\qquad u \geq 0 \end{cases} \Rightarrow \quad u_i = 0 \quad \text{for all } i \in I_{\text{oth}}(\hat{x}).$$

The system (J) is thus in particular solvable if the gradient vectors $\nabla g_i(\hat{x})$ are linearly independent for $i \in I(\hat{x})$.

Remark 2. If $I_{\text{oth}}(\hat{x}) = \emptyset$, then (J) is trivially satisfied by $z = 0$.

Proof. Let Z denote the set of solutions to the system (J). The first part of the proof consists in showing that Z is a subset of the conic halfspace $\{z \in \mathbf{R}^n \mid -\langle f'(\hat{x}), z\rangle \geq 0\}$.

Assume therefore that $z \in Z$ and consider the halfline $\hat{x} - tz$ for $t \geq 0$. We claim that $\hat{x} - tz \in X$ for all sufficiently small $t > 0$.

If g is an affine function, i.e. has the form $g(x) = \langle c, x\rangle + b$, then $g'(x) = c$ and $g(x + y) = \langle c, x + y\rangle + b = \langle c, x\rangle + b + \langle c, y\rangle = g(x) + \langle g'(x), y\rangle$ for all x and y. Hence, for all indices $i \in I_{\text{aff}}(\hat{x})$,

$$g_i(\hat{x} - tz) = g_i(\hat{x}) - t\langle g'_i(\hat{x}), z\rangle = -t\langle g'_i(\hat{x}), z\rangle \leq 0$$

for all $t \geq 0$.

For indices $i \in I_{\text{oth}}(\hat{x})$, we obtain instead, using the chain rule, the inequality

$$\frac{d}{dt}\, g_i(\hat{x} - tz)|_{t=0} = -\langle g'_i(\hat{x}), z\rangle < 0.$$

The function $t \mapsto g_i(\hat{x} - tz)$ is in other words decreasing at the point $t = 0$, whence $g_i(\hat{x} - tz) < g_i(\hat{x}) = 0$ for all sufficiently small $t > 0$.

If the i:th constraint is inactive at $\hat{x}$, i.e. if $i \notin I(\hat{x})$, then $g_i(\hat{x}) < 0$, and it follows from continuity that $g_i(\hat{x} - tz) < 0$ for all sufficiently small $t > 0$.

We have thus proved that the points $\hat{x} - tz$ belong to the constraint set X if $t > 0$ is sufficiently small. Since $\hat{x}$ is a local minimum point of f, it follows that $f(\hat{x} - tz) \geq f(\hat{x})$ for all sufficiently small $t > 0$. Consequently,

$$-\langle f'(\hat{x}), z\rangle = \frac{d}{dt}\, f(\hat{x} - tz)\Big|_{t=0} = \lim_{t \to 0^+} \frac{f(\hat{x} - tz) - f(\hat{x})}{t} \geq 0.$$

Download free eBooks at bookboon.com

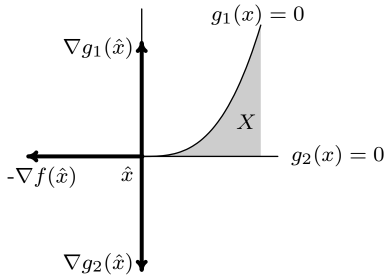


Figure 10.1. Illustration for Example 10.2.1: The vector $-\nabla f(\hat{x})$ does not belong to the cone generated by the gradients $\nabla g_1(\hat{x})$ and $\nabla g_2(\hat{x})$.

This proves the alleged inclusion

$$Z \subseteq \{z \in \mathbf{R}^n \mid -\langle f'(\hat{x}), z\rangle \geq 0\} = \{-f'(\hat{x})\}^+ = \big(\operatorname{con}\{-f'(\hat{x})\}\big)^+,$$

and it now follows from Theorem 3.2.1, Corollary 3.2.4 and Theorem 3.3.4 in Part I that

$$\operatorname{con}\{-f'(\hat{x})\} \subseteq Z^+ = \operatorname{con}\{g_i'(\hat{x}) \mid i \in I(\hat{x})\}.$$

So the vector $-f'(\hat{x})$ belongs to the cone generated by the vektors $g_i'(\hat{x})$, $i \in I(\hat{x})$, which means that there are nonnegative integers $\hat{\lambda}_i$, $i \in I(\hat{x})$, such that

$$-f'(\hat{x}) = \sum_{i \in I(\hat{x})} \hat{\lambda}_i g_i'(\hat{x}).$$

If we finally define $\hat{\lambda}_i = 0$ for $i \notin I(\hat{x})$, then

$$f'(\hat{x}) + \sum_{i=1}^{m} \hat{\lambda}_i g_i'(\hat{x}) = 0$$

and $\hat{\lambda}_i g_i(\hat{x}) = 0$ for $i = 1, 2, \ldots, m$. This means that the KKT-condition is satisfied. □

The condition in John's statement that the system (J) has a solution can be replaced with other qualifying constraints but can not be completely removed without the conclusion being lost. This is shown by the following example.

Example 10.2.1. Consider the problem

$$\begin{array}{ll} \min & f(x) = x_1 \\ \text{s.t.} & \begin{cases} g_1(x) = -x_1^3 + x_2 \leq 0 \\ g_2(x) = -x_2 \leq 0 \end{cases} \end{array}$$

Download free eBooks at bookboon.com

with Lagrange function $L(x,\lambda) = x_1 + \lambda_1(x_2 - x_1^3) - \lambda_2 x_2$. The unique optimal solution is $\hat{x} = (0,0)$, but the system $L'_x(\hat{x},\lambda) = 0$, i.e.

$$\begin{cases} \quad\quad 1 = 0 \\ \lambda_1 - \lambda_2 = 0, \end{cases}$$

has no solutions. This is explained by the fact that the system (J), i.e.

$$\begin{cases} -z_2 \geq 0 \\ \quad z_2 > 0, \end{cases}$$

has no solutions. □

EXAMPLE 10.2.2. We will solve the problem

$$\begin{array}{ll} \min & x_1 x_2 + x_3 \\ \text{s.t.} & \begin{cases} 2x_1 - 2x_2 + x_3 + 1 \leq 0 \\ \ x_1^2 + \ x_2^2 - x_3 \quad\ \ \leq 0 \end{cases} \end{array}$$

using John's theorem. Note first that the constraints define a compact set X, for the inequalities

$$x_1^2 + x_2^2 \leq x_3 \leq -2x_1 + 2x_2 - 1$$

Download free eBooks at bookboon.com

Click on the ad to read more

imply that $(x_1+1)^2+(x_2-1)^2 \leq 1$, and consequently, $-2 \leq x_1 \leq 0$, $0 \leq x_2 \leq 2$, and $0 \leq x_3 \leq 7$. Since the objective function is continuous, there is indeed an optimal solution.

Let us now first investigate whether the system (J) is solvable. We use the equivalent version (J′) in the remark after the theorem. First note that the gradients of the constraint functions are never equal to zero. The condition (J′) is thus met in the points where only one of the constraints is active.

Assume therefore that x is a point where $I(x)=\{1,2\}$, i.e. where both constraints are active, and that $u_1(2,-2,1)+u_2(2x_1,2x_2,-1)=(0,0,0)$. If $u_2>0$, we conclude from the above equation that $u_1=u_2$, $x_1=-1$ and $x_2=1$. Inserting $x_1=-1$ and $x_2=1$ into the two active constraints yields $x_3=3$ and $x_3=2$, respectively, which is contradictory. Thus, $u_2=0$, which means that the condition (J′) is fulfilled at all feasible points.

We conclude that the optimal point satisfies the KKT-condition, which in this instance is as follows

$$\begin{cases} \begin{aligned} x_2+2\lambda_1+2x_1\lambda_2 &= 0 & \text{(i)}\\ x_1-2\lambda_1+2x_2\lambda_2 &= 0 & \text{(ii)}\\ 1+\lambda_1-\lambda_2 &= 0 & \text{(iii)}\\ \lambda_1(2x_1-2x_2+x_3+1) &= 0 & \text{(iv)}\\ \lambda_2(x_1^2+x_2^2-x_3) &= 0 & \text{(v)} \end{aligned} \end{cases}$$

The further investigation is divided into two cases.

$\lambda_1=0:$ Equation (iii) implies that $\lambda_2=1$, which inserted into (i) and (ii) gives $x_1=x_2=0$, and from (v) now follows $x_3=0$. But this is a false solution, since $(0,0,0)\notin X$.

$\lambda_1>0:$ Equation (iv) now implies that

$$2x_1-2x_2+x_3+1=0. \qquad \text{(vi)}$$

From (i) and (ii) follows $(x_1+x_2)(1+2\lambda_2)=0$, and since $\lambda_2\geq 0$,

$$x_1+x_2=0. \qquad \text{(vii)}$$

By (iii,) $\lambda_2>0$. Condition (v) therefore implies that

$$x_1^2+x_2^2-x_3=0. \qquad \text{(viii)}$$

The system consisting of equations (vi), (vii), (viii) has two solutions, namely

$$\hat{x}=(-1+\sqrt{1/2},1-\sqrt{1/2},3-2\sqrt{2}) \text{ and } \overline{x}=(-1-\sqrt{1/2},1+\sqrt{1/2},3+2\sqrt{2}).$$

Using (i) and (iii), we compute the corresponding λ and obtain

$$\hat{\lambda}=(-1/2+\sqrt{1/2},1/2+\sqrt{1/2}) \text{ and } \overline{\lambda}=(-1/2-\sqrt{1/2},1/2-\sqrt{1/2}),$$

Download free eBooks at bookboon.com

respectively. Note that $\hat{\lambda} \geq 0$ and $\overline{\lambda} < 0$. The system KKT thus has a unique solution (x, λ) with $\lambda \geq 0$, namely $x = \hat{x}$, $\lambda = \hat{\lambda}$. By John's theorem, $\hat{x}$ is the unique optimal solution of our minimization problem, and the optimal value is $3/2 - \sqrt{2}$. □

Exercises

10.1 Determine the dual function for the optimization problem

$$\begin{array}{ll} \min & x_1^2 + x_2^2 \\ \text{s.t.} & x_1 + x_2 \geq 2, \end{array}$$

and prove that $(1, 1)$ is an optimal solution by showing that the optimality criterion is satisfied by $\hat{\lambda} = 2$. Also show that the KKT-condition is satisfied at the optimal point.

10.2 Consider the two minimization problems

$$(P_a) \quad \begin{array}{c} \min e^{-x_1} \\ x_1^2/x_2 \leq 0 \end{array} \qquad \text{and} \qquad (P_b) \quad \begin{array}{c} \min e^{-x_1} \\ |x_1| \leq 0 \end{array}$$

both with $\Omega = \{(x_1, x_2) \mid x_2 > 0\}$ as implicit domain. The two problems have the same set $X = \{(0, x_2) \mid x_2 > 0\}$ of feasible points and the same optimal value $v_{\text{min}} = 1$. Find their dual functions and dual problems, and show that strong duality holds for (P_b) but not for (P_a).

10.3 Suppose the function $f\colon X \times Y \to \mathbf{R}$ has two saddle points $(\hat{x}_1, \hat{y}_1)$ and $(\hat{x}_2, \hat{y}_2)$. Prove that

a) $f(\hat{x}_1, \hat{y}_1) = f(\hat{x}_2, \hat{y}_2)$;

b) $(\hat{x}_1, \hat{y}_2)$ and $(\hat{x}_2, \hat{y}_1)$ are saddle points, too.

10.4 Let $f\colon X \times Y \to \mathbf{R}$ be an arbitrary function.

a) Prove that

$$\sup_{y \in Y} \inf_{x \in X} f(x, y) \leq \inf_{x \in X} \sup_{y \in Y} f(x, y).$$

b) Suppose there is a point $(\hat{x}, \hat{y}) \in X \times Y$ such that

$$\sup_{y \in Y} \inf_{x \in X} f(x, y) = \inf_{x \in X} f(x, \hat{y}) \quad \text{and} \quad \inf_{x \in X} \sup_{y \in Y} f(x, y) = \sup_{y \in Y} f(\hat{x}, y).$$

Prove that $(\hat{x}, \hat{y})$ is a saddle point of the function f if and only if

$$\inf_{x \in X} f(x, \hat{y}) = \sup_{y \in Y} f(\hat{x}, y),$$

and that the common value then is equal to $f(\hat{x}, \hat{y})$.

Download free eBooks at bookboon.com

10.5 Consider a minimization problem

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \le 0, \qquad i = 1, 2, \dots, m \end{array}$$

with *convex* differentiable constraint functions $g_1, g_2, \dots, g_m$, and suppose there is a point $x_0 \in X = \{x \mid g_1(x) \le 0, \dots, g_m(x) \le 0\}$ which satisfies all non-affine constraints with strict inequality. Show that the system (J) is solvable at all points $\hat{x} \in X$.
[Hint: Show that $z = \hat{x} - x_0$ satisfies (J).]

10.6 Solve the following optimization problems

a) $$\begin{array}{ll} \min & x_1^3 + x_1 x_2^2 \\ \text{s.t.} & \begin{cases} x_1^2 + 2x_2^2 \le 1 \\ x_2 \ge 0 \end{cases} \end{array}$$

b) $$\begin{array}{ll} \max & x_1^2 + x_2^2 + \arctan x_1 x_2 \\ \text{s.t.} & \begin{cases} x_1^2 + x_2^2 \le 2 \\ 0 \le x_1 \le x_2 \end{cases} \end{array}$$

c) $$\begin{array}{ll} \min & x_1 x_2 \\ \text{s.t.} & \begin{cases} x_1^2 + x_1 x_2 + 4x_2^2 \le 1 \\ x_1 + 2x_2 \ge 0 \end{cases} \end{array}$$

d) $$\begin{array}{ll} \max & x_1^2 x_2 x_3 \\ \text{s.t.} & \begin{cases} 2x_1 + x_1 x_2 + x_3 \le 1 \\ x_1,\ x_2,\ x_3 \ge 0. \end{cases} \end{array}$$

Need help with your dissertation?

Get in-depth feedback & advice from experts in your topic area. Find out what you can do to improve the quality of your dissertation!

Get Help Now

Go to www.helpmyassignment.co.uk for more info

Helpmyassignment

Download free eBooks at bookboon.com

Click on the ad to read more

Chapter 11

Convex optimization

11.1 Strong duality

We recall that the minimization problem

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & \begin{cases} g_i(x) \le 0, & i = 1, 2, \ldots, p \\ g_i(x) = 0, & i = p+1, \ldots, m \end{cases} \end{array}$$

is called convex if

- the implicit constraint set Ω is convex,
- the objective function f is convex,
- the constraint functions g_i are convex for $i = 1, 2, \ldots, p$ and affine for $i = p+1, \ldots, m$.

The set X of feasible points is convex in a convex optimization problem, and the Lagrange function

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$$

is convex in the variable x for each fixed $\lambda \in \Lambda = \mathbf{R}_+^p \times \mathbf{R}^{m-p}$, since it is a conic combination of convex functions.

We have already noted that the optimality criterion in Theorem 10.1.2 need not be fulfilled at an optimal point, not even for convex problems, because of the trivial counterexample in Example 10.1.2. For the criterion to be met some additional condition is needed, and a weak one is given in the next definition.

Definition. The problem (P) satisfies *Slater's condition* if there is a feasible point $\overline{x}$ in the relative interior of Ω such that $g_i(\overline{x}) < 0$ for each *non-affine* constraint function g_i.

Download free eBooks at bookboon.com

Slater's condition is of course vacously fulfilled if all constraint functions are affine.

For convex problems that satisfy Slater's condition, the optimality criterion is both sufficient and necessary for optimality. We have namely the following result.

Theorem 11.1.1 (Duality theorem). *Suppose that the problem (P) is convex and satisfies Slater's condition, and that the optimal value $v_{\min}$ is finite. Let $\phi\colon \Lambda \to \underline{\mathbf{R}}$ denote the dual function of the problem. Then there is a point $\hat{\lambda} \in \Lambda$ such that*

$$\phi(\hat{\lambda}) = v_{\min}.$$

Proof. First suppose that all constraints are inequalities, i.e. that $p = m$, and renumber the constraints so that the functions g_i are convex and non-affine for $i = 1, 2, \dots, k$ and affine for $i = k+1, \dots, m$.

Because of Slater's condition, the system

$$\begin{cases} g_i(x) < 0, & i = 1, 2, \dots, k \\ g_i(x) \le 0, & i = k+1, \dots, m \end{cases}$$

has a solution in the relative interior of Ω, whereas the system

$$\begin{cases} f(x) - v_{\min} < 0 & \\ \qquad g_i(x) < 0, & i = 1, 2, \dots, k \\ \qquad g_i(x) \le 0, & i = k+1, \dots, m \end{cases}$$

lacks solutions in Ω, due to the definition of $v_{\min}$. Therefore, it follows from Theorem 6.5.1 in Part I that there exist nonnegative scalars $\hat{\lambda}_0, \hat{\lambda}_1, \dots, \hat{\lambda}_m$ such that at least one of the numbers $\hat{\lambda}_0, \hat{\lambda}_1, \dots, \hat{\lambda}_k$ is positive and

$$\hat{\lambda}_0(f(x) - v_{\min}) + \hat{\lambda}_1 g_1(x) + \hat{\lambda}_2 g_2(x) + \dots + \hat{\lambda}_m g_m(x) \ge 0$$

for all $x \in \Omega$. Here, the coefficient $\hat{\lambda}_0$ has to be positive, because if $\hat{\lambda}_0 = 0$ then $\hat{\lambda}_1 g_1(x) + \dots + \hat{\lambda}_m g_m(x) \ge 0$ for all $x \in \Omega$, which contradicts the fact that the first mentioned system of inequalities has a solution in Ω. We may therefore assume, by dividing by $\hat{\lambda}_0$ if necessary, that $\hat{\lambda}_0 = 1$, and this gives us the inequality

$$L(x, \hat{\lambda}) = f(x) + \sum_{i=1}^{m} \hat{\lambda}_i g_i(x) \ge v_{\min}$$

for all $x \in \Omega$. It follows that

$$\phi(\hat{\lambda}) = \inf_{x \in \Omega} L(x, \hat{\lambda}) \ge v_{\min},$$

which combined with Theorem 10.1.1 yields the desired equality $\phi(\hat{\lambda}) = v_{\min}$.

Download free eBooks at bookboon.com

If the problem has affine equality constraints, i.e. if $p < m$, we replace each equality $g_i(x) = 0$ with the two inequalities $\pm g_i(x) \leq 0$, and it follows from the already proven case of the theorem that there exist nonnegative Lagrange multipliers $\hat{\lambda}_1, \ldots, \hat{\lambda}_p, \hat{\mu}_{p+1}, \ldots, \hat{\mu}_m, \hat{\nu}_{p+1}, \ldots, \hat{\nu}_m$ such that

$$f(x) + \sum_{i=1}^{p} \hat{\lambda}_i g_i(x) + \sum_{i=p+1}^{m} (\hat{\mu}_i - \hat{\nu}_i) g_i(x) \geq v_{\min}$$

for all $x \in \Omega$, By defining $\hat{\lambda}_i = \hat{\mu}_i - \hat{\nu}_i$ for $i = p+1, \ldots, m$, we obtain a point $\hat{\lambda} \in \Lambda = \mathbf{R}_+^p \times \mathbf{R}^{m-p}$ which satisfies $\phi(\hat{\lambda}) \geq v_{\min}$, and this completes the proof of the theorem. □

By combining Theorem 11.1.1 with Theorem 10.1.2 we get the following corollary.

Corollary 11.1.2. *Suppose that the problem (P) is convex and that it satisfies Slater's condition. Then, a feasible point $\hat{x}$ is optimal if and only if it satisfies the optimality criterion, i.e. if and only if there exists a $\hat{\lambda} \in \Lambda$ such that $\phi(\hat{\lambda}) = f(\hat{x})$.*

Click on the ad to read more

11.2 The Karush–Kuhn–Tucker theorem

Variants of the following theorem were first proved by Karush and Kuhn–Tucker, and the theorem is therefore usually called the Karush–Kuhn–Tucker theorem.

Theorem 11.2.1. *Let*

$$(P) \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & \begin{cases} g_i(x) \le 0, & i = 1, 2, \dots, p \\ g_i(x) = 0, & i = p+1, \dots, m \end{cases} \end{array}$$

be a convex problem, and suppose that the objective and constraint functions are differentiable at the feasible point $\hat{x}$.

(i) If $\hat{\lambda}$ is a point in Λ and the pair $(\hat{x}, \hat{\lambda})$ satisfies the KKT-condition

$$\begin{cases} L'_x(\hat{x}, \hat{\lambda}) = 0 \\ \hat{\lambda}_i g_i(\hat{x}) = 0 \qquad \textit{for } i = 1, 2, \dots, p \end{cases}$$

then strong duality prevails; $\hat{x}$ is an optimal solution to the problem (P) and $\hat{\lambda}$ is an optimal solution to the dual problem.

(ii) Conversely, if Slater's condition is fulfilled and $\hat{x}$ is an optimal solution, then there exist Lagrange multipliers $\hat{\lambda} \in \Lambda$ such that $(\hat{x}, \hat{\lambda})$ satisfies the KKT-condition.

Proof. (i) The KKT-condition implies that $\hat{x}$ is a stationary point of the convex function $x \mapsto L(x, \hat{\lambda})$, and an interior stationary point of a convex function is a minimum point, according to Theorem 7.2.2 in Part I. Condition (iii) in Theorem 10.1.4 is thus fulfilled, and this means that the optimality criterion is satisfied by the pair $(\hat{x}, \hat{\lambda})$.

(ii) Conversely, if Slater's condition is satisfied and $\hat{x}$ is an optimal solution, then the optimality criterion $f(\hat{x}) = \phi(\hat{\lambda})$ is satisfied by some $\hat{\lambda} \in \Lambda$, according to Theorem 11.1.1. The KKT-condition is therefore met because of Corollary 10.1.5. □

The KKT-condition has a natural geometrical interpretation. Assume for simplicity that all constraints are inequalities, i.e. that $p = m$, and let $I(\hat{x})$ denote the index set for the constraints that are active at the optimal point $\hat{x}$. The KKT-condition means that $\hat{\lambda}_i = 0$ for all indices $i \notin I(\hat{x})$ and that

$$-\nabla f(\hat{x}) = \sum_{i \in I(\hat{x})} \hat{\lambda}_i \nabla g_i(\hat{x}),$$

Download free eBooks at bookboon.com

where all coefficients $\hat{\lambda}_i$ occuring in the sum are nonnegative. The geometrical meaning of the above equality is that the vector $-\nabla f(\hat{x})$ belongs to the cone generated by the gradients $\nabla g_i(\hat{x})$ of the active inequality constraints. Cf. figure 11.1 and figure 11.2.

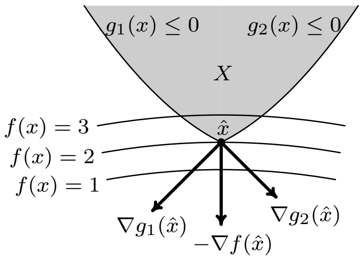


Figure 11.1. The point $\hat{x}$ is optimal since both constraints are active at the point and $-\nabla f(\hat{x}) \in \operatorname{con}\{\nabla g_1(\hat{x}), \nabla g_2(\hat{x})\}$.

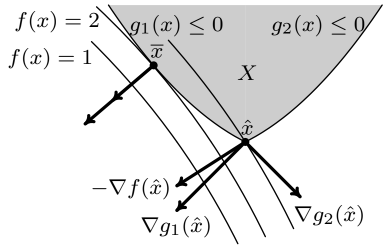


Figure 11.2. Here the point $\hat{x}$ is not optimal since $-\nabla f(\hat{x}) \notin \operatorname{con}\{\nabla g_1(\hat{x}), \nabla g_2(\hat{x})\}$. The optimum is instead attained at $\overline{x}$, where $-\nabla f(\overline{x}) = \lambda_1 \nabla g_1(\overline{x})$ for some $\lambda_1 > 0$.

Example 11.2.1. Consider the problem

$$\min \mathrm{e}^{x_1 - x_3} + \mathrm{e}^{-x_2}$$
$$\begin{cases} (x_1 - x_2)^2 - x_3 \leq 0 \\ x_3 - 4 \leq 0. \end{cases}$$

The objective and the constraint functions are convex. Slater's condition is satisfied, since for instance $(1, 1, 1)$ satisfies both constraints strictly. According to Theorem 11.2.1, x is therefore an optimal solution to the problem if and only if x solves the system

$$\begin{cases} \mathrm{e}^{x_1 - x_3} + 2\lambda_1(x_1 - x_2) = 0 & \text{(i)} \\ -\mathrm{e}^{-x_2} - 2\lambda_1(x_1 - x_2) = 0 & \text{(ii)} \\ -\mathrm{e}^{x_1 - x_3} - \lambda_1 + \lambda_2 = 0 & \text{(iii)} \\ \lambda_1\big((x_1 - x_2)^2 - x_3\big) = 0 & \text{(iv)} \\ \lambda_2(x_3 - 4) = 0 & \text{(v)} \\ \lambda_1, \lambda_2 \geq 0 & \text{(vi)} \end{cases}$$

It follows from (i) and (vi) that $\lambda_1 > 0$, from (iii) and (vi) that $\lambda_2 > 0$, and from (iv) and (v) that $x_3 = 4$ and $x_1 - x_2 = \pm 2$. But $x_1 - x_2 < 0$, because of (i) and (vi), and hence $x_1 - x_2 = -2$. By comparing (i) and (ii) we see that $x_1 - x_3 = -x_2$, i.e. $x_1 + x_2 = 4$. It follows that $x = (1, 3, 4)$

Download free eBooks at bookboon.com

and $\lambda = (e^{-3}/4, 5e^{-3}/4)$ is the unique solution of the system. The problem therefore has a unique optimal solution, namely $(1, 3, 4)$. The optimal value is equal to $2e^{-3}$. $\square$

11.3 The Lagrange multipliers

In this section we will study how the optimal value $v_{\min}(b)$ of an arbitrary minimization problem of the type

$$(\mathrm{P}_b) \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & \begin{cases} g_i(x) \le b_i, & i = 1, 2, \dots, p \\ g_i(x) = b_i, & i = p+1, \dots, m \end{cases} \end{array}$$

depends on the constraint parameters $b_1, b_2, \dots, b_m$. The functions f and $g_1, g_2, \dots, g_m$ are, as previously, defined on a subset Ω of $\mathbf{R}^n$, $b = (b_1, \dots, b_m)$ is a vector in $\mathbf{R}^m$, and

$$X(b) = \{x \in \Omega \mid g_i(x) \le b_i \text{ for } 1 \le i \le p \text{ and } g_i(x) = b_i \text{ for } p < i \le m\}$$

is the set of feasible points.

Download free eBooks at bookboon.com

Click on the ad to read more

The Lagrange function and the dual function associated to the minimization problem (P_b) are denoted by L_b and ϕ_b, respectively. By definition,

$$L_b(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i(g_i(x) - b_i),$$

and the relationship between the Lagrange functions L_b and $L_{\overline{b}}$ belonging to two different parameter vectors b and $\overline{b}$, is therefore given by the equation

$$L_b(x,\lambda) = L_{\overline{b}}(x,\lambda) + \sum_{i=1}^{m} \lambda_i(\overline{b}_i - b_i) = L_{\overline{b}}(x,\lambda) + \langle \lambda, \overline{b} - b\rangle.$$

By forming the infimum over $x \in \Omega$, we immediately get the following relation for the dual functions:

$$\phi_b(\lambda) = \phi_{\overline{b}}(\lambda) + \langle \lambda, \overline{b} - b\rangle. \tag{11.1}$$

The following theorem gives an interpretation of the Lagrange parameters in problems which satisfy the optimality criterion in Theorem 10.1.2, and thus especially for convex problems which satisfy Slater's condition.

Theorem 11.3.1. *Suppose that the minimization problem $(P_{\overline{b}})$ has an optimal solution $\overline{x}$ and that the optimality criterion is satisfied at the point, i.e. that there are Lagrange multipliers $\overline{\lambda}$ such that $\phi_{\overline{b}}(\overline{\lambda}) = f(\overline{x})$. Then:*

- *(i) The objective function f is bounded below on $X(b)$ for each $b \in \mathbf{R}^m$, so the optimal value $v_{\min}(b)$ of problem (P_b) is finite if the set $X(b)$ of feasible points is nonempty, and equal to $+\infty$ if $X(b) = \emptyset$.*
- *(ii) The vector $-\overline{\lambda}$ is a subgradient at the point $\overline{b}$ of the optimal value function $v_{\min}\colon \mathbf{R}^m \to \overline{\mathbf{R}}$.*
- *(iii) Suppose that the optimality criterion is satisfied in the problem (P_b) for all b in an open convex set U. The restriction of the function $v_{\min}$ to U is then a convex function.*

Proof. By using weak duality for problem (P_b), the identity (11.1) and the optimality criterion for problem $(P_{\overline{b}})$, we obtain the following inequality:

$$\begin{aligned} v_{\min}(b) &= \inf_{x\in X(b)} f(x) \ge \phi_b(\overline{\lambda}) = \phi_{\overline{b}}(\overline{\lambda}) + \langle \overline{\lambda}, \overline{b} - b\rangle = f(\overline{x}) + \langle \overline{\lambda}, \overline{b} - b\rangle \\ &= v_{\min}(\overline{b}) - \langle \overline{\lambda}, b - \overline{b}\rangle. \end{aligned}$$

It follows, first, that the optimal value $v_{\min}(b)$ can not be equal to $-\infty$, and second, that $-\overline{\lambda}$ is a subgradient of the function $v_{\min}$ at the point $\overline{b}$.

If the optimality criterion is satisfied at all $b \in U$, then $v_{\min}$ has a subgradient at all points in U, and such a function is convex. □

Download free eBooks at bookboon.com

Now suppose that the function $v_{\min}$ is differentiable at the point $\bar{b}$. The gradient at the point $\bar{b}$ is then, by Theorem 8.1.3 in Part I, the unique subgradient at the point, so it follows from (ii) in the above theorem that $v'_{\min}(\bar{b}) = -\bar{\lambda}$. This gives us the approximation

$$v_{\min}(\bar{b}_1 + \Delta b_1, \dots, \bar{b}_m + \Delta b_m) \approx v_{\min}(\bar{b}_1, \dots, \bar{b}_m) - \bar{\lambda}_1 \Delta b_1 \cdots - \bar{\lambda}_m \Delta b_m$$

for small increments Δb_j. So the Lagrange multipliers provide information about how the optimal value is affected by small changes in the parameters.

Example 11.3.1. As an illustration of Theorem 11.3.1, let us study the convex problem

$$\begin{aligned} &\min \;\; x_1^2 + x_2^2 \\ &\text{s.t.} \begin{cases} \;\; x_1 + 2x_2 \le b_1 \\ 2x_1 + \;\; x_2 \le b_2. \end{cases} \end{aligned}$$

Since it is about minimizing the distance squared from the origin to a polyhedron, there is certainly an optimal solution for each right-hand side b, and since the constraints are affine, it follows from the Karush–Kuhn–Tucker theorem that the optimal solution satisfies the KKT-condition, which in the present case is the system

$$\begin{cases} 2x_1 + \;\; \lambda_1 + 2\lambda_2 = 0 & \text{(i)} \\ 2x_2 + 2\lambda_1 + \;\; \lambda_2 = 0 & \text{(ii)} \\ \lambda_1(x_1 + 2x_2 - b_1) = 0 & \text{(iii)} \\ \lambda_2(2x_1 + x_2 - b_2) = 0 & \text{(iv)} \\ \lambda_1, \, \lambda_2 \ge 0. \end{cases}$$

We now solve this system by considering four separate cases:

$\lambda_1 = \lambda_2 = 0$: In this case, $x_1 = x_2 = 0$ is the unique solution to the KKT-system. Thus, the point $(0,0)$ is optimal provided it is feasible, and so is the case if and only if $b_1 \ge 0$ and $b_2 \ge 0$. The optimal value for these parameter values is $v_{\min}(b) = 0$.

$\lambda_1 > 0$, $\lambda_2 = 0$: From (i) and (ii), it follows first that $x_2 = 2x_1 = -\lambda_1$, and (iii) then gives $x = \frac{1}{5}(b_1, 2b_1)$. This point is feasible if $2x_1 + x_2 = \frac{4}{5}b_1 \le b_2$, and for the Lagrange multiplier $\lambda_1 = -\frac{2}{5}b_1$ to be positive, we must also have $b_1 < 0$. Thus, the point $x = \frac{1}{5}(b_1, 2b_1)$ is optimal if $b_1 < 0$ and $4b_1 \le 5b_2$, and the corresponding value is $v_{\min}(b) = \frac{1}{5}b_1^2$.

$\lambda_1 = 0$, $\lambda_2 > 0$: From (i) and (ii), it now follows that $x_1 = 2x_2 = -\lambda_2$, which inserted into (iv) gives $x = \frac{1}{5}(2b_2, b_2)$. This is a feasible point if $x_1 + 2x_2 = \frac{4}{5}b_2 \le b_1$. The Lagrange multiplier $\lambda_2 = -\frac{2}{5}b_2$ is positive if $b_2 < 0$.

Download free eBooks at bookboon.com

Hence, the point $x = \frac{1}{5}(2b_2, b_2)$ is optimal and the optimal value is $v(b) = \frac{1}{5}b_2^2$, if $b_2 < 0$ och $4b_2 \leq 5b_1$.

$\lambda_1 > 0$, $\lambda_2 > 0$: By solving the subsystem obtained from (iii) and (iv), we get $x = \frac{1}{3}(2b_2 - b_1, 2b_1 - b_2)$, and the equations (i) and (ii) then result in $\lambda = \frac{2}{9}(4b_2 - 5b_1, 4b_1 - 5b_2)$. The two Lagrange multipliers are positive if $\frac{5}{4}b_1 < b_2 < \frac{4}{5}b_1$. For these parameter values, x is the optimal point and $v_{\min}(b) = \frac{1}{9}(5b_1^2 - 8b_1b_2 + 5b_2^2)$ is the optimal value.

The result of our investigation is summarized in the following table:

	$v_{\min}(b)$	$-\lambda_1 = \dfrac{\partial v}{\partial b_1}$	$-\lambda_2 = \dfrac{\partial v}{\partial b_2}$
$b_1 \geq 0,\ b_2 \geq 0$	0	0	0
$b_1 < 0,\ b_2 \geq \frac{4}{5}b_1$	$\frac{1}{5}b_1^2$	$\frac{2}{5}b_1$	0
$b_2 < 0,\ b_2 \leq \frac{5}{4}b_1$	$\frac{1}{5}b_2^2$	0	$\frac{2}{5}b_2$
$\frac{5}{4}b_1 < b_2 < \frac{4}{5}b_1$	$\frac{1}{9}(5b_1^2 - 8b_1b_2 + 5b_2^2)$	$\frac{2}{9}(5b_1 - 4b_2)$	$\frac{2}{9}(5b_2 - 4b_1)$

$\square$

Click on the ad to read more

Exercises

11.1 Let $b > 0$ and consider the following trivial convex optimization problem

$$\begin{array}{ll} \min & x^2 \\ \text{s.t.} & x \geq b. \end{array}$$

Slater's condition is satisfied and the optimal value is attained at the point $\hat{x} = b$. Find the number $\hat{\lambda}$ which, according to Theorem 11.1.1, satisfies the optimality criterion.

11.2 Verify in the previous exercise that $v'(b) = \hat{\lambda}$.

11.3 Consider the minimization problem

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & \begin{cases} g_i(x) \leq 0, & i = 1, 2, \dots, p \\ g_i(x) = 0, & i = p+1, \dots, m \end{cases} \end{array}$$

with $x \in \Omega$ as implicit constraint, and the equivalent epigraph formulation

$$\text{(P}'\text{)} \qquad \begin{array}{ll} \min & t \\ \text{s.t.} & \begin{cases} f(x) - t \leq 0, & \\ \quad g_i(x) \leq 0, & i = 1, 2, \dots, p \\ \quad g_i(x) = 0, & i = p+1, \dots, m \end{cases} \end{array}$$

of the problem with $(t, x) \in \mathbf{R} \times \Omega$ as implicit constraint.

a) Show that (P′) satisfies Slater's condition if and only if (P) does.

b) Determine the relation between the Lagrange functions of the two problems and the relation between their dual functions.

c) Prove that the two dual problems have the same optimal value, and that the optimality criterion is satisfied in the minimization problem (P) if and only if it is satisfied in the problem (P′).

11.4 Prove for convex problems that Slater's condition is satisfied if and only if, for each non-affine constraint $g_i(x) \leq 0$, there is a feasible point $\overline{x}_i$ in the relative interior of Ω such that $g_i(\overline{x}_i) < 0$.

11.5 Let

$$\text{(P}_b\text{)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & \begin{cases} g_i(x) \leq b_i, & i = 1, 2, \dots, p \\ g_i(x) = b_i, & i = p+1, \dots, m \end{cases} \end{array}$$

be a convex problem, and suppose that its optimal value $v_{\min}(b)$ is $> -\infty$ for all right-hand sides b that belong to some convex subset U of $\mathbf{R}^m$. Prove that the restriction of $v_{\min}$ to U is a convex function.

Download free eBooks at bookboon.com

11.6 Solve the following convex optimization problems.

a) $\min \ e^{x_1-x_2}+e^{x_2}-x_1$ s.t. $x \in \mathbf{R}^2$

b) $\min \ e^{x_1-x_2}+e^{x_2}-x_1$ s.t. $\begin{cases} x_1^2+x_2^2 \le 1 \\ x_1+x_2 \ge -1 \end{cases}$

c) $\min \ -x_1-2x_2$ s.t. $\begin{cases} e^{x_1}+x_2 \le 1 \\ x_2 \ge 0 \end{cases}$

d) $\min \ x_1+2x_2$ s.t. $\begin{cases} x_1^2+x_2^2 \le 5 \\ x_1-x_2 \le 1 \end{cases}$

e) $\min \ x_1-x_2$ s.t. $\begin{cases} 0 < x_1 \le 2 \\ 0 \le x_2 \le \ln x_1 \end{cases}$

f) $\min \ e^{x_1}+e^{x_2}+x_1x_2$ s.t. $\begin{cases} x_1+x_2 \ge 1 \\ x_1,\, x_2 \ge 0 \end{cases}$

11.7 Solve the convex optimization problem

$$\begin{array}{ll} \min & x_1^2+x_2^2-\ln(x_1+x_2) \\ \text{s.t.} & \begin{cases} (x_1-1)^2+x_2^2 \le 9 \\ x_1+x_2 \ge 2 \\ x_1,\, x_2 \ge 0. \end{cases} \end{array}$$

11.8 Solve the convex optimization problem

$$\begin{array}{ll} \min & \displaystyle\sum_{j=1}^n v_j^{-1}\sqrt{y_j^2+a_j^2} \\ \text{s.t.} & \begin{cases} \sum_{j=1}^n y_j = b \\ y \in \mathbf{R}^n \end{cases} \end{array}$$

that occurred in our discussion of light refraction in Section 9.4, and verify *Snell's law of refraction:* $\sin\theta_i/\sin\theta_j = v_i/v_j$, where $\theta_j = \arctan y_j/a_j$.

11.9 Lisa has inherited 1 million dollars that she intends to invest by buying shares in three companies: A, B and C. Company A manufactures mobile phones, B manufactures antennas for mobile phones, and C manufactures ice cream. The annual return on an investment in the companies is a random variable, and the expected return for each company is estimated to be

	A	B	C
Expected return:	20%	12%	4%

Lisa's expected return if she invests x_1, x_2, x_3 million dollars in the three companies, is thus equal to

$$0.2\,x_1 + 0.12\,x_2 + 0.04\,x_3.$$

The investment risk is by definition the variance of the return. To calculate this we need to know the variance of each company's return and the correlation between the returns of the various companies. For obvious reasons,

Download free eBooks at bookboon.com

there is a strong correlation between sales in companies A and B, while sales of the company C only depend on whether the summer weather is beautiful or not, and not on the number of mobile phones sold. The so-called covariance matrix is in our case the matrix

$$\begin{bmatrix} 50 & 40 & 0 \\ 40 & 40 & 0 \\ 0 & 0 & 10 \end{bmatrix}$$

For those who know some basic probability theory, it is now easy to calculate the risk – it is given by the expression

$$50x_1^2 + 80x_1x_2 + 40x_2^2 + 10x_3^2.$$

Lisa, who is a careful person, wants to minimize her investment risk but she also wants to have an expected return of at least 12 %. Formulate and solve Lisa's optimization problem.

This e-book
is made with
SetaPDF

SETASIGN

PDF components for **PHP** developers

www.setasign.com

Download free eBooks at bookboon.com

Click on the ad to read more

11.10 Consider the consumer problem

$$\begin{array}{ll} \max & f(x) \\ \text{s.t.} & \begin{cases} \langle p, x\rangle \le I \\ x \ge 0 \end{cases} \end{array}$$

discussed in Section 9.4, where $f(x)$ is the consumer's utility function, assumed to be concave and differentiable, I is her disposable income, $p = (p_1, p_2, \dots, p_n)$ is the price vector and $x = (x_1, x_2, \dots, x_n)$ denotes a consumption bundle.

Suppose that $\hat{x}$ is an optimal solution. The optimal utility v, as well as $\hat{x}$, depends on the income I, of course; let us assume that $v = v(I)$ is a differentiable function. Show that under these assumptions

$$\begin{aligned} \hat{x}_j,\, \hat{x}_k > 0 \ &\Rightarrow\ \frac{1}{p_j}\frac{\partial f}{\partial x_j}\Big|_{\hat{x}} = \frac{1}{p_k}\frac{\partial f}{\partial x_k}\Big|_{\hat{x}} = \frac{dv}{dI} \\ \hat{x}_j = 0,\, \hat{x}_k > 0 \ &\Rightarrow\ \frac{1}{p_j}\frac{\partial f}{\partial x_j}\Big|_{\hat{x}} \le \frac{1}{p_k}\frac{\partial f}{\partial x_k}\Big|_{\hat{x}}. \end{aligned}$$

In words, this means:

The ratio between the marginal utility and the price of a commodity is for the optimal solution the same for all goods that are actually purchased, and it equals the marginal increase of utility at an increase of income. For goods that are not purchased, the corresponding ratio is not larger.

The conclusion is rather trivial, for it $x_k > 0$ and $\dfrac{1}{p_j}\dfrac{\partial f}{\partial x_j} > \dfrac{1}{p_k}\dfrac{\partial f}{\partial x_k}$, then the consumer benefits from changing a small quantity ϵ/p_k of commodity no. k to the quantity ϵ/p_j of commodity no. j.

Chapter 12

Linear programming

Linear programming (LP) is the art of optimizing linear functions over polyhedra, described as solution sets to systems of linear inequalities. In this chapter, we describe and study the basic mathematical theory of linear programming, above all the very important duality concept.

12.1 Optimal solutions

The optimal value of a general optimization problem was defined in Chapter 9. In particular, each LP problem

$$\text{(P)} \qquad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & x \in X \end{array}$$

has an optimal value, which in this section will be denoted by $v_{\min}(c)$ to indicate its dependence of the objective function.

LP problems with finite optimal values always have optimal solutions. The existence of an optimal solution is of course obvious if the polyhedron of feasible points is bounded, i.e. compact, since the objective function is continuous. For arbitrary LP problems, we rely on the representation theorem for polyhedra to prove the existence of optimal solutions.

Theorem 12.1.1. *Suppose that the polyhedron X of feasible solutions in the LP problem (P) is nonempty and a subset of $\mathbf{R}^n$. Then we have:*

(i) The value function $v_{\min}\colon \mathbf{R}^n \to \underline{\mathbf{R}}$ is concave with effective domain

$$\operatorname{dom} v_{\min} = (\operatorname{recc} X)^+.$$

The objective function $\langle c, x\rangle$ is, in other words, bounded below on X if and only if c belongs to the dual cone of the recession cone of X.

Download free eBooks at bookboon.com

(ii) The problem has optimal solutions for each $c \in (\operatorname{recc} X)^+$*, and the set of optimal solutions is a polyhedron. Moreover, the optimum is attained at some extreme point of* X *if* X *is a line-free polyhedron.*

Proof. By definition, the optimal value $v_{\min}(c) = \inf\{\langle c, x\rangle \mid x \in X\}$ is the pointwise infimum of a family of concave functions, namely the linear functions $c \mapsto \langle c, x\rangle$, with x running through X. So the value function $v_{\min}$ is concave by Theorem 6.2.4 in Part I.

Let us now determine $\operatorname{dom} v_{\min}$, i.e. the set of c such that $v_{\min}(c) > -\infty$. By the structure theorem for polyhedra (Theorem 5.3.1 in Part I), there is a finite nonempty set A such that $X = \operatorname{cvx} A + \operatorname{recc} X$, where $A = \operatorname{ext} X$ if the polyhedron is line-free. The optimal value $v_{\min}(c)$ can therefore be calculated as follows:

$$\begin{aligned} v_{\min}(c) &= \inf\{\langle c, y+z\rangle \mid y \in \operatorname{cvx} A,\ z \in \operatorname{recc} X\} \\ &= \inf\{\langle c, y\rangle \mid y \in \operatorname{cvx} A\} + \inf\{\langle c, z\rangle \mid z \in \operatorname{recc} X\} \\ &= \min\{\langle c, y\rangle \mid y \in A\} + \inf\{\langle c, z\rangle \mid z \in \operatorname{recc} X\}, \end{aligned} \tag{12.1}$$

The equality $\inf\{\langle c, y\rangle \mid y \in \operatorname{cvx} A\} = \min\{\langle c, y\rangle \mid y \in A\}$ holds because of Theorem 6.3.3 in Part I, since linear functions are concave.

If c belongs to the dual cone $(\operatorname{recc} X)^+$, then $\langle c, z\rangle \geq 0$ for all vectors $z \in \operatorname{recc} X$ with equality for $z = 0$, and it follows from equation (12.1) that

$$v_{\min}(c) = \min\{\langle c, y\rangle \mid y \in A\} > -\infty.$$

This proves the inclusion $(\operatorname{recc} X)^+ \subseteq \operatorname{dom} v_{\min}$, and that the optimal value is attained at a point in A, and then in particular at some extreme point of X if the polyhedron X is line-free.

If $c \notin (\operatorname{recc} X)^+$, then $\langle c, z_0\rangle < 0$ for some vector $z_0 \in \operatorname{recc} X$. Since $tz_0 \in \operatorname{recc} X$ for $t > 0$ and $\lim_{t\to\infty}\langle c, tz_0\rangle = -\infty$, it follows that

$$\inf\{\langle c, z\rangle \mid z \in \operatorname{recc} X\} = -\infty,$$

and equation (12.1) now implies that $v_{\min}(c) = -\infty$. This concludes the proof of the equality $\operatorname{dom} v_{\min} = (\operatorname{recc} X)^+$.

The set of minimum points to an LP problem with finite value $v_{\min}$ is equal to the intersection

$$X \cap \{x \in \mathbf{R}^n \mid \langle c, x\rangle = v_{\min}\}$$

between the polyhedron X and a hyperplane, and it is consequently a polyhedron. □

Download free eBooks at bookboon.com

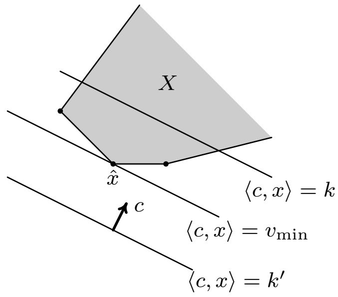


Figure 12.1. The minimum of $\langle c, x\rangle$ over the line-free polyhedron X is attained at an extreme point.

EXAMPLE 12.1.1. The polyhedron X of feasible points for the LP problem

$$\begin{array}{ll} \min & x_1 + x_2 \\ \text{s.t.} & \begin{cases} x_1 - x_2 \geq -2 \\ x_1 + x_2 \geq 1 \\ -x_1 \geq -3 \end{cases} \end{array}$$

has three extreme points, namely $(3, 5)$, $(-\frac{1}{2}, \frac{3}{2})$ and $(3, -2)$. The values of the objective function $f(x) = x_1 + x_2$ at these points are $f(3, 5) = 8$ and $f(-\frac{1}{2}, \frac{3}{2}) = f(3, -2) = 1$. The least of these is 1, which is the optimal value. The optimal value is attained at two extreme points, $(\frac{1}{2}, \frac{3}{2})$ och $(3, -2)$, and thus also at all points on the line segment between those two points.

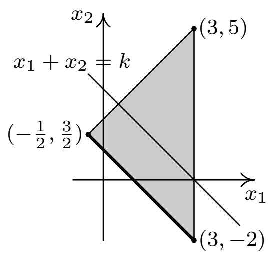


Figure 12.2. Illustration for Example 12.1.1. □

Suppose that $X = \{x \in \mathbf{R}^n \mid Ax \geq b\}$ is a line-free polyhedron and that we want to minimize a given linear function over X. To determine the optimal value of this LP problem, we need according to the previous theorem, assuming that the objective function is bounded below on X, only calculate function values at the finitely many extreme points of X. In theory, this

Download free eBooks at bookboon.com

is easy, but in practice it can be an insurmountable problem, because the number of extreme points may be extremely high. The number of potential extreme points of X when A is an $m \times n$-matrix, equals $\binom{m}{n}$, which for $m = 100$ and $n = 50$ is a number that is greater than 10^{29}. The simplex algorithm, which we will study in Chapter 13, is based on the idea that it is not necessary to search through all the extreme points; the algorithm generates instead a sequence $x_1, x_2, x_3, \ldots$ of extreme points with decreasing objective function values $\langle c, x_1\rangle \geq \langle c, x_2\rangle \geq \langle c, x_3\rangle \geq \ldots$ until the minimum point is found. The number of extreme points that needs to be investigated is therefore generally relatively small.

Sensitivity analysis

Let us rewrite the polyhedron of feasible points in the LP problem

$$\text{(P)} \qquad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & x \in X \end{array}$$

as

$$X = \operatorname{cvx} A + \operatorname{con} B$$

Download free eBooks at bookboon.com

Click on the ad to read more

with finite sets A and B. We know from the preceding theorem and its proof that a feasible point $\overline{x}$ is optimal for the LP problem if and only if

$$\begin{cases} \langle c, a\rangle \geq \langle c, \overline{x}\rangle & \text{for all } a \in A \\ \langle c, b\rangle \geq 0 & \text{for all } b \in B, \end{cases}$$

and these inequalities define a convex cone $C_{\overline{x}}$ in the variable c. The set of all c for which a given feasible point is optimal, is thus a convex cone.

Now suppose that $\overline{x}$ is indeed an optimal solution to (P). How much can we change the coefficients of the objective function without changing the optimal solution? The study of this issue is an example of *sensitivity analysis.*

Expressed in terms of the cone $C_{\overline{x}}$, the answer is simple: If we change the coefficients of the objective function to $c + \Delta c$, then $\overline{x}$ is also an optimal solution to the perturbed LP problem

$$\text{(P}'\text{)} \qquad \begin{array}{ll} \min & \langle c + \Delta c, x\rangle \\ \text{s.t.} & x \in X \end{array}$$

if and only if $c + \Delta c$ belongs to the cone $C_{\overline{x}}$, i.e. if and only if Δc lies in the polyhedron $-c + C_{\overline{x}}$.

In summary, we have thus come to the following conclusions.

Theorem 12.1.2. *(i) The set of all c for which a given feasible point is optimal in the LP problem (P), is a convex cone.*

(ii) If $\overline{x}$ is an optimal solution to problem (P), then there is a polyhedron such that $\overline{x}$ is also an optimal solution to the perturbed LP problem (P$'$) for all Δc in the polyhedron.

The set $\{\Delta c_k \mid \Delta c \in -c + C_{\overline{x}} \text{ and } \Delta c_j = 0 \text{ for } j \neq k\}$ is a (possibly unbounded) closed interval $[-d_k, e_k]$ around 0. An optimal solution to the problem (P) is therefore also optimal for the perturbed problem that is obtained by only varying the objective coefficient c_k, provided that the perturbation Δc_k lies in the interval $-d_k \leq \Delta c_k \leq e_k$. Many computer programs for LP problems, in addition to generating the optimal value and the optimal solution, also provide information about these intervals.

Sensitivity analysis will be studied in connection with the simplex algorithm in Chapter13.7.

Example 12.1.2. The printout of a computer program that was used to solve an LP problem with $c = (20, 30, 40, \dots)$ contained among other things the following information:

Download free eBooks at bookboon.com

Optimal value: 4000 **Optimal solution:** $\overline{x} = (50, 40, 10, \dots)$

Sensitivity report:

Variable	Value	Objective coeff.	Allowable decrease	Allowable increase
x_1	50	20	15	5
x_2	40	30	10	10
x_3	10	40	15	20
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Use the printout to determine the optimal solution and the optimal value if the coefficients c_1, c_2 and c_3 are changed to 17, 35 and 45, respectively, and the other objective coefficients are left unchanged.

Solution: The columns "Allowable decrease" and "Allowable increase" show that the polyhedron of changes Δc that do not affect the optimal solution contains the vectors $(-15, 0, 0, 0, \dots)$, $(0, 10, 0, 0, \dots)$ and $(0, 0, 20, 0, \dots)$. Since

$$(-3, 5, 5, 0, \dots) = \tfrac{1}{5}(-15, 0, 0, 0, \dots) + \tfrac{1}{2}(0, 10, 0, 0, \dots) + \tfrac{1}{4}(0, 0, 20, 0, \dots)$$

and $\frac{1}{5} + \frac{1}{2} + \frac{1}{4} = \frac{19}{20} < 1$, $\Delta c = (-3, 5, 5, 0, \dots)$ is a convex combination of changes that do not affect the optimal solutions, namley the three changes mentioned above and $(0, 0, 0, 0, \dots)$. The solution $\overline{x} = (50, 40, 10, \dots)$ is therefore still optimal for the LP problem with $c = (17, 35, 45, \dots)$. However, the new optimal value is of course $4000 - 20 \cdot 3 + 30 \cdot 5 + 40 \cdot 5 = 4290$. $\square$

12.2 Duality

Dual problems

By describing the polyhedron X in a linear minimization problem

$$\begin{array}{ll} \min & \langle c, x \rangle \\ \text{s.t.} & x \in X \end{array}$$

as the solution set of a system of linear inequalities, we get a problem with a corresponding Lagrange function, and hence also a dual function and a dual problem. The description of X as a solution set is of course not unique, so the dual problem is not uniquely determined by X as a polyhedron, but whichever description we choose, we get, according to Theorem 11.1.1, a dual problem, where strong duality holds, because Slater's condition is satisfied for convex problems with affine constraints.

Download free eBooks at bookboon.com

In this section, we describe the dual problem for some commonly occurring polyhedron descriptions, and we give an alternative proof of the duality theorem. Our premise is that the polyhedron X is given as

$$X = \{x \in U^+ \mid Ax - b \in V^+\},$$

where

- U and V are finitely generated cones in $\mathbf{R}^n$ and $\mathbf{R}^m$, respectively;
- A is an $m \times n$-matrix;
- b is a vector in $\mathbf{R}^m$.

As usual, we identify vectors with column matrices and matrices with linear transformations. The set X is of course a polyhedron, for by writing

$$X = U^+ \cap A^{-1}(b + V^+)$$

we see that X is an intersection of two polyhedra – the conical polyhedron U^+ and the inverse image $A^{-1}(b + V^+)$ under the linear map A of the polyhedron $b + V^+$.

Deloitte.

Discover the truth at www.deloitte.ca/careers

© Deloitte & Touche LLP and affiliated entities.

Download free eBooks at bookboon.com

Click on the ad to read more

The LP problem of minimizing $\langle c, x\rangle$ over the polyhedron X with the above description will now be written

$$\text{(P)} \qquad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax - b \in V^+,\ x \in U^+ \end{array}$$

and in order to form a suitable dual problem we will perceive the condition $x \in U^+$ as an implicit constraint and express the other condition $Ax-b \in V^+$ as a system of linear inequalities. Assume therefore that the finitely generated cone V is generated by the columns of the $m \times k$-matrix D, i.e. that

$$V = \{Dz \mid z \in \mathbf{R}^k_+\}.$$

The dual cone V^+ can then be written as

$$V^+ = \{y \in \mathbf{R}^m \mid D^T y \geq 0\},$$

and the constraint $Ax - b \in V^+$ can now be expressed as a system of inequalities, namely $D^T Ax - D^T b \geq 0$.

Our LP problem (P) has thus been transformed into

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & D^T b - D^T Ax \leq 0,\ x \in U^+. \end{array}$$

The associated Lagrange function $L\colon U^+ \times \mathbf{R}^k_+ \to \mathbf{R}$ is defined by

$$L(x, \lambda) = \langle c, x\rangle + \langle \lambda, D^T b - D^T Ax\rangle = \langle c - A^T D\lambda, x\rangle + \langle b, D\lambda\rangle,$$

and the corresponding dual function $\phi\colon \mathbf{R}^k_+ \to \underline{\mathbf{R}}$ is given by

$$\phi(\lambda) = \inf_{x \in U^+} L(x, \lambda) = \begin{cases} \langle b, D\lambda\rangle\,, & \text{if } c - A^T D\lambda \in U \\ -\infty\,, & \text{otherwise.} \end{cases}$$

This gives us a dual problem of the form

$$\begin{array}{ll} \max & \langle b, D\lambda\rangle \\ \text{s.t.} & c - A^T D\lambda \in U,\ \lambda \in \mathbf{R}^k_+. \end{array}$$

Since $D\lambda$ describes the cone V as λ runs through $\mathbf{R}^k_+$, we can by setting $y = D\lambda$ reformulate the dual problem so that it becomes

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & c - A^T y \in U,\ y \in V. \end{array}$$

It is therefore natural to define duality for LP problems of the form (P) as follows.

Download free eBooks at bookboon.com

Definition. Given the LP problem

$$\text{(P)} \qquad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax - b \in V^+,\ x \in U^+, \end{array}$$

which we call the *primal* problem, we call the problem

$$\text{(D)} \qquad \begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & c - A^T y \in U,\ y \in V \end{array}$$

the *dual* LP problem.

The optimal values of the two problems are denoted by $v_{\min}(P)$ and $v_{\max}(D)$. The polyhedron of feasible points will be denoted by X for the primal problem and by Y for the dual problem.

EXAMPLE 12.2.1. Different choices of the cones U and V give us different concrete dual problems (P) and (D). We exemplify with four important special cases.

1. The choice $U = \{0\}$, $U^+ = \mathbf{R}^n$ and $V = V^+ = \mathbf{R}^m_+$ gives us the following dual pair:

$$(\text{P}_1) \quad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax \geq b \end{array} \qquad \text{and} \qquad (\text{D}_1) \quad \begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & A^T y = c,\ y \geq 0. \end{array}$$

Every LP problem can be expressed in the form (P_1), because every polyhedron can be expressed as an intersection of halfspaces, i.e. be written as $Ax \geq b$.

2. The choice $U = U^+ = \mathbf{R}^n_+$ and $V = V^+ = \mathbf{R}^m_+$ gives instead the dual pair:

$$(\text{P}_2) \quad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax \geq b,\ x \geq 0 \end{array} \qquad \text{and} \qquad (\text{D}_2) \quad \begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & A^T y \leq c,\ y \geq 0. \end{array}$$

This is the most symmetric formulation of duality, and the natural formulation for many application problems with variables that represent physical quantities or prices, which of course are nonnegative. The diet problem and the production planning problem in Chapter 9.4 are examples of such problems.

3. $U = U^+ = \mathbf{R}^n_+$, $V = \mathbf{R}^m$ and $V^+ = \{0\}$ result in the dual pair:

$$(\text{P}_3) \quad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b,\ x \geq 0 \end{array} \qquad \text{and} \qquad (\text{D}_3) \quad \begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & A^T y \leq c. \end{array}$$

The formulation (P_3) is the natural starting point for the simplex algorithm.

Download free eBooks at bookboon.com

4. The choice $U = \{0\}$, $U^+ = \mathbf{R}^n$, $V = \mathbf{R}^m$ and $V^+ = \{0\}$ gives us the pair

$$
(\mathrm{P}_4) \quad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b \end{array} \qquad \text{and} \qquad (\mathrm{D}_4) \quad \begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & A^T y = c. \end{array}
$$

□

Example 12.2.2. A trivial example of dual LP problems in one variable is

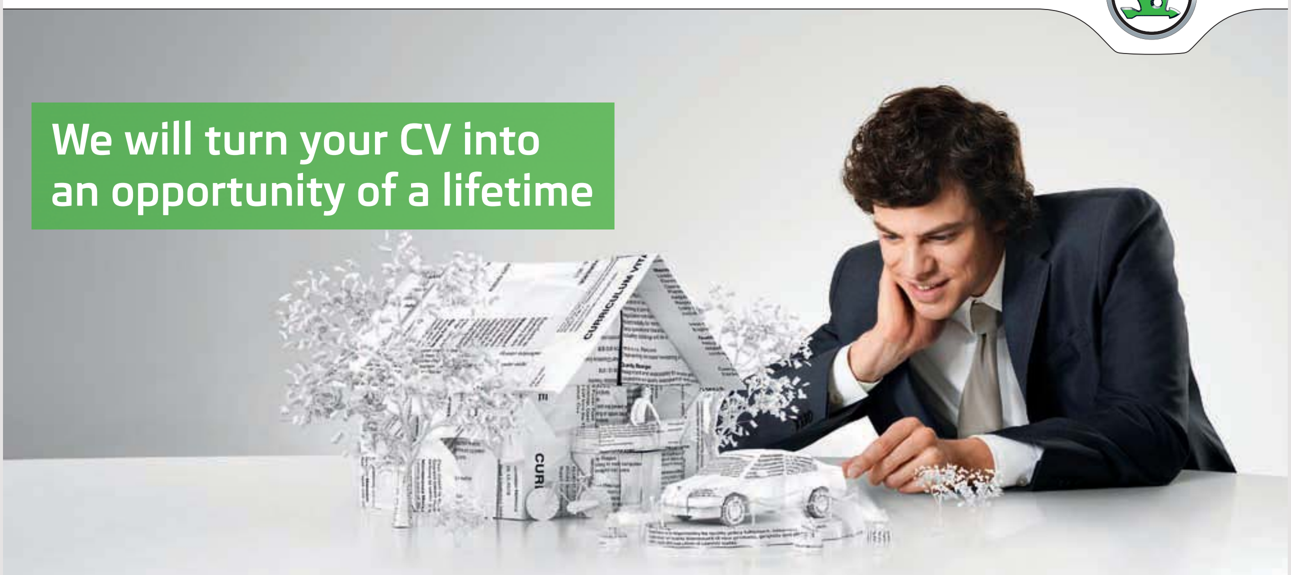
$$
\begin{array}{ll} \min & 5x \\ \text{s.t.} & 2x \ge 4 \end{array} \qquad \text{and} \qquad \begin{array}{ll} \max & 4y \\ \text{s.t.} & 2y = 5,\ y \ge 0 \end{array}
$$

Both problems have the optimal value 10. □

Example 12.2.3. The problems

$$
\begin{array}{ll} \min & x_1 + x_2 \\ \text{s.t.} & \begin{cases} x_1 - x_2 \ge -2 \\ x_1 + x_2 \ge 1 \\ -x_1 \ge -3 \end{cases} \end{array} \qquad \text{and} \qquad \begin{array}{ll} \max & -2y_1 + y_2 - 3y_3 \\ \text{s.t.} & \begin{cases} y_1 + y_2 - y_3 = 1 \\ -y_1 + y_2 = 1 \\ y_1, y_2, y_3 \ge 0 \end{cases} \end{array}
$$

are dual. The optimal solutions to the primal minimization problem were determined in Example 12.1.1 and the optimal value was found to be 1. The

SIMPLY CLEVER
ŠKODA


Do you like cars? Would you like to be a part of a successful brand?
We will appreciate and reward both your enthusiasm and talent.
Send us your CV. You will be surprised where it can take you.

Send us your CV on
www.employerforlife.com

Download free eBooks at bookboon.com
Click on the ad to read more

feasible points for the dual maximization problem are of the form

$$y = (t, 1 + t, 2t)$$

with $t \geq 0$, and the corresponding values of the objective function are $1 - 7t$. The maximum value is attained for $t = 0$ at the point $(0, 1, 0)$, and the maximum value is equal to 1. □

The Duality Theorem

The primal and dual problems in Examples 12.2.2 and 12.2.3 have the same optimal value, and this is no coincidence but a consequence of the duality theorem, which is formulated below and is a special case of the duality theorem for general convex problems (Theorem 11.1.1). In this section we give an alternative proof of this important theorem, and we start with the trivial result about weak duality.

Theorem 12.2.1 (Weak duality). *The optimal values of the two dual LP problems (P) and (D) satisfy the inequality*

$$v_{\max}(D) \leq v_{\min}(P).$$

Proof. The inequality is trivially satisfied if any of the two polyhedra X and Y of feasible points is empty, because if $Y = \emptyset$ then $v_{\max}(D) = -\infty$, by definition, and if $X = \emptyset$ then $v_{\min}(P) = +\infty$, by definition.

Assume therefore that both problems have feasible points. If $x \in X$ and $y \in Y$, then $y \in V$, $(Ax - b) \in V^+$, $(c - A^T y) \in U$ and $x \in U^+$, by definition, and hence $\langle Ax - b, y\rangle \geq 0$ and $\langle c - A^T y, x\rangle \geq 0$. It follows that

$$\begin{aligned}\langle b, y\rangle &\leq \langle b, y\rangle + \langle c - A^T y, x\rangle = \langle b, y\rangle + \langle c, x\rangle - \langle y, Ax\rangle \\ &= \langle c, x\rangle + \langle b, y\rangle - \langle Ax, y\rangle = \langle c, x\rangle - \langle Ax - b, y\rangle \leq \langle c, x\rangle.\end{aligned}$$

The objective function $\langle b, y\rangle$ in the maximization problem (D) is in other words bounded above on Y by $\langle c, x\rangle$ for each $x \in X$, and hence

$$v_{\max}(D) = \sup_{y \in Y} \langle b, y\rangle \leq \langle c, x\rangle.$$

The objective function $\langle c, x\rangle$ in the minimization problem (P) is therefore bounded below on X by $v_{\max}(D)$. This implies that $v_{\max}(D) \leq v_{\min}(P)$ and completes the proof of the theorem. □

The following optimality criterion follows from weak duality.

Download free eBooks at bookboon.com

Theorem 12.2.2 (Optimality criterion). *Suppose that $\hat{x}$ is a feasible point for the minimization problem (P), that $\hat{y}$ is a feasible point for the dual maximization problem (D), and that*

$$\langle c, \hat{x} \rangle = \langle b, \hat{y} \rangle.$$

Then $\hat{x}$ and $\hat{y}$ are optimal solutions of the respective problems.

Proof. The assumptions on $\hat{x}$ and $\hat{y}$ combined with Theorem 12.2.1 give us the following chain of inequalities

$$v_{\max}(D) \geq \langle b, \hat{y} \rangle = \langle c, \hat{x} \rangle \geq v_{\min}(P) \geq v_{\max}(D).$$

Since the two extreme ends are equal, there is equality everywhere, which means that $\hat{y}$ is a maximum point and $\hat{x}$ is a minimum point. □

Theorem 12.2.3 (Duality theorem). *Suppose that at least one of the two dual LP problems*

$$\text{(P)} \qquad \begin{array}{ll} \min & \langle c, x \rangle \\ \text{s.t.} & Ax - b \in V^+,\ x \in U^+ \end{array}$$

and

$$\text{(D)} \qquad \begin{array}{ll} \max & \langle b, y \rangle \\ \text{s.t.} & c - A^T y \in U,\ y \in V \end{array}$$

has feasible points. Then, the two problem have the same optimal value.

Thus, provided that at least one of the two dual problems has feasible points:
(i) $X = \emptyset \;\Leftrightarrow\;$ the objective function $\langle b, y \rangle$ is not bounded above on Y.
(ii) $Y = \emptyset \;\Leftrightarrow\;$ the objective function $\langle c, x \rangle$ is not bounded below on X.
(iii) If $X \neq \emptyset$ and $Y \neq \emptyset$, then there exist points $\hat{x} \in X$ and $\hat{y} \in Y$ such that $\langle b, y \rangle \leq \langle b, \hat{y} \rangle = \langle c, \hat{x} \rangle \leq \langle c, x \rangle$ for all $x \in X$ and all $y \in Y$.

The duality theorem for linear programming problems is a special case of the general duality theorem for convex problems, but we give here an alternative proof based directly on the following variant of Farkas's lemma.

Lemma. *The system*

$$\text{(12.2)} \qquad \begin{cases} \langle c, x \rangle \leq \alpha \\ x \in X \end{cases}$$

has a solution if and only if the systems

$$\text{(12.3-A)} \qquad \begin{cases} \langle b, y \rangle > \alpha \\ y \in Y \end{cases} \qquad \text{and} \qquad \text{(12.3-B)} \qquad \begin{cases} \langle b, y \rangle = 1 \\ -A^T y \in U \\ y \in V \end{cases}$$

both have no solutions,

Download free eBooks at bookboon.com

Proof. The system (12.2), i.e.

$$\begin{cases} \langle c, x\rangle \le \alpha \\ Ax - b \in V^+ \\ x \in U^+, \end{cases}$$

is solvable if and only if the following homogenized system is solvable:

$$\begin{cases} \langle c, x\rangle \le \alpha t \\ Ax - bt \in V^+ \\ x \in U^+ \\ t \in \mathbf{R} \\ t > 0. \end{cases} \tag{12.2$'$}$$

(If x solves the system (12.2), then $(x, 1)$ solves the system (12.2′), and if (x, t) solves the system (12.2′), then x/t solves the system (12.2).) We can write the system (12.2′) more compactly by introducing the matrix

$$\tilde{A} = \begin{bmatrix} \alpha & -c^T \\ -b & A \end{bmatrix}$$

Download free eBooks at bookboon.com

Click on the ad to read more

and the vectors $\tilde{x} = (t, x) \in \mathbf{R} \times \mathbf{R}^n$ and $d = (-1, 0) \in \mathbf{R} \times \mathbf{R}^n$, namely as

$$(12.2'') \qquad \begin{cases} \tilde{A}\tilde{x} \in \mathbf{R}_+ \times V^+ \\ \tilde{x} \in \mathbf{R} \times U^+ \\ d^T\tilde{x} < 0. \end{cases}$$

By Theorem 3.3.2 in Part I, the system (12.2″) is solvable if and only if the following dual system has no solutions:

$$(12.3'') \qquad \begin{cases} d - \tilde{A}^T\tilde{y} \in \{0\} \times U \\ \tilde{y} \in \mathbf{R}_+ \times V. \end{cases}$$

Since

$$\tilde{A}^T = \begin{bmatrix} \alpha & -b^T \\ -c & A^T \end{bmatrix},$$

we obtain the following equivalent system from (12.3″) by setting $\tilde{y} = (s, y)$ with $s \in \mathbf{R}$ and $y \in \mathbf{R}^m$:

$$(12.3') \qquad \begin{cases} -1 - \alpha s + \langle b, y\rangle = 0 \\ cs - A^Ty \in U \\ y \in V \\ s \geq 0. \end{cases}$$

The system (12.2) is thus solvable if and only if the system (12.3′) has no solutions, and by considering the cases $s > 0$ and $s = 0$ separately, we see that the system (12.3′) has no solutions if and only if the two systems

$$\begin{cases} \langle b, y/s\rangle = \alpha + 1/\mathrm{s} \\ c - A^T(y/s) \in U \\ y/s \in V \\ s > 0 \end{cases} \quad \text{and} \quad \begin{cases} \langle b, y\rangle = 1 \\ -A^Ty \in U \\ y \in V \end{cases}$$

have no solutions, and this is obviously the case if and only if the systems (12.3-A) and (12.3-B) both lack solutions. □

Proof of the duality theorem. We now return to the proof of the duality theorem, and because of weak duality, we only need to show the inequality

$$(12.4) \qquad v_{\min}(P) \leq v_{\max}(D).$$

We divide the proof of this inequality in three separate cases.

Download free eBooks at bookboon.com

Case 1. $Y \neq \emptyset$ **and the system (12.3-B) has no solution.**

The inequality (12.4) is trivially true if $v_{\max}(D) = \infty$. Therefore, assume that $v_{\max}(D) < \infty$. Then, because of the definition of $v_{\max}(D)$, the system (12.3-A) has no solution when $\alpha = v_{\max}(D)$. So neither of the two systems in (12.3) has a solution for $\alpha = v_{\max}(D)$. Thus, the system (12.2) has a solution for this α-value by the lemma, which means that there is a feasible point $\hat{x}$ such that $\langle c, \hat{x}\rangle \leq v_{\max}(D)$. Consequently, $v_{\min}(P) \leq \langle c, \hat{x}\rangle \leq v_{\max}(D)$.

Note that it follows from the proof that the minimization problem actually has an optimal solution $\hat{x}$.

Case 2. $Y = \emptyset$ **and the system (12.3-B) has no solution.**

The system (12.3-A) now lacks solutions for all values of α, so it follows from the lemma that the system (12.2) is solvable for all α-values, and this means that the objective function $\langle c, x\rangle$ is unbounded below on X. Hence, $v_{\min}(P) = -\infty = v_{\max}(D)$ in this case.

Case 3. The system (12.3-B) has a solution

It now follows from the lemma that the system (12.2) has no solution for all values of α, and this implies that the set X of feasible solutions is empty. The polyhedron Y of feasible points in the dual problem is consequently nonempty. Choose a point $y_0 \in Y$, let $\overline{y}$ be a solution to the system (12.3-B) and consider the points $y^t = y_0 + t\overline{y}$ for $t > 0$. The vectors y^t belong to V, because they are conical combinations of vectors in V. Moreover, the vectors $c - A^T y^t = (c - A^T y_0) - tA^T\overline{y}$ are conic combinations of vectors in U and thus belong to U. This means that the vector y^t lies in Y for $t > 0$, and since

$$\langle b, y^t\rangle = \langle b, y_0\rangle + t\langle b, \overline{y}\rangle = \langle b, y_0\rangle + t \to +\infty$$

as $t \to \infty$, we conclude that $v_{\max}(D) = \infty$. The inequality (12.4) is in other words trivially fulfilled. □

The Complementary Theorem

Theorem 12.2.4 (Complementary theorem)**.** *Suppose that $\hat{x}$ is a feasible point for the LP problem (P) and that $\hat{y}$ is a feasible point for the dual LP problem (D). Then, the two points are optimal for their respective problems if and only if*

$$\langle c - A^T\hat{y}, \hat{x}\rangle = \langle A\hat{x} - b, \hat{y}\rangle = 0.$$

Proof. Note first that due to the definition of the polyhedra X and Y of feasible points, we have $\langle Ax - b, y\rangle \geq 0$ for all points $x \in X$ and $y \in V$, while $\langle c - A^T y, x\rangle \geq 0$ for all points $y \in Y$ and $x \in U$.

Download free eBooks at bookboon.com

In particular, $\langle A\hat{x} - b, \hat{y}\rangle \geq 0$ and $\langle c - A^T\hat{y}, \hat{x}\rangle \geq 0$ if $\hat{x}$ is an optimal solution to the primal problem (P) and $\hat{y}$ is an optimal solution to the dual problem (D). Moreover, $\langle c, \hat{x}\rangle = \langle b, \hat{y}\rangle$ because of the Duality theorem, so it follows that

$$\langle c, \hat{x}\rangle - \langle A\hat{x} - b, \hat{y}\rangle \leq \langle c, \hat{x}\rangle = \langle b, \hat{y}\rangle \leq \langle b, \hat{y}\rangle + \langle c - A^T\hat{y}, \hat{x}\rangle = \langle c, \hat{x}\rangle - \langle A\hat{x} - b, \hat{y}\rangle.$$

Since the two extreme ends of this inequality are equal, we have equality everywhere, i.e. $\langle A\hat{x} - b, \hat{y}\rangle = \langle c - A^T\hat{y}, \hat{x}\rangle = 0$.

Conversely, if $\langle c - A^T\hat{y}, \hat{x}\rangle = \langle A\hat{x} - b, \hat{y}\rangle = 0$, then $\langle c, \hat{x}\rangle = \langle A^T\hat{y}, \hat{x}\rangle$ and $\langle b, \hat{y}\rangle = \langle A\hat{x}, \hat{y}\rangle$, and since $\langle A^T\hat{y}, \hat{x}\rangle = \langle A\hat{x}, \hat{y}\rangle$, we conclude that $\langle c, \hat{x}\rangle = \langle b, \hat{y}\rangle$. The optimality of the two points now follows from the Optimality criterion. □

Let us for clarity formulate the Complementarity theorem in the important special case when the primal and dual problems have the form described as Case 2 in Example 12.2.1.

Corollary 12.2.5. *Suppose that $\hat{x}$ and $\hat{y}$ are feasible points for the dual problems*

$$(\mathrm{P}_2) \qquad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax \geq b,\ x \geq 0 \end{array}$$

and

$$(\mathrm{D}_2) \qquad \begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & A^Ty \leq c,\ y \geq 0. \end{array}$$

respectively. Then, they are optimal solutions if and only if

$$\begin{cases} (A\hat{x})_i > b_i \ \Rightarrow \ \hat{y}_i = 0 \\ \quad \hat{x}_j > 0 \ \Rightarrow \ (A^T\hat{y})_j = c_j \end{cases} \tag{12.5}$$

In words we can express condition (12.5) as follows, which explains the term 'complementary slackness': If $\hat{x}$ satisfies an individual inequality in the system $Ax \geq b$ strictly, then the corresponding dual variable $\hat{y}_i$ has to be equal to zero, and if $\hat{y}$ satisfies an individual inequality in the system $A^Ty \leq c$ strictly, then the corresponding primal variable x_j has to be equal to zero.

Proof. Since $\langle A\hat{x} - b, \hat{y}\rangle = \sum_{i=1}^m ((A\hat{x})_i - b_i)\hat{y}_i$ is a sum of nonnegative terms, we have $\langle A\hat{x} - b, \hat{y}\rangle = 0$ if and only if all the terms are equal to zero, i.e. if and only if $(A\hat{x})_i > b_i \Rightarrow \hat{y}_i = 0$.

Similarly, $\langle c - A^T\hat{y}, \hat{x}\rangle = 0$ if and only if $\hat{x}_j > 0 \Rightarrow (A^T\hat{y})_j = c_j$. The corollary is thus just a reformulation of Theorem 12.2.4 for dual problems of type (P_2)–(D_2). □

Download free eBooks at bookboon.com

The curious reader may wonder whether the implications in the condition (12.5) can be replaced by equivalences. The following trivial example shows that this is not the case.

Example 12.2.4. Consider the dual problems

$$\begin{array}{ll} \min & x_1 + 2x_2 \\ \text{s.t.} & x_1 + 2x_2 \geq 2, \;\; x \geq 0 \end{array} \qquad \text{and} \qquad \begin{array}{ll} \max & 2y \\ \text{s.t.} & \begin{cases} y \leq 1 \\ 2y \leq 2, \;\; y \geq 0 \end{cases} \end{array}$$

with $A = c^T = \begin{bmatrix} 1 & 2 \end{bmatrix}$ and $b = \begin{bmatrix} 2 \end{bmatrix}$. The condition (12.5) is not fulfilled with equivalence at the optimal points $\hat{x} = (2, 0)$ and $\hat{y} = 1$, because $\hat{x}_2 = 0$ and $(A^T\hat{y})_2 = 2 = c_2$.

However, there are other optimal solutions to the minimization problem; all points on the line segment between $(2, 0)$ and $(0, 1)$ are optimal, and the optimal pairs $\hat{x} = (2 - 2t, t)$ and $\hat{y} = 1$ satisfy the condition (12.5) with equivalence for $0 < t < 1$. □

The last conclusion in the above example can be generalized. All dual problems with feasible points have a pair of optimal solutions $\hat{x}$ and $\hat{y}$ that satisfy the condition (12.5) with implications replaced by equivalences. See exercise 12.8.

Download free eBooks at bookboon.com

Click on the ad to read more

EXAMPLE 12.2.5. The LP problem

$$\begin{array}{ll} \min & -x_1 + 2x_2 + x_3 + 2x_4 \\ \text{s.t.} & \begin{cases} -x_1 - x_2 - 2x_3 + x_4 \geq 4 \\ -2x_1 + x_2 + 3x_3 + x_4 \geq 8 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases} \end{array}$$

is easily solved by first solving the dual problem

$$\begin{array}{ll} \max & 4y_1 + 8y_2 \\ \text{s.t.} & \begin{cases} -y_1 - 2y_2 \leq -1 \\ -y_1 + y_2 \leq 2 \\ -2y_1 + 3y_2 \leq 1 \\ y_1 + y_2 \leq 2 \\ y_1, y_2 \geq 0 \end{cases} \end{array}$$

graphically and then using the Complementary theorem.

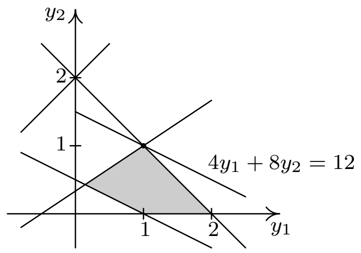


Figure 12.3. A graphical solution to the maximization problem in Ex. 12.2.5.

A graphical solution is obtained from figure 12.3, which shows that $\hat{y} = (1,1)$ is the optimal point and that the value is 12. Since $\hat{y}$ satisfies the first two constraints with strict inequality and $\hat{y}_1 > 0$ and $\hat{y}_2 > 0$, we obtain the optimal solution $\hat{x}$ to the minimization problem as a solution to the system

$$\begin{cases} -x_1 - x_2 - 2x_3 + x_4 = 4 \\ -2x_1 + x_2 + 3x_3 + x_4 = 8 \\ x_1 = 0 \\ x_2 = 0 \\ x_1, x_2, x_3, x_4 \geq 0. \end{cases}$$

The solution to this system is $\hat{x} = \left(0, 0, \frac{4}{5}, \frac{28}{5}\right)$, and the optimal value is 12, which it of course has to be according to the Duality theorem. □

Download free eBooks at bookboon.com

Exercises

12.1 The matrix A and the vector c are assumed to be fixed in the LP problem

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax \geq b \end{array}$$

but the right hand side vector b is allowed to vary. Suppose that the problem has a finite value for some right hand side b. Prove that for each b, the value is either finite or there are no feasible points. Show also that the optimal value is a convex function of b.

12.2 Give an example of dual problems which both have no feasible points.

12.3 Use duality to show that $(3, 0, 1)$ is an optimal solution to the LP problem

$$\begin{array}{ll} \min & 2x_1 + 4x_2 + 3x_3 \\ \text{s.t.} & \begin{cases} 2x_1 + 3x_2 + 4x_3 \geq 10 \\ \ x_1 + 2x_2 \qquad\quad\ \geq \ 3 \\ 2x_1 + 7x_2 + 2x_3 \geq \ 5, \quad x \geq 0. \end{cases} \end{array}$$

12.4 Show that the column player's problem and the row player's problem in a two-person zero-sum game (see Chapter 9.4) are dual problems.

12.5 Investigate how the optimal solution to the LP problem

$$\begin{array}{ll} \max & x_1 + x_2 \\ \text{s.t.} & \begin{cases} tx_1 + x_2 \geq -1 \\ x_1 \qquad\ \leq \ \ 2 \\ x_1 - x_2 \geq -1 \end{cases} \end{array}$$

depends on the parameter t.

12.6 The Duality theorem follows from Farkas's lemma (Corollary 3.3.3 in Part I). Show conversely that Farkas's lemma follows from the Duality theorem by considering the dual problems

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax \geq 0 \end{array} \qquad \text{and} \qquad \begin{array}{ll} \max & \langle 0, y\rangle \\ \text{s.t.} & A^T y = c,\ y \geq 0 \end{array}$$

12.7 Let $Y = \{y \in \mathbf{R}^m \mid c - A^T y \in U,\, y \in V\}$, where U and V are closed convex cones, and suppose that $Y \neq \emptyset$.

a) Show that $\operatorname{recc} Y = \{y \in \mathbf{R}^m \mid -A^T y \in U,\, y \in V\}$.

b) Show that the system (12.3-B) has a solution if and only if the vector $-b$ does not belong to the dual cone of $\operatorname{recc} Y$.

c) Show, using the result in b), that the conclusion in case 3 of the proof of the Duality theorem follows from Theorem 12.1.1, i.e. that $v_{\max}(D) = \infty$ if (and only if) the system (12.3-B) has a solution.

Download free eBooks at bookboon.com

12.8 Suppose that the dual problems

$$\begin{array}{llcll} \min & \langle c,x\rangle & \quad\text{and}\quad & \max & \langle b,y\rangle \\ \text{s.t.} & Ax\geq b,\ x\geq 0 & & \text{s.t.} & A^Ty\leq c,\ y\geq 0 \end{array}$$

both have feasible points. Prove that there exist optimal solutions $\hat{x}$ and $\hat{y}$ to the problems that satisfy

$$\begin{cases} (A\hat{x})_i > b_i & \Leftrightarrow \quad \hat{y}_i = 0 \\ \hat{x}_j > 0 & \Leftrightarrow \quad (A^T\hat{y})_j = c_j. \end{cases}$$

[Hint: Because of the Complementarity theorem it suffices to show that the following system of inequalities has a solution: $Ax \geq b$, $x \geq 0$, $A^Ty \leq c$, $y \geq 0$, $\langle b,y\rangle \geq \langle c,x\rangle$, $Ax + y > b$, $Ay - c < x$. And this system is solvable if and only if the following homogeneous system is solvable: $Ax - bt \geq 0$, $x \geq 0$, $-A^Ty + ct \geq 0$, $y \geq 0$, $-\langle c,x\rangle + \langle b,y\rangle \leq 0$, $Ax + y - bt > 0$, $x - A^Ty + ct > 0$, $t > 0$. The solvability can now be decided by using Theorem 3.3.7 in Part I.]

Download free eBooks at bookboon.com
Click on the ad to read more

Chapter 13

The simplex algorithm

For practical purposes, there are somewhat simplified two kinds of methods for solving LP problems. Both generate a sequence of points with progressively better objective function values. Simplex methods, which were introduced by Dantzig in the late 1940s, generate a sequence of extreme points of the polyhedron of feasible points in the primal (or dual) problem by moving along the edges of the polyhedron. Interior-point methods generate instead, as the name implies, points in the interior of the polyhedron. These methods are derived from techniques for non-linear programming, developed by Fiacco and McCormick in the 1960s, but it was only after Karmarkars innovative analysis in 1984 that the methods began to be used for LP problems.

In this chapter, we describe and analyze the simplex algorithm.

13.1 Standard form

The simplex algorithm requires that the LP problem is formulated in a special way, and the variant of the algorithm that we will study assumes that the problem is a minimization problem, that all variables are nonnegative and that all other constraints are formulated as equalities.

Definition. An LP problem has *standard form* if it has the form

$$\begin{array}{ll} \min & c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ \text{s.t.} & \begin{cases} a_{11}x_1 + \ a_{12}x_2 + \cdots + \ a_{1n}x_n = b_1 \\ \qquad\qquad\qquad\qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \\ \qquad\qquad\qquad x_1, x_2, \ldots, x_n \geq \ 0. \end{cases} \end{array}$$

Download free eBooks at bookboon.com

By introducing the matrices

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad \text{and} \quad c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

we get the following compact writing for an LP problem in standard form:

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b,\ x \geq 0. \end{array}$$

We noted in Chapter 9 that each LP problem can be transformed into an equivalent LP problem in standard form by using slack/surplus variables and by replacing unrestricted variables with differences of nonnegative variables.

Duality

We gave a general definition of the concept of duality in Chapter 12.2 and showed that dual LP problems have the same optimal value, except when both problems have no feasible points. In our description of the simplex algorithm, we will need a special case of duality, and to make the presentation independent of the results in the previous chapter, we now repeat the definition for this special case.

Definition. The LP problem

(D)
$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & A^T y \leq c \end{array}$$

is said to be *dual* to the LP problem

(P)
$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b,\ x \geq 0. \end{array}$$

We shall use the following trivial part of the Duality theorem.

Theorem 13.1.1 (Weak duality). *If x is a feasible point for the minimization problem (P) and y is a feasible point for the dual maximization problem (D), i.e. if $Ax = b$, $x \geq 0$ and $A^T y \leq c$, then*

$$\langle b, y\rangle \leq \langle c, x\rangle.$$

Proof. The inequalities $A^T y \leq c$ and $x \geq 0$ imply that att $\langle x, A^T y\rangle \leq \langle x, c\rangle$, and hence

$$\langle b, y\rangle = \langle Ax, y\rangle = \langle x, A^T y\rangle \leq \langle x, c\rangle = \langle c, x\rangle. \qquad \square$$

Download free eBooks at bookboon.com

Corollary 13.1.2 (Optimality criterion). *Suppose that $\hat{x}$ is a feasible point for the minimization problem (P), that $\hat{y}$ is a feasible point for the dual maximization problem (D), and that $\langle c, \hat{x}\rangle = \langle b, \hat{y}\rangle$. Then $\hat{x}$ and $\hat{y}$ are optimal solutions to the respective problems.*

Proof. It follows from the assumptions and Theorem 13.1.1, applied to the point $\overline{y}$ and an arbitrary feasible point x for the minimization problem, that

$$\langle c, \overline{x}\rangle = \langle b, \overline{y}\rangle \leq \langle c, x\rangle$$

for all feasible points x. This shows that $\overline{x}$ is a minimum point, and an analogous argument shows that $\overline{y}$ is a maximum point. □

13.2 Informal description of the simplex algorithm

In this section we describe the main features of the simplex algorithm with the help of some simple examples. The precise formulation of the algorithm and the proof that it works is given in sections 13.4 and 13.5.

Download free eBooks at bookboon.com
Click on the ad to read more

Example 13.2.1. We start with a completely trivial problem, namely

$$\begin{aligned} &\min \ f(x) = x_3 + 2x_4 \\ &\text{s.t.} \begin{cases} x_1 \qquad + 2x_3 - x_4 = 2 \\ \qquad x_2 - \ \ x_3 + x_4 = 3, \ \ x \geq 0. \end{cases} \end{aligned}$$

Since the coefficients of the objective function $f(x)$ are positive and $x \geq 0$, it is clear that $f(x) \geq 0$ for all feasible points x. There is also a feasible point x with $x_3 = x_4 = 0$, namely $x = (2, 3, 0, 0)$. The minimum is therefore equal to 0, and $(2, 3, 0, 0)$ is the (unique) minimum point. □

Now consider an arbitrary problem of the form

$$\begin{aligned} &\min \ f(x) = c_{m+1}x_{m+1} + \cdots + c_n x_n + d \\ &\text{s.t.} \begin{cases} x_1 + \ a_{1\,m+1}x_{m+1} + \ldots + \ a_{1n}x_n = b_1 \\ x_2 + \ a_{2\,m+1}x_{m+1} + \ldots + \ a_{2n}x_n = b_2 \\ \qquad \vdots \\ x_m + a_{m\,m+1}x_{m+1} + \ldots + a_{mn}x_n = b_m, \ \ x \geq 0 \end{cases} \end{aligned} \tag{13.1}$$

where

$$b_1, b_2, \ldots, b_m \geq 0.$$

If $c_{m+1}, c_{m+2}, \ldots, c_n \geq 0$, then obviously $f(x) \geq d$ for all feasible points x, and since $\overline{x} = (b_1, \ldots, b_m, 0, \ldots, 0)$ is a feasible point and $f(\overline{x}) = d$, it follows that d is the optimal value.

The constraint system in LP problem (13.1) has a very special form, for it is solved with respect to the basic variables $x_1, x_2, \ldots, x_m$, and these variables are not present in the objective function. Quite generally, we shall call a set of variables *basic* to a given system of linear equations if it is possible to solve the system with respect to the variables in the set.

Example 13.2.2. Let us alter the objective function in Example 13.2.1 by changing the sign of the x_3-coefficient. Our new problem thus reads as follows:

$$\begin{aligned} &\min \ f(x) = -x_3 + 2x_4 \\ &\text{s.t.} \begin{cases} x_1 \qquad + 2x_3 - x_4 = 2 \\ \qquad x_2 - \ \ x_3 + x_4 = 3, \ \ x \geq 0. \end{cases} \end{aligned} \tag{13.2}$$

The point $(2, 3, 0, 0)$ is of course still feasible and the corresponding value of the objective function is 0, but we can get a smaller value by choosing $x_3 > 0$ and keeping $x_4 = 0$. However, we must ensure that $x_1 \geq 0$ and

Download free eBooks at bookboon.com

$x_2 \geq 0$, so the first constraint equation yields the bound $x_1 = 2 - 2x_3 \geq 0$, i.e. $x_3 \leq 1$.

We now transform the problem by solving the system (13.2) with respect to the variables x_2 and x_3, i.e. by changing basic variables from x_1, x_2 to x_2, x_3. Using Gaussian elimination, we obtain

$$\begin{cases} \frac{1}{2}x_1 \qquad + x_3 - \frac{1}{2}x_4 = 1 \\ \frac{1}{2}x_1 + x_2 \qquad + \frac{1}{2}x_4 = 4. \end{cases}$$

The new basic variable x_3 is then eliminated from the objectiv function by using the first equation in the new system. This results in

$$f(x) = \tfrac{1}{2}x_1 + \tfrac{3}{2}x_4 - 1,$$

and our problem has thus been reduced to a problem of the form (13.1), namely

$$\begin{array}{ll} \min & \frac{1}{2}x_1 + \frac{3}{2}x_4 - 1 \\ \text{s.t.} & \begin{cases} \frac{1}{2}x_1 \qquad + x_3 - \frac{1}{2}x_4 = 1 \\ \frac{1}{2}x_1 + x_2 \qquad + \frac{1}{2}x_4 = 4, \ x \geq 0 \end{cases} \end{array}$$

with x_2 and x_3 as basic variables and with nonnegative coefficients for the other variables in the objectiv function. Hence, the optimal value is equal to -1, and $(0, 4, 1, 0)$ is the optimal point. □

The strategy for solving a problem of the form (13.1), where some coefficient c_{m+k} is negative, consists in replacing one of the basic variables $x_1, x_2, \dots, x_m$ with x_{m+k} so as to obtain a new problem of the same form. If the new c-coefficients are nonnegative, then we are done. If not, we have to repeat the procedure. We illustrate with another example.

Example 13.2.3. Consider the problem

$$\begin{array}{ll} \min & f(x) = 2x_1 - x_2 + x_3 - 3x_4 + x_5 \\ \text{s.t.} & \begin{cases} x_1 \qquad\quad + 2x_4 - \ x_5 = 5 \\ \quad x_2 \qquad + \ x_4 + 3x_5 = 4 \\ \qquad\quad x_3 - \ x_4 + \ x_5 = 3, \ x \geq 0. \end{cases} \end{array} \tag{13.3}$$

First we have to eliminate the basic variables x_1, x_2, x_3 from the objective function with

$$f(x) = -5x_4 + 5x_5 + 9 \tag{13.4}$$

as result. Since the coefficient of x_4 is negative, x_4 has to be eliminated from the objective function and from two constraint equations in such a way

Download free eBooks at bookboon.com

that the right hand side of the transformed system remains nonnegative. The third equation in (13.3) can not be used for this elimination, since the coefficient of x_4 is negative. Eliminating x_4 from the first equation by using the second equation results in the equation $x_1 - 2x_2 - 7x_5 = 5 - 2 \cdot 4 = -3$, which has an illegal right-hand side. It therefore only remains to use the first of the constraints in (13.3) for the elimination. We then get the following equivalent system

$$\text{(13.5)} \qquad \begin{cases} \frac{1}{2}x_1 & & & + x_4 - \frac{1}{2}x_5 = \frac{5}{2} \\ -\frac{1}{2}x_1 + x_2 & & & + \frac{7}{2}x_5 = \frac{3}{2} \\ \frac{1}{2}x_1 & & + x_3 & + \frac{1}{2}x_5 = \frac{11}{2}, \quad x \geq 0 \end{cases}$$

with x_2, x_3, x_4 as new basic variables.

The reason why the right-hand side of the system remains positive when the first equation of (13.3) is used for the elimination of x_4, is that the ratio of the right-hand side and the x_4-coefficient is smaller for the first equation than for the second ($5/2 < 4/1$).

We now eliminate x_4 from the objective function, using equation (13.4) and the first equation of the system (13.5), and obtain

$$f(x) = \tfrac{5}{2}x_1 + \tfrac{5}{2}x_5 - \tfrac{7}{2}$$

Download free eBooks at bookboon.com
Click on the ad to read more

which is to be minimized under the constraints (13.5). The minimum value is clearly equal to $-\frac{7}{2}$, and $(0, \frac{3}{2}, \frac{11}{2}, \frac{5}{2}, 0)$ is the minimum point.

To reduce the writing it is customary to omit the variables and only work with coefficients in tabular form. The problem (13.3) is thus represented by the following *simplex tableau:*

$$\begin{array}{rrrrr|c}
1 & 0 & 0 & 2 & -1 & 5 \\
0 & 1 & 0 & 1 & 3 & 4 \\
0 & 0 & 1 & -1 & 1 & 3 \\
\hline
2 & -1 & 1 & -3 & 1 & f
\end{array}$$

The upper part of the tableau represents the system of equations, and the lower row represents the objective function f. The vertical line corresponds to the equality signs in (13.3).

To eliminate the basic variables x_1, x_2, x_3 from the objective function we just have to add -2 times row 1, row 2 and -1 times row 3 to the objective function row in the above tableau. This gives us the new tableau

$$\begin{array}{rrrrr|c}
1 & 0 & 0 & \underline{2} & -1 & 5 \\
0 & 1 & 0 & 1 & 3 & 4 \\
0 & 0 & 1 & -1 & 1 & 3 \\
\hline
\underline{0} & \underline{0} & \underline{0} & -5 & 5 & f-9
\end{array}$$

The bottom row corresponds to equation (13.4). Note that the constant term 9 appears on the other side of the equality sign compared to (13.4), and this explains the minus sign in the tableau. We have also highlighted the basic variables by underscoring.

Since the x_4-coefficient of the objective function is negative, we have to transform the tableau in such a way that x_4 becomes a new basic variable. By comparing the ratios $5/2$ and $4/1$ we conclude that the first row has to be the *pivot row*, i.e. has to be used for the eliminations. We have indicated this by underscoring the coefficient in the first row and the fourth column of the tableau, the so-called *pivot element*.

Gaussian elimination gives rise to the new simplex tableau

$$\begin{array}{rrrrr|c}
\frac{1}{2} & 0 & 0 & 1 & -\frac{1}{2} & \frac{5}{2} \\
-\frac{1}{2} & 1 & 0 & 0 & \frac{7}{2} & \frac{3}{2} \\
\frac{1}{2} & 0 & 1 & 0 & \frac{1}{2} & \frac{11}{2} \\
\hline
\frac{5}{2} & \underline{0} & \underline{0} & \underline{0} & \frac{5}{2} & f+\frac{7}{2}
\end{array}$$

Since the coefficients of the objective function are now nonnegative, we can read the minimum, with *reversed sign*, in the lower right corner of the tableau.

Download free eBooks at bookboon.com

The minimum point is obtained by assigning the value 0 to the non-basic variables x_1 and x_5, which gives $x = (0, \frac{3}{2}, \frac{11}{2}, \frac{5}{2}, 0)$. $\square$

EXAMPLE 13.2.4. Let us solve the LP problem

$$\begin{array}{ll} \min & x_1 - 2x_2 + x_3 \\ \text{s.t.} & \begin{cases} x_1 + 2x_2 + 2x_3 + x_4 & = 5 \\ x_1 \qquad\quad + \ x_3 \qquad + x_5 & = 2 \\ \qquad x_2 - 2x_3 \qquad\qquad + x_6 & = 1, \ x \geq 0. \end{cases} \end{array}$$

The corresponding simplex tableau is

$$\begin{array}{cccccc|c} 1 & 2 & 2 & 1 & 0 & 0 & 5 \\ 1 & 0 & 1 & 0 & 1 & 0 & 2 \\ 0 & \underline{1} & -2 & 0 & 0 & 1 & 1 \\ \hline 1 & -2 & 1 & \underline{0} & \underline{0} & \underline{0} & f \end{array}$$

with x_4, x_5, x_6 as basic variables, and these are already eliminated from the objective function. Since the x_2-coefficient of the objective function is negative, we have to introduce x_2 as a new basic variable, and we have to use the underscored element as pivot element, since $1/1 < 5/2$. Using the third row, the tableau is transformed into

$$\begin{array}{cccccc|c} 1 & 0 & \underline{6} & 1 & 0 & -2 & 3 \\ 1 & 0 & 1 & 0 & 1 & 0 & 2 \\ 0 & 1 & -2 & 0 & 0 & 1 & 1 \\ \hline 1 & \underline{0} & -3 & \underline{0} & \underline{0} & 2 & f+2 \end{array}$$

and this tableau corresponds to the problem

$$\begin{array}{ll} \min & x_1 - 3x_3 + 2x_6 - 2 \\ \text{s.t.} & \begin{cases} x_1 \qquad + 6x_3 + x_4 \qquad - 2x_6 & = 3 \\ x_1 \qquad + \ \ x_3 \qquad + x_5 & = 2 \\ \qquad x_2 - 2x_3 \qquad\qquad + \ \ x_6 & = 1, \ x \geq 0. \end{cases} \end{array}$$

Since the x_3-coefficient in the objective function is now negative, we have to repeat the procedure. We must thus introduce x_3 as a new basic variable, and this time we have to use the first row as pivot row, for $3/6 < 2/1$. The new tableau has the following form

$$\begin{array}{cccccc|c} \frac{1}{6} & 0 & 1 & \frac{1}{6} & 0 & -\frac{1}{3} & \frac{1}{2} \\ \frac{5}{6} & 0 & 0 & -\frac{1}{6} & 1 & \frac{1}{3} & \frac{3}{2} \\ \frac{1}{3} & 1 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 2 \\ \hline \frac{3}{2} & \underline{0} & \underline{0} & \frac{1}{2} & \underline{0} & 1 & f+\frac{7}{2} \end{array}$$

Download free eBooks at bookboon.com

We can now read off the minimum $-\frac{7}{2}$ and the minimum point $(0, 2, \frac{1}{2}, 0, \frac{3}{2}, 0)$. □

So far, we have written the function symbol f in the lower right corner of our simplex tableaux. We have done this for pedagogical reasons to explain why the function value in the box gets a reverse sign. Remember that the last row of the previous simplex tableau means that

$$\tfrac{3}{2}x_1 + \tfrac{1}{2}x_4 + x_6 = f(x) + \tfrac{7}{2}.$$

Since the symbol has no other function, we will omit it in the future.

EXAMPLE 13.2.5. The problem

$$\begin{array}{ll} \min & f(x) = -2x_1 + x_2 \\ \text{s.t.} & \begin{cases} x_1 - x_2 + x_3 = 3 \\ -x_1 + x_2 + x_4 = 4, \ \ x \geq 0 \end{cases} \end{array}$$

gives rise to the following simplex tableaux:

$$\begin{array}{rrrr|r} \underline{1} & -1 & 1 & 0 & 3 \\ -1 & 1 & 0 & 1 & 4 \\ \hline -2 & 1 & \underline{0} & \underline{0} & 0 \end{array}$$

$$\begin{array}{rrrr|r} 1 & -1 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 & 7 \\ \hline \underline{0} & -1 & 2 & \underline{0} & 6 \end{array}$$

Since the objective function has a negative x_2-coefficient, we are now supposed to introduce x_2 as a basic variable, but no row will work as a pivot row since the entire x_2-column is non-positive. This implies that the objective function is unbounded below, i.e. there is no minimum. To see this, we rewrite the last tableau with variables in the form

$$\begin{array}{ll} \min & f(x) = -x_2 + 2x_3 - 6 \\ \text{s.t.} & \begin{cases} x_1 = x_2 - x_3 + 3 \\ x_4 = - x_3 + 7. \end{cases} \end{array}$$

By choosing $x_2 = t$ and $x_3 = 0$ we get a feasible point $x^t = (3 + t, t, 0, 7)$ for each $t \geq 0$, and since $f(x^t) = -t - 6 \to -\infty$ as $t \to \infty$, we conclude that the objective function is unbounded below. □

Examples 13.2.4 and 13.2.5 are typical for LP problems of the form (13.1). In Section 13.5, namely, we show that one can always perform the iterations

Download free eBooks at bookboon.com

so as to obtain a final tableau similar to the one in Example 13.2.4 or the one in Example 13.2.5, and in Section 13.6 we will show how to get started, i.e. how to transform an arbitrary LP problem in standard form into a problem of the form (13.1).

13.3 Basic solutions

In order to describe and understand the simplex algorithm it is necessary first to know how to produce solutions to a linear system of equations. We assume that Gaussian elimination is familiar and concentrate on describing how to switch from one basic solution to another. We begin by reviewing the notation that we will use in the rest of this chapter.

The columns of an $m \times n$-matrix A will be denoted $A_{*1}, A_{*2}, \dots, A_{*n}$ so that

$$A = \begin{bmatrix} A_{*1} & A_{*2} & \dots & A_{*n} \end{bmatrix}.$$

We will often have to consider submatrices comprised of certain columns in an $m \times n$-matrix A. So if $1 \le k \le n$ and

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)$$

Download free eBooks at bookboon.com
Click on the ad to read more

is a permutation of k elements chosen from the set $\{1, 2, \dots, n\}$, we let $A_{*\alpha}$ denote the $m \times k$-matrix consisting of the columns $A_{*\alpha_1}$, $A_{*\alpha_2}$, ..., $A_{*\alpha_k}$ in the matrix A, i.e.

$$A_{*\alpha} = \begin{bmatrix} A_{*\alpha_1} & A_{*\alpha_2} & \dots & A_{*\alpha_k} \end{bmatrix}.$$

And if

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

is a column matrix with n entries, then x_α denotes the column matrix

$$\begin{bmatrix} x_{\alpha_1} \\ x_{\alpha_2} \\ \vdots \\ x_{\alpha_k} \end{bmatrix}.$$

As usual, we make no distinction between column matrices with n entries and vectors in $\mathbf{R}^n$.

We consider permutations $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)$ as ordered sets and allow us therefore to write $j \in \alpha$ if j is any of the numbers $\alpha_1, \alpha_2, \dots, \alpha_k$. This also allows us to write sums of the type

$$\sum_{i=1}^{k} x_{\alpha_i} A_{*\alpha_i}$$

as

$$\sum_{j\in\alpha} x_j A_{*j},$$

or with matrices as

$$A_{*\alpha} x_\alpha.$$

Definition. Let A be an $m \times n$-matrix of rank m, and let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)$ be a permutation of m numbers from the set $\{1, 2, \dots, n\}$. The permutation α is called a *basic index set* of the matrix A if the columns of the $m \times m$-matrix $A_{*\alpha}$ form a basis for $\mathbf{R}^m$.

The condition that the columns $A_{*\alpha_1}, A_{*\alpha_2}, \dots, A_{*\alpha_m}$ form a basis is equivalent to the condition that the submatrix

$$A_{*\alpha} = \begin{bmatrix} A_{*\alpha_1} & A_{*\alpha_2} & \dots & A_{*\alpha_m} \end{bmatrix}$$

Download free eBooks at bookboon.com

is invertible. The inverse of the matrix $A_{*\alpha}$ will be denoted by $A_{*\alpha}^{-1}$. This matrix, which thus means $(A_{*\alpha})^{-1}$, will appear frequently in the sequel – do not confuse it with $(A^{-1})_{*\alpha}$, which is not generally well defined.

If $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)$ is a basic index set, so too is of course every permutation of α.

Example 13.3.1. The matrix

$$\begin{bmatrix} 3 & 1 & 1 & -3 \\ 3 & -1 & 2 & -6 \end{bmatrix}$$

has the following basic index sets: $(1,2)$, $(2,1)$, $(1,3)$, $(3,1)$, $(1,4)$, $(4,1)$ $(2,3)$, $(3,2)$, $(2,4)$, and $(4,2)$. □

We also need a convenient way to show the result of replacing an element in an ordered set with some other element. Therefore, let $M = (a_1, a_2, \dots, a_n)$ be an arbitrary n-tuple (ordered set). The n-tuple obtained by replacing the item a_r at location r with an arbitrary object x will be denoted by $M_{\hat{r}}[x]$. In other words,

$$M_{\hat{r}}[x] = (a_1, \dots, a_{r-1}, x, a_{r+1}, \dots, a_n).$$

An $m \times n$-matrix can be regarded as an ordered set of columns. If b is a column matrix with m entries and $1 \le r \le n$, we therefore write $A_{\hat{r}}[b]$ for the matrix

$$\begin{bmatrix} A_{*1} & \dots & A_{*r-1} & b & A_{*r+1} & \dots & A_{*n} \end{bmatrix}.$$

Another context in which we will use the above notation for replacement of elements, is when $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)$ is a permutation of m elements taken from the set $\{1, 2, \dots, n\}$. If $1 \le r \le m$, $1 \le k \le n$ and $k \notin \alpha$, then $\alpha_{\hat{r}}[k]$ denotes the new permutation

$$(\alpha_1, \dots, \alpha_{r-1}, k, \alpha_{r+1}, \dots, \alpha_m).$$

Later we will need the following simple result, where the above notation is used.

Lemma 13.3.1. *Let E be the unit matrix of order m, and let b be a column matrix with m elements. The matrix $E_{\hat{r}}[b]$ is invertible if and only if $b_r \neq 0$, and in this case*

$$E_{\hat{r}}[b]^{-1} = E_{\hat{r}}[c],$$

where

$$c_j = \begin{cases} -b_j/b_r & \text{for } j \neq r, \\ 1/b_r & \text{for } j = r. \end{cases}$$

Download free eBooks at bookboon.com

Proof. The proof is left to the reader as a simple exercise. □

EXAMPLE 13.3.2.

$$\begin{bmatrix} 1 & 4 & 0 \\ 0 & 3 & 0 \\ 0 & 5 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -4/3 & 0 \\ 0 & 1/3 & 0 \\ 0 & -5/3 & 1 \end{bmatrix}$$ □

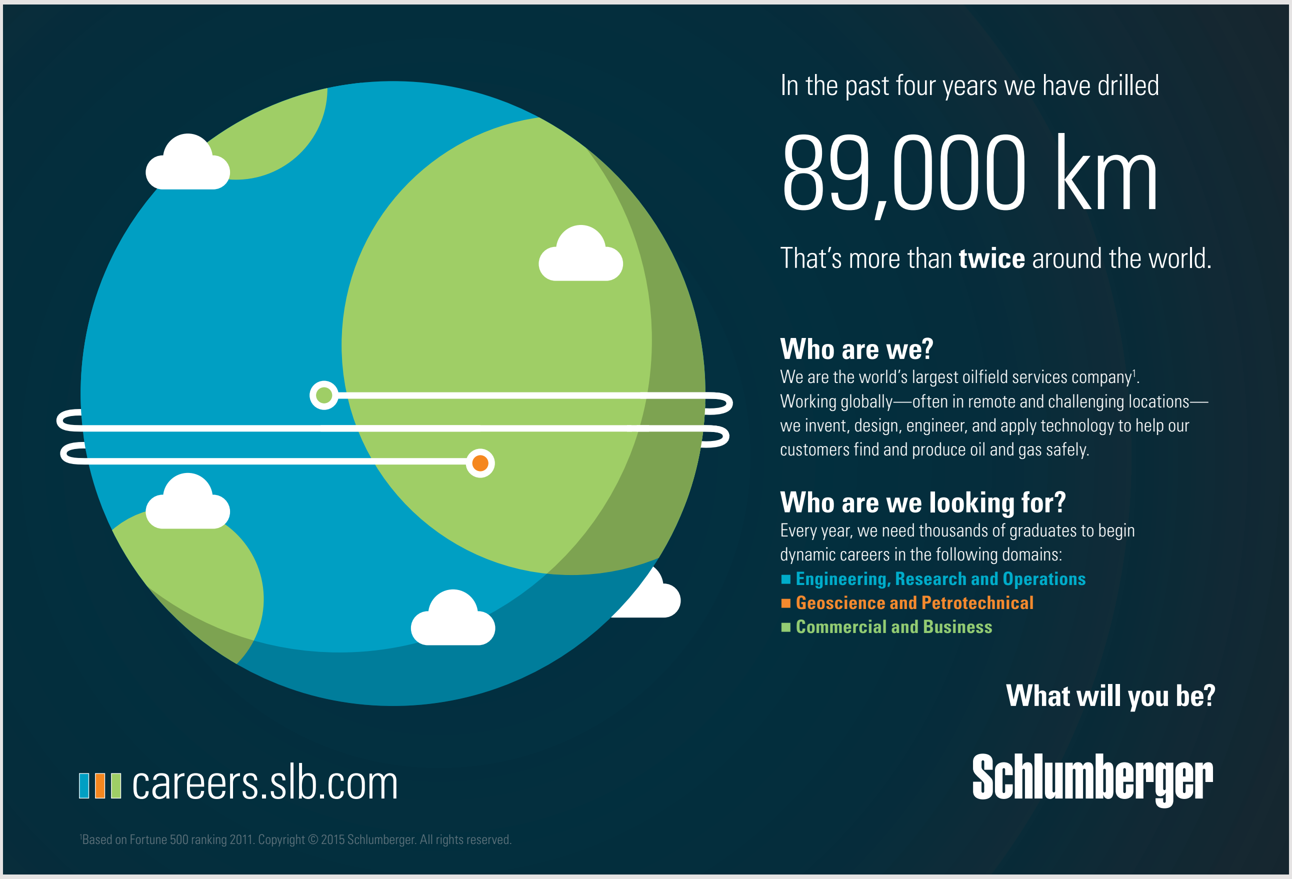
Systems of linear equations and basic solutions

Consider a system of linear equations

$$\text{(13.6)} \qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

with coefficient matrix A of rank m and right-hand side matrix b. Such a system can equivalently be regarded as a vector equation

$$\text{(13.6}'\text{)} \qquad \sum_{j=1}^{n} x_j A_{*j} = b$$



Download free eBooks at bookboon.com

Click on the ad to read more

or as a matrix equation

$$(13.6'') \qquad Ax = b.$$

Both alternative approaches are, as we shall see, fruitful.

We solve the system (13.6), preferably using Gaussian elimination, by expressing m of the variables, $x_{\alpha_1}, x_{\alpha_2}, \dots, x_{\alpha_m}$ say, as linear combinations of the remaining $n-m$ variables $x_{\beta_1}, x_{\beta_2}, \dots, x_{\beta_{n-m}}$ and $b_1, b_2, \dots, b_m$. Each assignment of values to the latter β-variables results in a unique set of values for the former α-variables. In particular, we get a unique solution by setting all β-variables equal to 0.

This motivates the following definition.

Definition. Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)$ be a permutation of m numbers chosen from the set $\{1, 2, \dots, n\}$, and let $\beta = (\beta_1, \beta_2, \dots, \beta_{n-m})$ be a permutation of the remaining $n-m$ numbers. The variables $x_{\alpha_1}, x_{\alpha_2}, \dots, x_{\alpha_m}$ are called *basic variables* and the variables $x_{\beta_1}, x_{\beta_2}, \dots, x_{\beta_{n-m}}$ are called *free variables* in the system (13.6), if for each $c = (c_1, c_2, \dots, c_{n-m}) \in \mathbf{R}^{n-m}$ there is a unique solution x to the system (13.6) such that $x_\beta = c$. The unique solution obtained by setting all free variables equal to 0 is called a *basic solution.*

Any m variables can not be chosen as basic variables; to examine which ones can be selected, let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)$ be a permutation of m numbers from the set $\{1, 2, \dots, n\}$ and let $\beta = (\beta_1, \beta_2, \dots, \beta_{n-m})$ be an arbitrary permutation of the remaining $n-m$ numbers, and rewrite equation (13.6′) as

$$(13.6''') \qquad \sum_{j=1}^{m} x_{\alpha_j} A_{*\alpha_j} = b - \sum_{j=1}^{n-m} x_{\beta_j} A_{*\beta_j}.$$

If α is a basic index set, i.e. if the columns $A_{*\alpha_1}, A_{*\alpha_2}, \dots, A_{*\alpha_m}$ form a basis of $\mathbf{R}^m$, then equation (13.6‴) has clearly a unique solution for each assignment of values to the β-variables, and $(x_{\alpha_1}, x_{\alpha_2}, \dots, x_{\alpha_m})$ is in fact the coordinates of the vector $b - \sum_{j=1}^{n-m} x_{\beta_j} A_{*\beta_j}$ in this basis. In particular, the coordinates of the vector b are equal to $(\overline{x}_{\alpha_1}, \overline{x}_{\alpha_2}, \dots, \overline{x}_{\alpha_m})$, where $\overline{x}$ is the corresponding basic solution, defined by the condition that $\overline{x}_{\beta_j} = 0$ for all j.

Conversely, suppose that each assignment of values to the β-variables determines uniquely the values of the α-variables. In particular, the equation

$$(13.7) \qquad \sum_{j=1}^{m} x_{\alpha_j} A_{*\alpha_j} = b$$

Download free eBooks at bookboon.com

has then a unique solution, and this implies that the equation

$$\sum_{j=1}^{m} x_{\alpha_j} A_{*\alpha_j} = 0 \tag{13.8}$$

has no other solution then the trivial one, $x_{\alpha_j} = 0$ for all j, because we would otherwise get several solutions to equation (13.7) by to a given one adding a non-trivial solution to equation (13.8). The columns $A_{*\alpha_1}, A_{*\alpha_2}, \dots, A_{*\alpha_m}$ are in other words linearly independent, and they form a basis for $\mathbf{R}^m$ since they are m in number. Hence, α is a basic index set.

In summary, we have proved the following result.

Theorem 13.3.2. *The variables $x_{\alpha_1}, x_{\alpha_2}, \dots, x_{\alpha_m}$ are basic variables in the system* (13.6) *if and only if α is a basic index set of the coefficient matrix A.*

Let us now express the basic solution corresponding to the basic index set α in matrix form. By writing the matrix equation (13.6″) in the form

$$A_{*\alpha} x_\alpha + A_{*\beta} x_\beta = b$$

and multiplying from the left by the matrix $A_{*\alpha}^{-1}$, we get

$$\begin{aligned} x_\alpha + A_{*\alpha}^{-1} A_{*\beta} x_\beta &= A_{*\alpha}^{-1} b, \qquad \text{i.e.} \\ x_\alpha &= A_{*\alpha}^{-1} b - A_{*\alpha}^{-1} A_{*\beta} x_\beta, \end{aligned}$$

which expresses the basic variables as linear combinations of the free variables and the coordinates of b. The basic solution is obtained by setting $x_\beta = 0$ and is given by

$$\overline{x}_\alpha = A_{*\alpha}^{-1} b\,, \;\; \overline{x}_\beta = 0.$$

We summarize this result in the following theorem.

Theorem 13.3.3. *Let α be a basic index set of the matrix A. The corresponding basic solution $\overline{x}$ to the system $Ax = b$ is given by the conditions*

$$\overline{x}_\alpha = A_{*\alpha}^{-1} b \quad \textit{and} \quad \overline{x}_k = 0 \quad \textit{for } k \notin \alpha.$$

The $n-m$ free variables in a basic solution are equal to zero by definition. Of course, some basic variable may also happen to be equal to zero, and since this results in certain complications for the simplex algorithm, we make the following definition.

Definition. A basic solution $\overline{x}$ is called *non-degenerate* if $\overline{x}_i \neq 0$ for m indices i and *degenerate* if $\overline{x}_i \neq 0$ for less than m indices i.

Download free eBooks at bookboon.com

Two basic index sets α and α', which are permutations of each other, naturally give rise to the same basic solution $\overline{x}$. So the number of different basic solutions to a system $Ax = b$ with m equations and n unknowns is at most equal to the number of subsets with m elements that can be chosen from the set $\{1, 2, \dots, n\}$, i.e. at most equal to $\binom{n}{m}$. The number is smaller if the matrix A contains m linearly dependent columns.

EXAMPLE 13.3.3. The system

$$\begin{cases} 3x_1 + x_2 + x_3 - 3x_4 = 3 \\ 3x_1 - x_2 + 2x_3 - 6x_4 = 3 \end{cases}$$

has – apart from permutations – the following basic index sets: $(1, 2)$, $(1, 3)$, $(1, 4)$, $(2, 3)$ and $(2, 4)$, and the corresponding basic solutions are in turn $(1, 0, 0, 0)$, $(1, 0, 0, 0)$, $(1, 0, 0, 0)$, $(0, 1, 2, 0)$ and $(0, 1, 0, -\frac{2}{3})$. The basic solution $(1, 0, 0, 0)$ is degenerate, and the other two basic solutions are non-degenerate. □

The reason for our interest in basic index sets and basic solutions is that optimal values of LP problems are attained at extreme points, and these points are basic solutions, because we have the following characterisation of extreme points.

Download free eBooks at bookboon.com
Click on the ad to read more

Theorem 13.3.4. *Suppose that A is an $m \times n$-matrix of rank m, that $b \in \mathbf{R}^m$ and that $c \in \mathbf{R}^n$. Then:*

(i) $\overline{x}$ is an extreme point of the polyhedron $X = \{x \in \mathbf{R}^n \mid Ax = b,\ x \geq 0\}$ if and only if $\overline{x}$ is a nonnegative basic solution to the system $Ax = b$, i.e. if and only if there is a basic index set α of the matrix A such that $\overline{x}_\alpha = A_{*\alpha}^{-1} b \geq 0$ and $\overline{x}_k = 0$ for $k \notin \alpha$.

(ii) $\overline{y}$ is an extreme point of the polyhedron $Y = \{y \in \mathbf{R}^m \mid A^T y \leq c\}$ if and only if $A^T \overline{y} \leq c$ and there is a basic index set α of the matrix A such that $\overline{y} = (A_{*\alpha}^{-1})^T c_\alpha$.

Proof. (i) According to Theorem 5.1.1 in Part I, $\overline{x}$ is an extreme point of the polyhedron X if and only if $\overline{x} \geq 0$ and $\overline{x}$ is the unique solution of a system of linear equations consisting of the equation $Ax = b$ and $n - m$ equations out of the n equations $x_1 = 0$, $x_2 = 0$, $\dots$, $x_n = 0$. Let $\alpha_1, \alpha_2, \dots, \alpha_m$ be the indices of the m equations $x_i = 0$ that are not used in this system. Then, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)$ is a basic index set and $\overline{x}$ is the corresponding basic solution.

(ii) Because of the same theorem, $\overline{y}$ is an extreme point of the polyhedron Y if and only if $\overline{y} \in Y$ and $\overline{y}$ is the unique solution of a quadratic system of linear equations obtained by selecting m out of the n equations in the system $A^T y = c$. Let $\alpha_1, \alpha_2, \dots, \alpha_m$ denote the indices of the selected equations. The quadratic system is then of the form $(A_{*\alpha})^T y = c_\alpha$, and this system of equations has a unique solution $\overline{y} = (A_{*\alpha}^{-1})^T c_\alpha$ if and only if $A_{*\alpha}$ is an invertible matrix, i.e. if and only if $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)$ is a basic index set of A. □

Example 13.3.4. It follows from Theorem 13.3.4 and Example 13.3.3 that the polyhedron X of solutions to the system

$$\begin{cases} 3x_1 + x_2 + x_3 - 3x_4 = 3 \\ 3x_1 - x_2 + 2x_3 - 6x_4 = 3,\ x \geq 0 \end{cases}$$

has two extreme points, namely $(1, 0, 0, 0)$ and $(0, 1, 2, 0)$.

The "dual" polyhedron Y of solutions to the system

$$\begin{cases} 3y_1 + 3y_2 \leq 2 \\ y_1 - y_2 \leq 1 \\ y_1 + 2y_2 \leq 1 \\ -3y_1 - 6y_2 \leq -1 \end{cases}$$

has three extreme points, namely $(\frac{5}{6}, -\frac{1}{6})$, $(\frac{1}{3}, \frac{1}{3})$ and $(\frac{7}{9}, -\frac{2}{9})$, corresponding to the basic index sets $(1, 2)$, $(1, 3)$ and $(2, 4)$. (The points associated with the other two basic index sets $(1, 4)$ and $(2, 3)$, $y = (1, -\frac{1}{3})$ and $y = (1, 0)$, respectively, are not extreme points since they lie outside Y.) □

Download free eBooks at bookboon.com

Changing basic index sets

We will now discuss how to generate a suite of basic solutions by successively replacing one element at a time in the basic index set.

Theorem 13.3.5. *Suppose that $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)$ is a basic index set of the system $Ax = b$ and let $\overline{x}$ denote the corresponding basic solution. Let k be a column index not belonging to the basic index set α, and let $v \in \mathbf{R}^n$ be the column vector defined by*

$$v_\alpha = A_{*\alpha}^{-1} A_{*k}, \quad v_k = -1 \quad \text{and} \quad v_j = 0 \quad \text{for } j \notin \alpha \cup \{k\}.$$

(i) Then $Av = 0$, so it follows that $\overline{x} - tv$ is a solution to the system $Ax = b$ for all $t \in \mathbf{R}$.

(ii) Suppose that $1 \le r \le m$ and define a new ordered set α' by replacing the element α_r in α with the number k, i.e.

$$\alpha' = \alpha_{\hat{r}}[k] = (\alpha_1, \dots, \alpha_{r-1}, k, \alpha_{r+1}, \dots, \alpha_m).$$

Then, α' is a basic index set if and only if $v_{\alpha_r} \neq 0$. In this case,

$$A_{*\alpha'}^{-1} = E_{\hat{r}}[v_\alpha]^{-1} A_{*\alpha}^{-1}$$

and if $\overline{x}'$ is the basic solution corresponding to the basic index set α', then

$$\overline{x}' = \overline{x} - \tau v,$$

where $\tau = \overline{x}_{\alpha_r} / v_{\alpha_r}$.

(iii) The two basic solutions $\overline{x}$ and $\overline{x}'$ are identical if and only if $\tau = 0$. So if $\overline{x}$ is a non-degenerate basic solution, then $\overline{x} \neq \overline{x}'$.

We will call v the *search vector* associated with the basic index set α and the index k, since we obtain the new basic solution $\overline{x}'$ from the old one $\overline{x}$ by searching in the direction of minus v.

Proof. (i) It follows immediately from the definition of v that

$$Av = \sum_{j \in \alpha} v_j A_{*j} + \sum_{j \notin \alpha} v_j A_{*j} = A_{*\alpha} v_\alpha - A_{*k} = A_{*k} - A_{*k} = 0.$$

(ii) The set α' is a basic index set if and only if $A_{*\alpha'}$ is an invertible matrix. But

$$\begin{aligned} A_{*\alpha}^{-1} A_{*\alpha'} &= A_{*\alpha}^{-1} \big[\, A_{*\alpha_1} \ \dots \ A_{*\alpha_{r-1}} \ A_{*k} \ A_{*\alpha_{r+1}} \ \dots \ A_{*\alpha_m} \,\big] \\ &= \big[\, A_{*\alpha}^{-1} A_{*\alpha_1} \ \dots \ A_{*\alpha}^{-1} A_{*\alpha_{r-1}} \ A_{*\alpha}^{-1} A_{*k} \ A_{*\alpha}^{-1} A_{*\alpha_{r+1}} \ \dots \ A_{*\alpha}^{-1} A_{*\alpha_m} \,\big] \\ &= \big[\, E_{*1} \ \dots \ E_{*r-1} \ v_\alpha \ E_{*r+1} \ \dots \ E_{*m} \,\big] = E_{\hat{r}}[v_\alpha], \end{aligned}$$

Download free eBooks at bookboon.com

where of course E denotes the unit matrix of order m. Hence

$$A_{*\alpha'} = A_{*\alpha} E_{\hat{r}}[v_\alpha].$$

The matrix $A_{*\alpha'}$ is thus invertible if and only if the matrix $E_{\hat{r}}[v_\alpha]$ is invertible, and this is the case if and only if $v_{\alpha_r} \neq 0$, according to Lemma 13.3.1. If the inverse exists, then

$$A^{-1}_{*\alpha'} = \big(A_{*\alpha} E_{\hat{r}}[v_\alpha]\big)^{-1} = E_{\hat{r}}[v_\alpha]^{-1} A^{-1}_{*\alpha}.$$

Now, define $x^\tau = \overline{x} - \tau v$. Then x^τ is a solution to the equation $Ax = b$, by part (i) of the theorem, so in order to prove that x^τ is the basic solution corresponding to the basic index set α', it suffices to show that $x^\tau_j = 0$ for all $j \notin \alpha'$, i.e. for $j = \alpha_r$ and for $j \notin \alpha \cup \{k\}$.

But $x^\tau_{\alpha_r} = \overline{x}_{\alpha_r} - \tau v_{\alpha_r} = 0$, because of the definition of τ, and if $j \notin \alpha \cup \{k\}$ then $\overline{x}_j$ and v_j are both equal to 0, whence $x^\tau_j = \overline{x}_j - \tau v_j = 0$.

(iii) Since $v_k = -1$, we have $\tau v = 0$ if and only if $\tau = 0$. Hence , $\overline{x}' = \overline{x}$ if and only if $\tau = 0$.

If the basic solution $\overline{x}$ is non-degenerate, then $\overline{x}_j \neq 0$ for all $j \in \alpha$ and in particular $\overline{x}_{\alpha_r} \neq 0$, which implies that $\tau \neq 0$, and that $\overline{x}' \neq \overline{x}$. $\square$

Click on the ad to read more

Corollary 13.3.6. *Keep the asumptions of Theorem 13.3.5 and suppose in addition that $\overline{x} \geq 0$, that the index set*

$$I_+ = \{j \in \alpha \mid v_j > 0\}$$

is nonempty, and that the index r is chosen so that $\alpha_r \in I_+$ and

$$\tau = \overline{x}_{\alpha_r}/v_{\alpha_r} = \min\{\overline{x}_j/v_j \mid j \in I_+\}.$$

Then $\overline{x}' \geq 0$.

Proof. Since $\overline{x}'_j = 0$ for all $j \notin \alpha'$, it suffices to show that $\overline{x}'_j \geq 0$ for all $j \in \alpha \cup \{k\}$.

We begin by noting that $\tau \geq 0$ since $\overline{x} \geq 0$, and therefore

$$\overline{x}'_k = \overline{x}_k - \tau v_k = 0 + \tau \geq 0.$$

For indices $j \in \alpha \setminus I_+$ we have $v_j \leq 0$, and this implies that

$$\overline{x}'_j = \overline{x}_j - \tau v_j \geq \overline{x}_j \geq 0.$$

Finally, if $j \in I_+$, then $\overline{x}_j/v_j \geq \tau$, and it follows that

$$\overline{x}'_j = \overline{x}_j - \tau v_j \geq 0.$$

This completes the proof. □

13.4 The simplex algorithm

The variant of the simplex algorithm that we shall describe assumes that the LP problem is given in standard form. So we start from the problem

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b,\ x \geq 0 \end{array}$$

where A is an $m \times n$-matrix, $b \in \mathbf{R}^m$ and $c \in \mathbf{R}^n$.

We assume that

$$\operatorname{rank} A = m = \text{the number of rows in } A.$$

Of course, this is no serious restriction, because if $\operatorname{rank} A < m$ and the system $Ax = b$ is consistent, then we can delete $(m - \operatorname{rank} A)$ constraint equations without changing the set of solutions, and this leaves us with an equivalent system $A'x = b$, where the rank of A' is equal to the number of rows in A'.

Let us call a basic index set α of the matrix A and the corresponding basic solution $\overline{x}$ to the system $Ax = b$ *feasible*, if $\overline{x}$ is a feasible point for our standard problem, i.e. if $\overline{x} \geq 0$.

Download free eBooks at bookboon.com

The simplex algorithm starts from a feasible basic index set α of the matrix A, and we shall show in Section 13.6 how to find such an index set by applying the simplex algorithm to a so-called artificial problem.

First compute the corresponding feasible basic solution $\overline{x}$, i.e.

$$\overline{x}_\alpha = A_{*\alpha}^{-1} b \geq 0,$$

and then the number $\lambda \in \mathbf{R}$ and the column vectors $\overline{y} \in \mathbf{R}^m$ and $z \in \mathbf{R}^n$, defined as

$$\begin{aligned} \lambda &= \langle c, \overline{x} \rangle = \langle c_\alpha, \overline{x}_\alpha \rangle \\ \overline{y} &= (A_{*\alpha}^{-1})^T c_\alpha \\ z &= c - A^T \overline{y}. \end{aligned}$$

The number λ is thus equal to the value of the objective function at $\overline{x}$.

Note that $z_\alpha = c_\alpha - (A^T\overline{y})_{\alpha *} = c_\alpha - (A_{*\alpha})^T\overline{y} = c_\alpha - c_\alpha = 0$, so in order to compute the vector z we only have to compute its coordinates

$$z_j = c_j - (A_{*j})^T \overline{y} = c_j - \langle A_{*j}, \overline{y} \rangle$$

for indices $j \notin \alpha$. The numbers z_j are usually called *reduced costs.*

Lemma 13.4.1. *The number λ and the vectors $\overline{x}$, $\overline{y}$ and z have the following properties:*

(i) $\langle z, \overline{x} \rangle = 0$, *i.e. the vectors z and $\overline{x}$ are orthogonal.*

(ii) $Ax = 0 \Rightarrow \langle c, x \rangle = \langle z, x \rangle$.

(iii) $Ax = b \Rightarrow \langle c, x \rangle = \lambda + \langle z, x \rangle$.

(iv) If v is the search vector corresponding to the basic index set α and the index $k \notin \alpha$, then $\langle c, \overline{x} - tv \rangle = \lambda + tz_k$.

Proof. (i) Since $z_j = 0$ for $j \in \alpha$ and $\overline{x}_j = 0$ for $j \notin \alpha$,

$$\langle z, \overline{x} \rangle = \sum_{j \in \alpha} z_j \overline{x}_j + \sum_{j \notin \alpha} z_j \overline{x}_j = 0 + 0 = 0.$$

(ii) It follows immediately from the definition of z that

$$\langle z, x \rangle = \langle c, x \rangle - \langle A^T\overline{y}, x \rangle = \langle c, x \rangle - \langle \overline{y}, Ax \rangle = \langle c, x \rangle$$

for all x satisfying the equation $Ax = 0$.

(iii) If $Ax = b$, then

$$\begin{aligned} \langle c, x \rangle - \langle z, x \rangle &= \langle A^T\overline{y}, x \rangle = \langle \overline{y}, Ax \rangle = \langle (A_{*\alpha}^{-1})^T c_\alpha, b \rangle = \langle c_\alpha, A_{*\alpha}^{-1} b \rangle \\ &= \langle c_\alpha, \overline{x}_\alpha \rangle = \lambda. \end{aligned}$$

(iv) Since $Av = 0$, it follows from (ii) that

$$\langle c, \overline{x} - tv \rangle = \langle c, \overline{x} \rangle - t\langle c, v \rangle = \lambda - t\langle z, v \rangle = \lambda + tz_k. \qquad \square$$

Download free eBooks at bookboon.com

The following theorem contains all the essential ingredients of the simplex algorithm.

Theorem 13.4.2. *Let α, $\overline{x}$, λ, $\overline{y}$ and z be defined as above.*

(i) ***(Optimality)*** *If $z \geq 0$, then $\overline{x}$ is an optimal solution to the minimization problem*

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b,\ x \geq 0 \end{array}$$

and $\overline{y}$ is an optimal solution to the dual maximization problem

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & A^T y \leq c \end{array}$$

with λ as the optimal value. The optimal solution $\overline{x}$ to the minimization problem is unique if $z_j > 0$ for all $j \notin \alpha$.

(ii) Suppose that $z \not\geq 0$, and let k be an index such that $z_k < 0$. Let further v be the search vector associated to α and k, i.e.

$$v_\alpha = A_{*\alpha}^{-1} A_{*k}, \quad v_k = -1, \quad v_j = 0 \quad \textit{for } j \notin \alpha \cup \{k\},$$

and set $x^t = \overline{x} - tv$ for $t \geq 0$. Depending on whether $v \leq 0$ or $v \not\leq 0$, the following applies:

(ii a) ***(Unbounded objective function)*** *If $v \leq 0$, then the points x^t are feasible for the minimization problem for all $t \geq 0$ and $\langle c, x^t\rangle \to -\infty$ as $t \to \infty$. The objective function is thus unbounded below, and the dual maximization problem has no feasible points.*

(ii b) ***(Iteration step)*** *If $v \not\leq 0$, then define a new basic index set α' and the number τ as in Theorem 13.3.5 (ii) with the index r chosen as in Corollary 13.3.6. The basic index set α' is feasible with $\overline{x}' = \overline{x} - \tau v$ as the corresponding feasible basic solution, and*

$$\langle c, \overline{x}'\rangle = \langle c, \overline{x}\rangle + \tau z_k \leq \langle c, \overline{x}\rangle.$$

Hence, $\langle c, \overline{x}'\rangle < \langle c, \overline{x}\rangle$, if $\tau > 0$.

Proof. (i) Suppose that $z \geq 0$ and that x is an arbitrary feasible point for the minimization problem. Then $\langle z, x\rangle \geq 0$ (since $x \geq 0$), and it follows from part (iii) of Lemma 13.4.1 that $\langle c, x\rangle \geq \lambda = \langle c, \overline{x}\rangle$. The point $\overline{x}$ is thus optimal and the optimal value is equal to λ.

The condition $z \geq 0$ also implies that $A^T\overline{y} = c - z \leq c$, i.e. $\overline{y}$ is a feasible point for the dual maximization problem, and

$$\langle b, \overline{y}\rangle = \langle \overline{y}, b\rangle = \langle (A_{*\alpha}^{-1})^T c_\alpha, b\rangle = \langle c_\alpha, A_{*\alpha}^{-1} b\rangle = \langle c_\alpha, \overline{x}_\alpha\rangle = \langle c, \overline{x}\rangle,$$

so if follows from the optimality criterion (Corollary 13.1.2) that $\overline{y}$ is an optimal solution to the dual problem.

Download free eBooks at bookboon.com

Now suppose that $z_j > 0$ for all $j \notin \alpha$. If x is a feasible point $\neq \overline{x}$, then $x_{j_0} > 0$ for some index $j_0 \notin \alpha$, and it follows that $\langle z, x\rangle = \sum_{j\notin\alpha} z_j x_j \geq z_{j_0}x_{j_0} > 0$. Hence, $\langle c, x\rangle = \lambda + \langle z, x\rangle > \lambda = \langle c, \overline{x}\rangle$, by Lemma 13.4.1 (iii). This proves that the minimum point is unique.

(ii a) According to Theorem 13.3.5, x^t is a solution to the equation $Ax = b$ for all real numbers t, and if $v \leq 0$ then $x^t = \overline{x} - tv \geq \overline{x} \geq 0$ for $t \geq 0$. So the points x^t are feasible for all $t \geq 0$ if $v \leq 0$, and by Lemma 13.4.1 (iv),

$$\lim_{t\to\infty} \langle c, x^t\rangle = \lambda + \lim_{t\to\infty} z_k t = -\infty.$$

The objective function is thus not bounded below.

Suppose that the dual maximization problem has a feasible point y. Then, $\langle b, y\rangle \leq \langle c, x^t\rangle$ for all $t \geq 0$, by the weak duality theorem, and this is contradictory since the right hand side tends to $-\infty$ as $t \to \infty$. So it follows that the dual maximization problem has no feasible points.

(ii b) By Corollary 13.3.6, α' is a feasible basic solution with x^τ as the corresponding basic solution, and the inequality $\langle c, \overline{x}'\rangle \leq \langle c, \overline{x}\rangle$ now follows directly from Lemma 13.4.1 (iv), because $\tau \geq 0$. □

Click on the ad to read more

Theorem 13.4.2 gives rise to the following algorithm for solving the standard problem

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b,\ x \geq 0. \end{array}$$

The simplex algorithm

Given a feasible basic index set α.

1. Compute the matrix $A_{*\alpha}^{-1}$, the corresponding feasible basic solution $\overline{x}$, i.e. $\overline{x}_\alpha = A_{*\alpha}^{-1}b$ and $\overline{x}_j = 0$ for $j \notin \alpha$, and the number $\lambda = \langle c_\alpha, \overline{x}_\alpha\rangle$.

Repeat steps 2–8 until a stop occurs.

2. Compute the vector $\overline{y} = (A_{*\alpha}^{-1})^T c_\alpha$ and the numbers $z_j = c_j - \langle A_{*j}, \overline{y}\rangle$ for $j \notin \alpha$.
3. *Stopping criterion:* **quit** if $z_j \geq 0$ for all $j \notin \alpha$.
   Optimal solution: $\overline{x}$. Optimal value: λ. Optimal dual solution: $\overline{y}$.
4. Choose otherwise an index k such that $z_k < 0$, compute the corresponding search vector v, i.e. $v_\alpha = A_{*\alpha}^{-1}A_{*k}$, $v_k = -1$ and $v_j = 0$ for $j \notin \alpha \cup \{k\}$, and put $I_+ = \{j \in \alpha \mid v_j > 0\}$.
5. *Stopping criterion:* **quit** if $I_+ = \emptyset$.
   Optimal value: $-\infty$.
6. Define otherwise $\tau = \min\{\overline{x}_j/v_j \mid j \in I_+\}$ and determine an index r so that $\alpha_r \in I_+$ and $\overline{x}_{\alpha_r}/v_{\alpha_r} = \tau$.
7. Put $\alpha' = \alpha_{\hat{r}}[k]$ and compute the inverse $A_{*\alpha'}^{-1} = E_{\hat{r}}[v_\alpha]^{-1}A_{*\alpha}^{-1}$.
8. *Update:* $\alpha := \alpha'$, $A_{*\alpha}^{-1} := A_{*\alpha'}^{-1}$, $\overline{x} := \overline{x} - \tau v$, and $\lambda := \lambda + \tau z_k$.

Before we can call the above procedure an algorithm in the sense of a mechanical calculation that a machine can perform, we need to specify how to choose k in step 4 in the case when $z_j < 0$ for several indices j, and r in step 6 when $\overline{x}_j/v_j = \tau$ for more than one index $j \in I_+$.

A simple rule that works well most of the time, is to select the index j that minimizes z_j (and if there are several such indices the least of these) as the index k, and the smallest of all indices i for which $\overline{x}_{\alpha_i}/v_{\alpha_i} = \tau$ as the index r. We shall return to the choice of k and r later; for the immediate discussion of the algorithm, it does not matter how to make the choice.

We also need a method to find an initial feasible basic index set to start the simplex algorithm from. We shall treat this problem and solve it in Section 13.6.

Now suppose that we apply the simplex algorithm to an LP problem in standard form, starting from a feasible basic index set. It follows from Theorem 13.4.2 that the algorithm delivers an optimal solution if it stops

Download free eBooks at bookboon.com

during step 3, and that the objective function is unbounded from below if the algorithm stops during step 5.

So let us examine what happens if the algorithm does not stop. Since a feasible basic index set is generated each time the algorithm comes to step 7, we will obtain in this case an infinite sequence $\alpha^1, \alpha^2, \alpha^3, \dots$ of feasible basic index sets with associated feasible basic solutions $\overline{x}^1, \overline{x}^2, \overline{x}^3, \dots$. As the number of different basic index sets is finite, some index set α^p has to be repeated after a number of additional, say q, iterations. This means that $\alpha^p = \alpha^{p+q}$ and $\overline{x}^p = \overline{x}^{p+q}$ and in turn implies that the sequence $\alpha^p, \alpha^{p+1}, \dots, \alpha^{p+q-1}$ is repeated periodically in all infinity. We express this by saying that the algorithm *cycles*. According to (ii) in Theorem 13.4.2,

$$\langle c, \overline{x}^p\rangle \geq \langle c, \overline{x}^{p+1}\rangle \geq \cdots \geq \langle c, \overline{x}^{p+q}\rangle = \langle c, \overline{x}^p\rangle,$$

and this implies that

$$\langle c, \overline{x}^p\rangle = \langle c, \overline{x}^{p+1}\rangle = \cdots = \langle c, \overline{x}^{p+q-1}\rangle.$$

The number τ is hence equal to 0 for all the iterations of the cycle, and this implies that the basic solutions $\overline{x}^p, \overline{x}^{p+1}, \dots, \overline{x}^{p+q-1}$ are identical and degenerate. If the simplex algorithm does not stop, but continues indefinitely, it is so because the algortihm has got stuck in a degenerate basic solution.

The following theorem is now an immediate consequence of the above discussion.

Theorem 13.4.3. *The simplex algorithm stops when applied to an LP problem in which all feasible basic solutions are non-degenerate.*

Cycling can occur, and we shall give an example of this in the next section. Theoretically, this is a bit troublesome, but cycling seems to be a rare phenomenon in practical problems and therefore lacks practical significance. The small rounding errors introduced during the numerical treatment of an LP problem also have a beneficial effect since these errors usually turn degenerate basic solutions into non-degenerate solutions and thereby tend to prevent cycling. There is also a simple rule for the choice of indices k and r, *Bland's rule*, which prevents cycling and will be described in the next section.

Download free eBooks at bookboon.com

Example

Example 13.4.1. We now illustrate the simplex algorithm by solving the minimization problem

$$\begin{array}{ll} \min & x_1 - x_2 + x_3 \\ \text{s.t.} & \begin{cases} -2x_1 + x_2 + x_3 \le 3 \\ -x_1 + x_2 - 2x_3 \le 3 \\ 2x_1 - x_2 + 2x_3 \le 1, \quad x \ge 0. \end{cases} \end{array}$$

We start by writing the problem in standard form by introducing three slack variables:

$$\begin{array}{ll} \min & x_1 - x_2 + x_3 \\ \text{s.t.} & \begin{cases} -2x_1 + x_2 + x_3 + x_4 = 3 \\ -x_1 + x_2 - 2x_3 + x_5 = 3 \\ 2x_1 - x_2 + 2x_3 + x_6 = 1, \quad x \ge 0. \end{cases} \end{array}$$

Using matrices, this becomes

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b,\ x \ge 0 \end{array}$$

Download free eBooks at bookboon.com

Click on the ad to read more

with

$$A = \begin{bmatrix} -2 & 1 & 1 & 1 & 0 & 0 \\ -1 & 1 & -2 & 0 & 1 & 0 \\ 2 & -1 & 2 & 0 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix} \quad \text{and}$$
$$c^T = \begin{bmatrix} 1 & -1 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

We note that we can start the simplex algorithm with

$$\alpha = (4,5,6), \quad A_{*\alpha}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \overline{x}_\alpha = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix},$$
$$\lambda = \langle c_\alpha, \overline{x}_\alpha \rangle = c_\alpha^T \overline{x}_\alpha = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix} = 0.$$

1st iteration:

$$\overline{y} = (A_{*\alpha}^{-1})^T c_\alpha = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$z_{1,2,3} = c_{1,2,3} - (A_{*1,2,3})^T \overline{y} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - \begin{bmatrix} -2 & -1 & 2 \\ 1 & 1 & -1 \\ 1 & -2 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

Since $z_2 = -1 < 0$, we have to select $k = 2$ and then

$$v_\alpha = A_{*\alpha}^{-1} A_{*k} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \quad v_2 = -1$$
$$I_+ = \{j \in \alpha \mid v_j > 0\} = \{4,5\}$$
$$\tau = \min\{\overline{x}_j / v_j \mid j \in I_+\} = \min\{\overline{x}_4/v_4, \overline{x}_5/v_5\} = \min\{3/1, 3/1\} = 3$$
$$\text{for } \alpha_1 = 4, \text{ i.e.}$$
$$r = 1$$
$$\alpha' = \alpha_{\hat{r}}[k] = (4,5,6)_{\hat{1}}[2] = (2,5,6)$$
$$E_{\hat{r}}[v_\alpha]^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$
$$A_{*\alpha'}^{-1} = E_{\hat{r}}[v_\alpha]^{-1} A_{*\alpha}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Download free eBooks at bookboon.com

$$\overline{x}'_{\alpha'} = \overline{x}_{\alpha'} - \tau v_{\alpha'} = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix} - 3 \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}$$
$$\lambda' = \lambda + \tau z_k = 0 + 3\,(-1) = -3.$$

Update: $\alpha := \alpha'$, $A^{-1}_{*\alpha} := A^{-1}_{*\alpha'}$, $\overline{x}_\alpha := \overline{x}'_{\alpha'}$ and $\lambda := \lambda'$.

2nd iteration:

$$\overline{y} = (A^{-1}_{*\alpha})^T c_\alpha = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$$
$$z_{1,3,4} = c_{1,3,4} - (A_{*1,3,4})^T \overline{y} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} -2 & -1 & 2 \\ 1 & -2 & 2 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}.$$

Since $z_1 = -1 < 0$,

$$k = 1$$
$$v_\alpha = A^{-1}_{*\alpha} A_{*k} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \quad v_1 = -1$$
$$I_+ = \{j \in \alpha \mid v_j > 0\} = \{5\}$$
$$\tau = \overline{x}_5 / v_5 = 0/1 = 0 \quad \text{for } \alpha_2 = 5, \text{ i.e.}$$
$$r = 2$$
$$\alpha' = \alpha_{\hat{r}}[k] = (2,5,6)_{\hat{2}}[1] = (2,1,6)$$
$$E_{\hat{r}}[v_\alpha]^{-1} = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$A^{-1}_{*\alpha'} = E_{\hat{r}}[v_\alpha]^{-1} A^{-1}_{*\alpha} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 2 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$
$$\overline{x}'_{\alpha'} = \overline{x}_{\alpha'} - \tau v_{\alpha'} = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix} - 0 \begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}$$
$$\lambda' = \lambda + \tau z_k = -3 + 0\,(-1) = -3.$$

Update: $\alpha := \alpha'$, $A^{-1}_{*\alpha} := A^{-1}_{*\alpha'}$, $\overline{x}_\alpha := \overline{x}'_{\alpha'}$ and $\lambda := \lambda'$.

Download free eBooks at bookboon.com

3rd iteration:

$$\overline{y} = (A_{*\alpha}^{-1})^T c_\alpha = \begin{bmatrix} -1 & -1 & 1 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$$

$$z_{3,4,5} = c_{3,4,5} - (A_{*3,4,5})^T \overline{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 & -2 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Since $z_3 = -1 < 0$,

$$\begin{aligned}
k &= 3 \\
v_\alpha &= A_{*\alpha}^{-1} A_{*k} = \begin{bmatrix} -1 & 2 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} -5 \\ -3 \\ 3 \end{bmatrix}, \qquad v_3 = -1 \\
I_+ &= \{j \in \alpha \mid v_j > 0\} = \{6\} \\
\tau &= \overline{x}_6 / v_6 = 4/3 \quad \text{for } \alpha_3 = 6, \text{ i.e.} \\
r &= 3 \\
\alpha' &= \alpha_{\hat{r}}[k] = (2,1,6)_{\hat{3}}[3] = (2,1,3)
\end{aligned}$$

Click on the ad to read more

$$E_{\hat{r}}[v_\alpha]^{-1} = \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & -3 \\ 0 & 0 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & \frac{5}{3} \\ 0 & 1 & 1 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$$

$$A_{*\alpha'}^{-1} = E_{\hat{r}}[v_\alpha]^{-1} A_{*\alpha}^{-1} = \begin{bmatrix} 1 & 0 & \frac{5}{3} \\ 0 & 1 & 1 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} -1 & 2 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & 2 & \frac{5}{3} \\ 0 & 1 & 1 \\ \frac{1}{3} & 0 & \frac{1}{3} \end{bmatrix}$$

$$\overline{x}'_{\alpha'} = \overline{x}_{\alpha'} - \tau v_{\alpha'} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} - \frac{4}{3} \begin{bmatrix} -5 \\ -3 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{29}{3} \\ 4 \\ \frac{4}{3} \end{bmatrix}$$

$$\lambda' = \lambda + \tau z_k = -3 + \frac{4}{3}(-1) = -\frac{13}{3}.$$

Update: $\alpha := \alpha'$, $A_{*\alpha}^{-1} := A_{*\alpha'}^{-1}$, $\overline{x}_\alpha := \overline{x}'_{\alpha'}$ and $\lambda := \lambda'$.

4th iteration:

$$\overline{y} = (A_{*\alpha}^{-1})^T c_\alpha = \begin{bmatrix} \frac{2}{3} & 0 & \frac{1}{3} \\ 2 & 1 & 0 \\ \frac{5}{3} & 1 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} \\ -1 \\ -\frac{1}{3} \end{bmatrix}$$

$$z_{4,5,6} = c_{4,5,6} - (A_{*4,5,6})^T \overline{y} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{3} \\ -1 \\ -\frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ 1 \\ \frac{1}{3} \end{bmatrix}.$$

The solution $\overline{x} = (4, \frac{29}{3}, \frac{4}{3}, 0, 0, 0)$ is optimal with optimal value $-\frac{13}{3}$ since $z_{4,5,6} > 0$. The original minimization problem has the same optimal value, of course, and $(x_1, x_2, x_3) = (4, \frac{29}{3}, \frac{4}{3})$ is the optimal solution. $\square$

The version of the simplex algorithm that we have presented is excellent for computer calculations, but it is unnecessarily complicated for calculations by hand. Then it is better to use the tableau form which we utilized in Section 13.2, even if this entails performing unnecessary calculations. To the LP problem

$$\begin{array}{ll} \min & \langle c, x \rangle \\ \text{s.t.} & Ax = b,\ x \geq 0 \end{array}$$

we associate the following simplex tableau:

$$\text{(13.9)} \qquad \begin{array}{c|c|c} A & b & E \\ \hline c^T & 0 & 0^T \end{array}$$

We have included the column on the far right of the table only to explain how the tableau calculations work; it will be omitted later on.

Download free eBooks at bookboon.com

Let α be a feasible basic index set with $\overline{x}$ as the corresponding basic solution. The upper part $[\,A \mid b \mid E\,]$ of the tableau can be seen as a matrix, and by multiplying this matrix from the left by $A_{*\alpha}^{-1}$, we obtain the following new tableau:

$$\begin{array}{c|c|c} A_{*\alpha}^{-1}A & A_{*\alpha}^{-1}b & A_{*\alpha}^{-1} \\ \hline c^T & 0 & 0^T \end{array}$$

Now subtract the upper part of this tableau multiplied from the left by c_α^T from the bottom row of the tableau. This results in the tableau

$$\begin{array}{c|c|c} A_{*\alpha}^{-1}A & A_{*\alpha}^{-1}b & A_{*\alpha}^{-1} \\ \hline c^T - c_\alpha^T A_{*\alpha}^{-1}A & -c_\alpha^T A_{*\alpha}^{-1}b & -c_\alpha^T A_{*\alpha}^{-1} \end{array}$$

Using the notation introduced in the definition of the simplex algorithm, we have $A_{*\alpha}^{-1}b = \overline{x}_\alpha$, $c_\alpha^T A_{*\alpha}^{-1} = \big((A_{*\alpha}^{-1})^T c_\alpha\big)^T = \overline{y}^T$, $c^T - c_\alpha^T A_{*\alpha}^{-1}A = c^T - \overline{y}^T A = z^T$ and $c_\alpha^T A_{*\alpha}^{-1}b = c_\alpha^T \overline{x}_\alpha = \langle c_\alpha, \overline{x}_\alpha\rangle = \lambda$, which means that the above tableau can be written in the form

$$\begin{array}{c|c|c} A_{*\alpha}^{-1}A & \overline{x}_\alpha & A_{*\alpha}^{-1} \\ \hline z^T & -\lambda & -\overline{y}^T \end{array} \tag{13.10}$$

Note that the columns of the unit matrix appear as columns in the matrix $A_{*\alpha}^{-1}A$, because column number α_j in $A_{*\alpha}^{-1}A$ is identical with unit matrix column E_{*j}. Moreover, $z_{\alpha_j} = 0$.

When performing the actual calculations, we use Gaussian elimination to get from tableau (13.9) to tableau (13.10).

If $z^T \geq 0$, which we can determine with the help of the bottom line in (13.10), then $\overline{x}$ is an optimal solution, and we can also read off the optimal solution $\overline{y}$ to the dual maximization problem. (The matrix A will in many cases contain the columns of the unit matrix, and if so then it is of course possible to read off the solution to the dual problem in the final simplex tableau without first having to add the unit matrix on the right side of tableau (13.9).)

If $z^T \not\geq 0$, then we choose a column index k with $z_k < 0$, and consider the corresponding column $a = A_{*\alpha}^{-1}A_{*k}$ $(= v_\alpha)$ in the upper part of the tableau.

The minimization problem is unbounded if $a \leq 0$. In the opposite case, we choose an index $i = r$ that minimizes $\overline{x}_{\alpha_i}/a_i$ $(= \overline{x}_{\alpha_i}/v_{\alpha_i})$ among all ratios with positive a_i. This means that r is the index of a row with the least ratio $\overline{x}_{\alpha_i}/a_i$ among all rows with positive a_i. Finally, we transform the simplex tableau by pivoting around the element at location (r, k).

Download free eBooks at bookboon.com

EXAMPLE 13.4.2. We solve Example 13.4.1 again – this time by performing all calculations in tabular form. Our first tableau has the form

$$\begin{array}{rrrrrr|r}
-2 & \underline{1} & 1 & 1 & 0 & 0 & 3 \\
-1 & 1 & -2 & 0 & 1 & 0 & 3 \\
2 & -1 & 2 & 0 & 0 & 1 & 1 \\
\hline
1 & -1 & 1 & \underline{0} & \underline{0} & \underline{0} & 0
\end{array}$$

and in this case it is of course not necessary to repeat the columns of the unit matrix in a separate part of the tableau in order also to solve the dual problem.

The basic index set $\alpha = (4, 5, 6)$ is feasible, and since $A_{*\alpha} = E$ and $c_\alpha^T = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$, we can directly read off $z^T = \begin{bmatrix} 1 & -1 & 1 & 0 & 0 & 0 \end{bmatrix}$ and $-\lambda = 0$ from the bottom line of the tableau.

The optimality criterion is not satisfied since $z_2 = -1 < 0$, so we proceed by choosing $k = 2$. The positive ratios of corresponding elements in the right-hand side column and the second column are in this case the same and equal to $3/1$ for the first and the second row. Therefore, we can choose $r = 1$ or $r = 2$, and we decide to use the smaller of the two numbers, i.e. we put $r = 1$. The tableau is then transformed by pivoting around the element at

Need help with your dissertation?

Get in-depth feedback & advice from experts in your topic area. Find out what you can do to improve the quality of your dissertation!

Get Help Now

Go to www.helpmyassignment.co.uk for more info

Helpmyassignment

Click on the ad to read more

location $(1,2)$. By then continuing in the same style, we get the following sequence of tableaux:

$$\begin{array}{rrrrrr|r}
-2 & 1 & 1 & 1 & 0 & 0 & 3 \\
\underline{1} & 0 & -3 & -1 & 1 & 0 & 0 \\
0 & 0 & 3 & 1 & 0 & 1 & 4 \\
\hline
-1 & \underline{0} & 2 & 1 & \underline{0} & \underline{0} & 3
\end{array}$$

$\alpha = (2,5,6), \quad k = 1, \quad r = 2$

$$\begin{array}{rrrrrr|r}
0 & 1 & -5 & -1 & 2 & 0 & 3 \\
1 & 0 & -3 & -1 & 1 & 0 & 0 \\
0 & 0 & \underline{3} & 1 & 0 & 1 & 4 \\
\hline
\underline{0} & \underline{0} & -1 & 0 & 1 & \underline{0} & 3
\end{array}$$

$\alpha = (2,1,6), \quad k = 3, \quad r = 3$

$$\begin{array}{rrrrrr|r}
0 & 1 & 0 & \frac{2}{3} & 2 & \frac{5}{3} & \frac{29}{3} \\
1 & 0 & 0 & 0 & 1 & 1 & 4 \\
0 & 0 & 1 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{4}{3} \\
\hline
\underline{0} & \underline{0} & \underline{0} & \frac{1}{3} & 1 & \frac{1}{3} & \frac{13}{3}
\end{array}$$

$\alpha = (2,1,3)$

The optimality criterion is now satisfied with $\overline{x} = (4, \frac{29}{3}, \frac{4}{3}, 0, 0, 0)$ as optimal solution and $-\frac{13}{3}$ as optimal value. The dual problem has the optimal solution $(-\frac{1}{3}, -1, -\frac{1}{3})$. $\square$

Henceforth, we will use the tableau variant of the simplex algorithm to account for our calculations, because it is the most transparent method.

The optimality condition in step 2 of the simplex algorithm is a sufficient condition for optimality, but the condition is not necessary. A degenerate basic solution can be optimal without the optimality condition being satisfied. Here is a trivial example of this.

Example 13.4.3. The problem

$$\begin{array}{ll}
\min & -x_2 \\
\text{s.t.} & x_1 + x_2 = 0, \quad x \geq 0
\end{array}$$

has only one feasible point, $x = (0,0)$, which is therefore optimal. There are two feasible basic index sets, $\alpha = (1)$ and $\alpha' = (2)$, both with $(0,0)$ as the corresponding degenerate basic solution.

Download free eBooks at bookboon.com

The optimality condition is not fulfilled at the basic index set α, because $\overline{y} = 1 \cdot 0 = 0$ and $z_2 = -1 - 1 \cdot 0 = -1 < 0$. At the other basic index set α', $\overline{y} = 1 \cdot (-1) = -1$ and $z_2 = 0 - 1 \cdot (-1) = 1 > 0$, and the optimality criterion is now satisfied.

The corresponding simplex tableaux are

$$\begin{array}{cc|c} 1 & 1 & 0 \\ \hline \underline{0} & -1 & 0 \end{array} \qquad \text{and} \qquad \begin{array}{cc|c} 1 & 1 & 0 \\ \hline 1 & \underline{0} & 0 \end{array}$$
$$\alpha = (1) \qquad\qquad\qquad \alpha = (2)$$

□

We shall now study a simple example with a non-unique optimal solution.

EXAMPLE 13.4.4. The simplex tableaux associated with the problem

$$\begin{array}{ll} \min & x_1 + x_2 \\ \text{s.t.} & \begin{cases} x_1 + x_2 - x_3 \phantom{{}+x_4} = 1 \\ \phantom{x_1 +{}} 2x_2 - x_3 + x_4 = 1, \ \ x \geq 0 \end{cases} \end{array}$$

are as follows:

$$\begin{array}{cccc|c} 1 & 1 & -1 & 0 & 1 \\ 0 & 2 & -1 & 1 & 1 \\ \hline 1 & 1 & 0 & 0 & 0 \end{array}$$
$$\alpha = (1, 4)$$

$$\begin{array}{cccc|c} 1 & 1 & -1 & 0 & 1 \\ 0 & 2 & -1 & 1 & 1 \\ \hline \underline{0} & 0 & 1 & \underline{0} & -1 \end{array}$$
$$\alpha = (1, 4)$$

The optimality condition is met; $\overline{x} = (1, 0, 0, 1)$ is an optimal solution, and the optimal value is 1. However, coefficient number 2 in the last row, i.e. z_2, is equal to 0, so we can therefore perform another iteration of the simplex algorithm by choosing the second column as the pivot column and the second row as the pivot row, i.e. $k = 2$ and $r = 2$. This gives rise to the following new tableau:

$$\begin{array}{cccc|c} 1 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \hline \underline{0} & \underline{0} & 1 & 0 & -1 \end{array}$$
$$\alpha = (1, 2)$$

Download free eBooks at bookboon.com

The optimality condition is again met, now with $\hat{x} = (\frac{1}{2}, \frac{1}{2}, 0, 0)$ as optimal solution. Since the set of optimal solutions is convex, each point on the line segment between $\hat{x}$ and $\overline{x}$ is also an optimal point. $\square$

13.5 Bland's anti cycling rule

We begin with an example of Kuhn showing that cycling can occur in degenerate LP problems if the column index k and the row index r are not properly selected.

Example 13.5.1. Consider the problem

$$\begin{array}{ll} \min & -2x_1 - 3x_2 + x_3 + 12x_4 \\ \text{s.t.} & \begin{cases} -2x_1 - 9x_2 + x_3 + 9x_4 + x_5 = 0 \\ \frac{1}{3}x_1 + x_2 - \frac{1}{3}x_3 - 2x_4 + x_6 = 0 \\ 2x_1 + 3x_2 - x_3 - 12x_4 + x_7 = 2, \quad x \geq 0. \end{cases} \end{array}$$

We use the simplex algorithm with the additional rule that the column index k should be chosen so as to make z_k as negative as possible and the row index r should be the least among all allowed row indices. Our first tableau is

Download free eBooks at bookboon.com

Click on the ad to read more

$$\begin{array}{rrrrrrr|r}
-2 & -9 & 1 & 9 & 1 & 0 & 0 & 0 \\
\frac{1}{3} & \underline{1} & -\frac{1}{3} & -2 & 0 & 1 & 0 & 0 \\
2 & 3 & -1 & -12 & 0 & 0 & 1 & 2 \\
\hline
-2 & -3 & 1 & 12 & \underline{0} & \underline{0} & \underline{0} & 0
\end{array}$$

with $\alpha = (5, 6, 7)$ as feasible basic index set. According to our rule for the choice of of k, we must choose $k = 2$. There is only one option for the row index r, namely $r = 2$, so we use the element located at $(2, 2)$ as pivot element and obtain the following new tableau

$$\begin{array}{rrrrrrr|r}
\underline{1} & 0 & -2 & -9 & 1 & 9 & 0 & 0 \\
\frac{1}{3} & 1 & -\frac{1}{3} & -2 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & -6 & 0 & -3 & 1 & 2 \\
\hline
-1 & \underline{0} & 0 & 6 & \underline{0} & 3 & \underline{0} & 0
\end{array}$$

with $\alpha = (5, 2, 7)$. This time $k = 1$, but there are two row indices i with the same least value of the ratios $\overline{x}_{\alpha_i}/v_{\alpha_i}$, namely 1 and 2. Our additional rule tells us to choose $r = 1$. Pivoting around the element at location $(1, 1)$ results in the next tableau

$$\begin{array}{rrrrrrr|r}
1 & 0 & -2 & -9 & 1 & 9 & 0 & 0 \\
0 & 1 & \frac{1}{3} & \underline{1} & -\frac{1}{3} & -2 & 0 & 0 \\
0 & 0 & 2 & 3 & -1 & -12 & 1 & 2 \\
\hline
\underline{0} & \underline{0} & -2 & -3 & 1 & 12 & \underline{0} & 0
\end{array}$$

with $\alpha = (1, 2, 7)$, $k = 4$, $r = 2$.

The algorithms goes on with the following sequence of tableaux:

$$\begin{array}{rrrrrrr|r}
1 & 9 & \underline{1} & 0 & -2 & -9 & 0 & 0 \\
0 & 1 & \frac{1}{3} & 1 & -\frac{1}{3} & -2 & 0 & 0 \\
0 & -3 & 1 & 0 & 0 & -6 & 1 & 2 \\
\hline
\underline{0} & 3 & -1 & \underline{0} & 0 & 6 & \underline{0} & 0
\end{array}$$

$$\alpha = (1, 4, 7), \quad k = 3, \quad r = 1$$

$$\begin{array}{rrrrrrr|r}
1 & 9 & 1 & 0 & -2 & -9 & 0 & 0 \\
-\frac{1}{3} & -2 & 0 & 1 & \frac{1}{3} & \underline{1} & 0 & 0 \\
-1 & -12 & 0 & 0 & 2 & 3 & 1 & 2 \\
\hline
\underline{1} & 12 & \underline{0} & \underline{0} & -2 & -3 & \underline{0} & 0
\end{array}$$

$$\alpha = (3, 4, 7), \quad k = 6, \quad r = 2$$

Download free eBooks at bookboon.com

$$\begin{array}{rrrrrrr|r}
-2 & -9 & 1 & 9 & \underline{1} & 0 & 0 & 0 \\
-\frac{1}{3} & -2 & 0 & 1 & \frac{1}{3} & 1 & 0 & 0 \\
0 & -6 & 0 & -3 & 1 & 0 & 1 & 2 \\
\hline
0 & 6 & \underline{0} & 3 & -1 & \underline{0} & \underline{0} & 0
\end{array}$$

$$\alpha = (3,6,7), \quad k = 5, \quad r = 1$$

$$\begin{array}{rrrrrrr|r}
-2 & -9 & 1 & 9 & 1 & 0 & 0 & 0 \\
\frac{1}{3} & \underline{1} & -\frac{1}{3} & -2 & 0 & 1 & 0 & 0 \\
2 & 3 & -1 & -12 & 0 & 0 & 1 & 2 \\
\hline
-2 & -3 & 1 & 12 & \underline{0} & \underline{0} & \underline{0} & 0
\end{array}$$

$$\alpha = (5,6,7)$$

After six iterations we are back to the starting tableau. The simplex algorithm cycles! □

We now introduce a rule for the choice of indices k and r that prevents cycling.

Bland's rule: *Choose k in step 4 of the simplex algorithm so that*

$$k = \min\{j \mid z_j < 0\}$$

and r in step 6 so that

$$\alpha_r = \min\{j \in I_+ \mid \overline{x}_j/v_j = \tau\}.$$

EXAMPLE 13.5.2. Consider again the minimization problem in the previous example and now use the simplex algorithm with Bland's rule. This results in the following sequence of tableaux:

$$\begin{array}{rrrrrrr|r}
-2 & -9 & 1 & 9 & 1 & 0 & 0 & 0 \\
\underline{\frac{1}{3}} & 1 & -\frac{1}{3} & -2 & 0 & 1 & 0 & 0 \\
2 & 3 & -1 & -12 & 0 & 0 & 1 & 2 \\
\hline
-2 & -3 & 1 & 12 & \underline{0} & \underline{0} & \underline{0} & 0
\end{array}$$

$$\alpha = (5,6,7), \quad k = 1, \quad r = 2$$

Download free eBooks at bookboon.com

$$\begin{array}{rrrrrrr|r}
0 & -3 & -1 & -3 & 1 & 6 & 0 & 0 \\
1 & 3 & -1 & -6 & 0 & 3 & 0 & 0 \\
0 & -3 & \underline{1} & 0 & 0 & -6 & 1 & 2 \\
\hline
\underline{0} & 3 & -1 & 0 & \underline{0} & 6 & \underline{0} & 0
\end{array}$$

$$\alpha = (5, 1, 7), \quad k = 3, \quad r = 3$$

$$\begin{array}{rrrrrrr|r}
0 & -6 & 0 & -3 & 1 & 0 & 0 & 2 \\
1 & 0 & 0 & -6 & 0 & -3 & 1 & 2 \\
0 & -3 & 1 & 0 & 0 & -6 & 1 & 2 \\
\hline
\underline{0} & 0 & \underline{0} & 0 & \underline{0} & 12 & 1 & 2
\end{array}$$

$$\alpha = (5, 1, 3)$$

The optimality criterion is met with $\overline{x} = (2, 0, 2, 0, 2, 0, 0)$ as optimal solution and -2 as optimal value. □

Theorem 13.5.1. *The simplex algorithm always stops if Bland's rule is used.*

Click on the ad to read more

Proof. We prove the theorem by contradiction. So suppose that the simplex algorithm cycles when applied to some given LP problem, and let $\overline{x}$ be the common basic solution during the iterations of the cycle.

Let $\mathcal{C}$ denote the set of indices k of the varibles x_k that change from being basic to being free during the iterations of the cycle. Since these variables have to return as basic variables during the cycle, $\mathcal{C}$ is of course also equal to the set of indices of the variables x_k that change from being free to being basic during the cycle. Moreover, $\overline{x}_k = 0$ for all $k \in \mathcal{C}$.

Let

$$q = \max\{j \mid j \in \mathcal{C}\},$$

and let α be the basic index set which is in use during the iteration in the cycle when the variabel x_q changes from being basic to being free, and let x_k be the free variable that replaces x_q. The index q is in other words replaced by k in the basic index set that follows after α. The corresponding search vector v and reduced cost vector z satisfy the inequalities

$$z_k < 0 \quad \text{and} \qquad v_q > 0,$$

and

$$z_j \geq 0 \quad \text{for } j < k.$$

since the index k is chosen according to Bland's rule. Since $k \in \mathcal{C}$, we also have $k < q$, because of the definition of q.

Let us now consider the basic index set α' that belongs to an iteration when x_q returns as a basic variables after having been free. Because of Bland's rule for the choice of incoming index, in this case q, the corresponding reduced cost vector z' has to satisfy the following inequalities:

$$z'_j \geq 0 \quad \text{for } j < q \qquad \text{and} \qquad z'_q < 0. \tag{13.11}$$

Especially, thus $z'_k \geq 0$.

Since $Av = 0$, $v_k = -1$ and $v_j = 0$ for $j \notin \alpha \cup \{k\}$, and $z_j = 0$ for $j \in \alpha$, it follows from Lemma 13.4.1 that

$$\sum_{j\in\alpha} z'_j v_j - z'_k = \langle z', v\rangle = \langle c, v\rangle = \langle z, v\rangle = \sum_{j\in\alpha} z_j v_j + z_k v_k = -z_k > 0,$$

and hence

$$\sum_{j\in\alpha} z'_j v_j > z'_k \geq 0.$$

There is therefore an index $j_0 \in \alpha$ such that $z'_{j_0} v_{j_0} > 0$. Hence $z'_{j_0} \neq 0$, which means that j_0 can not belong to the index set α'. The variable x_{j_0} is

Download free eBooks at bookboon.com

in other words basic during one iteration of the cycle and free during another iteration. This means that j_0 is an index in the set $\mathcal{C}$, and hence $j_0 \le q$, by the definition of q. The case $j_0 = q$ is impossible since $v_q > 0$ and $z'_q < 0$. Thus $j_0 < q$, and it now follows from (13.11) that $z'_{j_0} > 0$. This implies in turn that $v_{j_0} > 0$, because the product $z'_{j_0} v_{j_0}$ is positive. So j_0 belongs to the set $I_+ = \{j \in \alpha \mid v_j > 0\}$, and since $\overline{x}_{j_0}/v_{j_0} = 0 = \tau$, it follows that

$$\min\{j \in I_+ \mid \overline{x}_j/v_j = \tau\} \le j_0 < q.$$

The choice of q thus contradicts Bland's rule for how to choose index to leave the basic index set α, and this contradiction proves the theorem. □

Remark. It is not necessary to use Bland's rule all the time in order to prevent cycling; it suffices to use it in iterations with $\tau = 0$.

13.6 Phase 1 of the simplex algorithm

The simplex algorithm assumes that there is a feasible basic index set to start from. For some problems we will automatically get one when the problem is written in standard form. This is the case for problems of the type

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax \le b,\ x \ge 0 \end{array}$$

where A is an $m \times n$-matrix and the right-hand side vector b is nonnegative. By introducing m slack variables $s_{n+1}, s_{n+2}, \dots, s_{n+m}$ and defining

$$s = (s_{n+1}, s_{n+2}, \dots, s_{n+m}),$$

we obtain the standard problem

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax + Es = b,\ \ x, s \ge 0, \end{array}$$

and it is now obvious how to start; the slack variables will do as basic variables, i.e. $\alpha = (n+1, n+2, \dots, n+m)$ is a feasible basic index set with $\overline{x} = 0$, $\overline{s} = b$ as the corresponding basic solution.

In other cases, it is not at all obvious how to find a feasible basic index set to start from, but one can always generate such a set by using the simplex algorithm on a suitable artificial problem.

Consider an arbitrary standard LP problem

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b,\ x \ge 0, \end{array}$$

Download free eBooks at bookboon.com

where A is an $m \times n$-matrix. We can assume without restriction that $b \geq 0$, for if any b_j is negative, we just multiply the corresponding equation by -1.

We begin by choosing an $m \times k$-matris B so that the matrix

$$A' = \begin{bmatrix} A & B \end{bmatrix}$$

gets rank equal to m and the system

$$A' \begin{bmatrix} x \\ y \end{bmatrix} = Ax + By = b$$

gets an obvious feasible basic index set α^0. The new y-variables are called *artificial variables*, and we number them so that $y = (y_{n+1}, y_{n+2}, \ldots, y_{n+k})$.

A trivial way to achieve this is to choose B equal to the unit matrix E of order m, for $\alpha^0 = (n+1, n+2, \ldots, n+m)$ is then a feasible basic index set with $(\overline{x}, \overline{y}) = (0, b)$ as the corresponding feasible basic solution. Often, however, A already contains a number of unit matrix columns, and then it is sufficient to add the missing unit matrix columns to A.

Now let

$$\mathbf{1} = \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}^T$$

Click on the ad to read more

be the $k \times 1$-matrix consisting of k ones, and consider the following artificial LP problems:

$$\begin{array}{ll} \min & \langle \mathbf{1}, y\rangle = y_{n+1} + \cdots + y_{n+k}. \\ \text{s.t.} & Ax + By = b,\ x, y \geq 0 \end{array}$$

The optimal value is obviously ≥ 0, and *the value is equal to zero if and only if there is a feasible solution of the form* $(x, 0)$*, i.e. if and only if there is a nonnegative solution to the system* $Ax = b$*.*

Therefore, we solve the artificial problem using the simplex algorithm with α^0 as the first feasible basic index set. Since the objective function is bounded below, the algorithm stops after a finite number of iterations (perhaps we need to use Bland's supplementary rule) in a feasible basic index set α, where the optimality criterion is satisfied. Let $(\overline{x}, \overline{y})$ denote the corresponding basic solution.

There are now two possibilities.

Case 1. *The artificial problem's optimal value is greater than zero.*

In this case, the original problem has no feasible solutions.

Case 2. *The artificial problem's value is equal to zero.*

In this case, $\overline{y} = 0$ and $A\overline{x} = b$.

If $\alpha \subseteq \{1, 2, \ldots, n\}$, then α is also a feasible basic index set of the matrix A, and we can start the simplex algorithm on our original problem from α and the corresponding feasible basic solution $\overline{x}$.

If $\alpha \not\subseteq \{1, 2, \ldots, n\}$, we set

$$\alpha' = \alpha \cap \{1, 2, \ldots, n\}.$$

The matrix columns $\{A_{*k} \mid k \in \alpha'\}$ are now linearly independent, and we can construct an index set $\beta \supseteq \alpha'$, which is maximal with respect to the property that the columns $\{A_{*k} \mid k \in \beta\}$ are linearly independent.

If $\operatorname{rank} A = m$, then β will consist of m elements, and β is then a basic index set of the matrix A. Since $\overline{x}_j = 0$ for all $j \notin \alpha'$, and thus especially for all $j \notin \beta$, it follows that $\overline{x}$ is the basic solution of the system $Ax = b$ that corresponds to the basic index set β. Hence, β is a feasible basic index set for our original problem. We can also note that $\overline{x}$ is a degenerate basic solution.

If $\operatorname{rank} A < m$, then β will consist of just $p = \operatorname{rank} A$ elements, but we can now delete $m - p$ equations from the system $Ax = b$ without changing the set of solutions. This results in a new equivalent LP problem with a coefficient matrix of rank p, and β is a feasible basic index set with $\overline{x}$ as the corresponding basic solution in this problem.

Download free eBooks at bookboon.com

To solve a typical LP problem, one thus normally needs to use the simplex algorithm twice. In Phase 1, we use the simplex algorithm to generate a feasible basic index set α for the original LP problem by solving an artificial LP problem, and in phase 2, the simplex algorithm is used to solve the original problem starting from the basic index set α.

EXAMPLE 13.6.1. We illustrate the technique on the simple problem

$$\begin{array}{ll}\min & x_1 + 2x_2 + x_3 - 2x_4 \\ \text{s.t.} & \begin{cases} x_1 + x_2 + x_3 - x_4 = 2 \\ 2x_1 + x_2 - x_3 + 2x_4 = 3 \\ x_1, x_2, x_3, x_4 \geq 0. \end{cases}\end{array}$$

Phase 1 consists in solving the artificial problem

$$\begin{array}{ll}\min & y_5 + y_6 \\ \text{s.t.} & \begin{cases} x_1 + x_2 + x_3 - x_4 + y_5 = 2 \\ 2x_1 + x_2 - x_3 + 2x_4 + y_6 = 3 \\ x_1, x_2, x_3, x_4, y_5, y_6 \geq 0. \end{cases}\end{array}$$

The computations are shown in tabular form, and the first simplex tableau is the following one.

$$\begin{array}{cccccc|c} 1 & 1 & 1 & -1 & 1 & 0 & 2 \\ 2 & 1 & -1 & 2 & 0 & 1 & 3 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{array}$$

We begin by eliminating the basic variables from the objective function and then obtain the following sequence of tableaux:

$$\begin{array}{cccccc|c} 1 & 1 & 1 & -1 & 1 & 0 & 2 \\ \underline{2} & 1 & -1 & 2 & 0 & 1 & 3 \\ \hline -3 & -2 & 0 & -1 & \underline{0} & \underline{0} & -5 \end{array}$$

$$\alpha = (5, 6), \quad k = 1, \quad r = 2$$

$$\begin{array}{cccccc|c} 0 & \frac{1}{2} & \underline{\frac{3}{2}} & -2 & 1 & -\frac{1}{2} & \frac{1}{2} \\ 1 & \frac{1}{2} & -\frac{1}{2} & 1 & 0 & \frac{1}{2} & \frac{3}{2} \\ \hline \underline{0} & -\frac{1}{2} & -\frac{3}{2} & 2 & \underline{0} & \frac{3}{2} & -\frac{1}{2} \end{array}$$

$$\alpha = (5, 1), \quad k = 3, \quad r = 1$$

Download free eBooks at bookboon.com

$$
\begin{array}{cccccc|c}
0 & \frac{1}{3} & 1 & -\frac{4}{3} & \frac{2}{3} & -\frac{1}{3} & \frac{1}{3} \\
1 & \frac{2}{3} & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{5}{3} \\
\hline
\underline{0} & 0 & \underline{0} & 0 & 1 & 1 & 0
\end{array}
$$

$$\alpha = (3,1)$$

The above final tableau for the artificial problem shows that $\alpha = (3,1)$ is a feasible basic index set for the original problem with $\overline{x} = (\frac{5}{3}, 0, \frac{1}{3}, 0)$ as corresponding basic solution. We can therefore proceed to phase 2 with the following tableau as our first tableau.

$$
\begin{array}{cccc|c}
0 & \frac{1}{3} & 1 & -\frac{4}{3} & \frac{1}{3} \\
1 & \frac{2}{3} & 0 & \frac{1}{3} & \frac{5}{3} \\
\hline
1 & 2 & 1 & -2 & 0
\end{array}
$$

By eliminating the basic variables from the objective function, we obtain the following tableau:

Download free eBooks at bookboon.com
Click on the ad to read more

$$\begin{array}{cccc|c}
0 & \frac{1}{3} & 1 & -\frac{4}{3} & \frac{1}{3} \\
1 & \frac{2}{3} & 0 & \underline{\frac{1}{3}} & \frac{5}{3} \\
\hline
\underline{0} & 1 & \underline{0} & -1 & -2
\end{array}$$

$$\alpha = (3,1), \quad k = 4, \quad r = 2$$

One iteration is enough to obtain a tableau satisfying the optimality criterion.

$$\begin{array}{cccc|c}
4 & 3 & 1 & 0 & 7 \\
3 & 2 & 0 & 1 & 5 \\
\hline
3 & 3 & \underline{0} & \underline{0} & 3
\end{array}$$

$$\alpha = (3,4)$$

The optimal value is thus equal to -3, and $\overline{x} = (0,0,7,5)$ is the optimal solution. □

Since the volume of work grows with the number of artificial variables, one should not introduce more artificial variables than necessary. The minimization problem

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax \le b,\ x \ge 0 \end{array}$$

requires no more than one artificial variable. By introducing slack variables $s = (s_{n+1}, s_{n+2}, \dots, s_{n+m})$, we first obtain an equivalent standard problem

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax + Es = b,\ x, s \ge 0 \end{array}.$$

If $b \ge 0$, this problem can be solved, as we have already noted, without artificial variables. Let otherwise i_0 be the index of the most negative coordinate of the right-hand side b, and subtract equation no. i_0 in the system $Ax + Es = b$ from all other equations with negative right-hand side, and change finally the sign of equation no. i_0.

The result is a system of equations of the form $A'x + E's = b'$, which is equivalent to the system $Ax+Es = b$ and where $b' \ge 0$ and all the columns of the matrix E', except column no. i_0, are equal to the corresponding columns of the unit matrix E. Phase 1 of the simplex algorithm applied to the problem

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & A'x + E's = b',\ x, s \ge 0 \end{array}$$

therefore requires only one artificial variable.

Existence of optimal solutions and the duality theorem

The simplex algorithm is of course first and foremost an efficient algorithm for solving concrete LP problems, but we can also use it to provide alternative proofs of important theoretical results. These are corollaries to the following theorem.

Theorem 13.6.1. *Each standard LP problem with feasible points has a feasible basic index set where one of the two stopping criteria in the simplex algorithm is satisfied.*

Proof. Bland's rule ensures that phase 1 of the simplex algorithm stops with a feasible basic index set from where to start phase 2, and Bland's rule also ensures that this phase stops in a feasible basic index set, where one of the two stopping criteria is satisfied. □

As first corollary we obtain a new proof that every LP problem with finite value has optimal solutions (Theorem 12.1.1).

Corollary 13.6.2. *Each linear minimization problem with feasible solutions and downwards bounded objective function has an optimal solution.*

Proof. Since each LP problem can be replaced by an equivalent LP problem in standard form, it is sufficient to consider such problems. The only way for the simplex algorithm to stop, when the objective function is bounded below, is to stop at a basic solution which satisfies the optimality criterion. So it follows at once from the above theorem that there exists an optimal solution if the objective function is bounded below and the set of feasible solutions is nonempty. □

We can also give an algorithmic proof of the Duality theorem.

Corollary 13.6.3 (Duality theorem). *If the linear optimization problem*

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b,\ x \geq 0 \end{array}$$

has feasible solutions, then it has the same optimal value as the dual maximization problem

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & A^T y \leq c. \end{array}$$

Proof. Let α be the feasible basic index set where the simplex algorithm stops. If the optimality criterion is satisfied at α, then it follows from Theorem 13.4.2 that the minimization problem and the dual maximization problem have the same finite optimal value. If instead the algorithm stops because

Download free eBooks at bookboon.com

the objective function is unbounded below, then the dual problem has no feasible points according to Theorem 13.4.2, and the value of both problems is equal to $-\infty$, by definition. □

By writing general minimization problems in standard form, one can also deduce the general form of the duality theorem from the above special case.

13.7 Sensitivity analysis

In Section 12.1, we studied how the optimal value and the optimal solution depend on the coefficients of the objective function. In this section we shall study the same issue in connection with the simplex algorithm and also study how the solution to the LP problem

$$\text{(P)} \qquad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b,\ x \geq 0 \end{array}$$

depends on the right-hand side b. In real LP problems, the coefficients of the objective function and the constraints are often not exactly known, some of them might even be crude estimates, and it is then of course important to know how sensitive the optimal solution is to errors in input data. And even if the input data are accurate, it is of course interesting to know how the optimum solution is affected by changes in one or more of the coefficients.

Let α be a basic index set of the matrix A, and let $\overline{x}(b)$ denote the corresponding basic solution to the system $Ax = b$, i.e.

$$\overline{x}(b)_\alpha = A_{*\alpha}^{-1}b \quad \text{and} \quad \overline{x}(b)_j = 0 \text{ for all } j \notin \alpha.$$

Suppose that the LP problem (P) has an optimal solution for certain given values of b and c, and that this solution has been obtained because the simplex algorithm stopped at the basic index set α. For that to be the case, the basic solution $\overline{x}(b)$ has to be feasible, i.e.

$$\text{(i)} \qquad A_{*\alpha}^{-1}b \geq 0,$$

and the optimality criterion $z \geq 0$ in the simplex algorithm has to be satisfied. Since

$$z = c - A^T\overline{y} \quad \text{and} \quad \overline{y} = (A_{*\alpha}^{-1})^T c_\alpha,$$

we have $z = c - (A_{*\alpha}^{-1}A)^T c_\alpha$, which means that the optimality criterion can be written as

$$\text{(ii)} \qquad z(c) = c - (A_{*\alpha}^{-1}A)^T c_\alpha \geq 0.$$

Download free eBooks at bookboon.com

Conversely, $\overline{x}(b)$ is an optimal solution to the LP problem (P) for all b and c that satisfy the conditions (i) and (ii), because the optimality criterion in the simplex algorithm is then satisfied.

Condition (i) is a system of homogeneous linear inequalities in the variables b_1, b_2, ..., b_m, and it defines a polyhedral cone B_α in $\mathbf{R}^m$, while (ii) is a system of homogeneous linear inequalities in the variables $c_1, c_2, \dots, c_n$ and defines a polyhedral cone C_α in $\mathbf{R}^n$. In summary, we have the following result:

$\overline{x}(b)$ is an optimal solution to the LP problem (P) for all $b \in B_\alpha$ and all $c \in C_\alpha$.

Now suppose that we have solved the problem (P) for given values of b and c with $\overline{x} = \overline{x}(b)$ as optimal solution and λ as optimal value. Condition (ii) determines how much we are allowed to change the coefficients of the objective function without changing the optimal solution; $\overline{x}$ is still an optimal solution to the perturbed problem

$$(\mathrm{P}') \qquad \begin{array}{ll} \min & \langle c + \Delta c, x \rangle \\ \text{s.t.} & Ax = b,\ x \geq 0 \end{array}$$

Download free eBooks at bookboon.com

Click on the ad to read more

if $z(c+\Delta c) = z(c) + z(\Delta c) \geq 0$, i.e. if

$$\Delta c - (A_{*\alpha}^{-1}A)^T(\Delta c)_\alpha \geq -z(c). \tag{13.12}$$

The optimal value is of course changed to $\lambda + \langle \Delta c, \overline{x}\rangle$.

Inequality (13.12) defines a polyhedron in the variables Δc_1, Δc_2, ..., Δc_n. If for instance $\Delta c_j = 0$ for all j except $j = k$, i.e. if only the c_k-coefficient of the objective function is allowed to change, then inequality (13.12) determines a (possibly unbounded) closed interval $[-d_k, e_k]$ around 0 for Δc_k.

If instead we change the right-hand side of the constraints replacing the vector b by $b+\Delta b$, then $\overline{x}(b+\Delta b)$ becomes an optimal solution to the problem

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b + \Delta b,\ x \geq 0 \end{array}$$

as long as the solution is feasible, i.e. as long as $A_{*\alpha}^{-1}(b+\Delta b) \geq 0$. After simplification, this results in the condition

$$A_{*\alpha}^{-1}(\Delta b) \geq -\overline{x}(b)_\alpha,$$

which is a system of linear inequalities that determines how to choose Δb. If $\Delta b_i = 0$ for all indices except $i = k$, then the set of solutions for Δb_k is an interval around 0 of the form $[-d_k, e_k]$.

The printouts of softwares for the simplex algorithm generally contain information on these intervals.

Example 13.7.1. A person is studying the diet problem

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax \geq b,\ \ x \geq 0 \end{array}$$

in a specific case with six foods and four nutrient requirements. The following computer printout is obtained when $c^T = (1, 2, 3, 4, 1, 6)$ and $b^T = (10, 15, 20, 18)$.

Optimal value: 8.52

Optimal solution:

Food 1: 5.73
Food 2: 0.00
Food 3: 0.93
Food 4: 0.00
Food 5: 0.00
Food 6: 0.00

Download free eBooks at bookboon.com

Sensitivity report

Variable	Value	Objective-coeff.	Allowable decrease	Allowable increase
Food 1:	5.73	1.00	0.14	0.33
Food 2:	0.00	2.00	1.07	∞
Food 3:	0.93	3.00	2.00	0.50
Food 4:	0.00	4.00	3.27	∞
Food 5:	0.00	1.00	0.40	∞
Food 6:	0.00	6.00	5.40	∞

Constraint	Final value	Shadow price	Bounds r.h. side	Allowable decrease	Allowable increase
Nutrient 1:	19.07	0.00	10.00	∞	9.07
Nutrient 2:	31.47	0.00	15.00	∞	16.47
Nutrient 3:	20.00	0.07	20.00	8.00	7.00
Nutrient 4:	18.00	0.40	18.00	4.67	28.67

The sensitivity report shows that the optimal solution remains unchanged as long as the price of food 1 stays in the interval $[5.73 - 0.14, 5.73 + 0.33]$, ceteris paribus. A price change of z units in this range changes the optimal value by $5.73\,z$ units.

A price reduction of food 4 with a maximum of 3.27, or an unlimited price increase of the same food, ceteris paribus, does not affect the optimal solution, nor the optimal value.

The set of price changes that leaves the optimal solution unchanged is a convex set, since it is a polyhedron according to inequality (13.12). The optimal solution of our example is therefore unchanged if for example the prices of foods 1, 2 and 3 are increased by 0.20, 1.20 and 0.10, respectively, because $\Delta c = (0.20,\ 1.20,\ 0.10,\ 0,\ 0,\ 0)$ is a convex combination of allowable increases, since

$$\frac{0.20}{0.33} + \frac{1.20}{\infty} + \frac{0.10}{0.50} \le 1.$$

The sensitivity report also shows how the optimal solution is affected by certain changes in the right-hand side b. The optimal solution remains unchanged, for example, if the need for nutrient 1 would increase from 10 to 15, since the constraint is not binding and the increase 5 is less than the permitted increase 9.07.

The sensitivity report also tells us that the new optimal solution will still be derived from the same basic index set as above, if b_4 is increased by say 20 units from 18 to 38, an increase that is within the scope of the permissible. So in this case, the optimal diet will also only consist of foods 1 and 3, but

Download free eBooks at bookboon.com

the optimal value will increase by $20 \cdot 0.40$ to 16.52 since the shadow price of nutrient 4 is equal to 0.40. $\square$

13.8 The dual simplex algorithm

The simplex algorithm, applied to a problem

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b,\ x \geq 0 \end{array}$$

with a bounded optimal value, starts from a given feasible basic index set α^0 and then generates a finite sequence $(\alpha^k, \overline{x}^k, \overline{y}^k)_{k=0}^p$ of basic index sets α^k, corresponding basic solutions $\overline{x}^k$ and vectors $\overline{y}^k$ with the following properties:

(i) The basic solutions $\overline{x}^k$ are extreme points of the polyhedron

$$X = \{x \in \mathbf{R}^n \mid Ax = b,\ x \geq 0\}$$

of feasible solutions.

(ii) The line segments $[\overline{x}^k, \overline{x}^{k+1}]$ are edges of the polyhedron X.

(iii) The objective function values $(\langle c, \overline{x}^k\rangle)_{k=0}^p$ form a decreasing sequence.

(iv) $\langle b, \overline{y}^k\rangle = \langle c, \overline{x}^k\rangle$ for all k.

Deloitte.

© Deloitte & Touche LLP and affiliated entities.

Discover the truth at www.deloitte.ca/careers

Download free eBooks at bookboon.com

Click on the ad to read more

(v) The algorithm stops after p iterations when the optimality criterion is met, and $\overline{y}^p$ is then an extreme point of the polyhedron

$$Y = \{y \in \mathbf{R}^m \mid A^T y \leq c\}.$$

(vi) $\overline{x}^p$ is an optimal solution to the primal problem, and $\overline{y}^p$ is an optimal solution to the dual problem

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & A^T y \leq c. \end{array}$$

(vii) The vectors $\overline{y}^k$ do not, however, belong to Y for $0 \leq k \leq p-1$.

The optimal solution $\overline{x}^p$ is obtained by moving along edges of the polyhedron X until an extreme point has been reached that also corresponds to an extreme point of the polyhedron Y. Instead, we could move along edges of the polyhedron Y, and this observation leads to the following method for solving the minimization problem.

The dual simplex algorithm

Given a basic index set α such that $z = c - A^T\overline{y} \geq 0$, where $\overline{y} = (A_{*\alpha}^{-1})^T c_\alpha$.

Repeat steps 1–4 until a stop occurs.

1. Compute the basic solution $\overline{x}$ corresponding to α.
2. *Stopping criterion:* **quit** if $\overline{x} \geq 0$.
   Optimal solution: $\overline{x}$. Optimal dual solution: $\overline{y}$.
   Also **quit** if any of the constraint equations has the form $a'_{i1}x_1 + a'_{i2}x_2 + \cdots + a'_{in}x_n = b'_i$ with $b'_i > 0$ and $a'_{ij} \leq 0$ for all j, because then there are no feasible solutions to the primal problem.
3. Generate a new basic index set α' by replacing one of the indices of α in such a way that the new reduced cost vector z' remains nonnegative and $\langle b, \overline{y}'\rangle \geq \langle b, \overline{y}\rangle$, where $\overline{y}' = (A_{*\alpha'}^{-1})^T c_{\alpha'}$.
4. *Update:* $\alpha := \alpha'$, $\overline{y} := \overline{y}'$.

We refrain from specifying the necessary pivoting rules. Instead, we consider a simple example.

Example 13.8.1. We shall solve the minimization problem

$$\begin{array}{ll} \min & x_1 + 2x_2 + 3x_3 \\ \text{s.t.} & \begin{cases} 2x_1 \phantom{{}+2x_2} + x_3 \geq 9 \\ x_1 + 2x_2 \phantom{{}+2x_3} \geq 12 \\ \phantom{x_1+{}} x_2 + 2x_3 \geq 15, \quad x \geq 0 \end{cases} \end{array}$$

Download free eBooks at bookboon.com

by using the dual simplex algorithm, and we begin by reformulating the problem in standard form as follows:

$$\begin{array}{ll} \min & x_1 + 2x_2 + 3x_3 \\ \text{s.t.} & \begin{cases} 2x_1 + x_3 - x_4 = 9 \\ x_1 + 2x_2 - x_5 = 12 \\ \phantom{x_1+{}} x_2 + 2x_3 - x_6 = 15, \ x \geq 0. \end{cases} \end{array}$$

The corresponding simplex tableau now looks like this:

$$\begin{array}{rrrrrr|r} 2 & 0 & 1 & -1 & 0 & 0 & 9 \\ 1 & 2 & 0 & 0 & -1 & 0 & 12 \\ 0 & 1 & 2 & 0 & 0 & -1 & 15 \\ \hline 1 & 2 & 3 & \underline{0} & \underline{0} & \underline{0} & 0 \end{array}$$

For comparison, we also state the corresponding dual maximization problem:

$$\begin{array}{ll} \max & 9y_1 + 12y_2 + 15y_3 \\ \text{s.t.} & \begin{cases} 2y_1 + y_2 \leq 1 \\ \phantom{2y_1+{}} 2y_2 + y_3 \leq 2 \\ y_1 + 2y_3 \leq 3, \ y \geq 0. \end{cases} \end{array}$$

We can start the dual simplex algorithm from the basic index set $\alpha = (4, 5, 6)$, and as usual, we have underlined the basic columns. The corresponding basic solution $\overline{x}$ is not feasible since the coordinates of $\overline{x}_\alpha = (-9, -12, -15)$ are negative. The bottom row $[1\ 2\ 3\ 0\ 0\ 0]$ of the tableau is the reduced cost vector $z^T = c^T - \overline{y}^T A$. The row vector $\overline{y}^T = c_\alpha^T A_{*\alpha}^{-1} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$ can also be read in the bottom row; it is found below the matrix $-E$, and $\overline{y}$ belongs to the polyhedron Y of feasible solutions to the dual problem, since $z^T \geq 0$.

We will now gradually replace one element at a time in the basic index set. As pivot row r, we choose the row that corresponds to the most negative coordinate of $\overline{x}_\alpha$, and in the first iteration, this is the third row in the above simplex tableau. To keep the reduced cost vector nonnegative, we must select as pivot column a column k, where the matrix element a_{rk} is positive and the ratio z_k/a_{rk} is as small as possible. In the above tableau, this is the third column, so we pivot around the element at location $(3, 3)$. This leads to the following tableau:

$$\begin{array}{rrrrrr|r} 2 & -\frac{1}{2} & 0 & -1 & 0 & \frac{1}{2} & \frac{3}{2} \\ 1 & 2 & 0 & 0 & -1 & 0 & 12 \\ 0 & \frac{1}{2} & 1 & 0 & 0 & -\frac{1}{2} & \frac{15}{2} \\ \hline 1 & \frac{1}{2} & \underline{0} & \underline{0} & \underline{0} & \frac{3}{2} & -\frac{45}{2} \end{array}$$

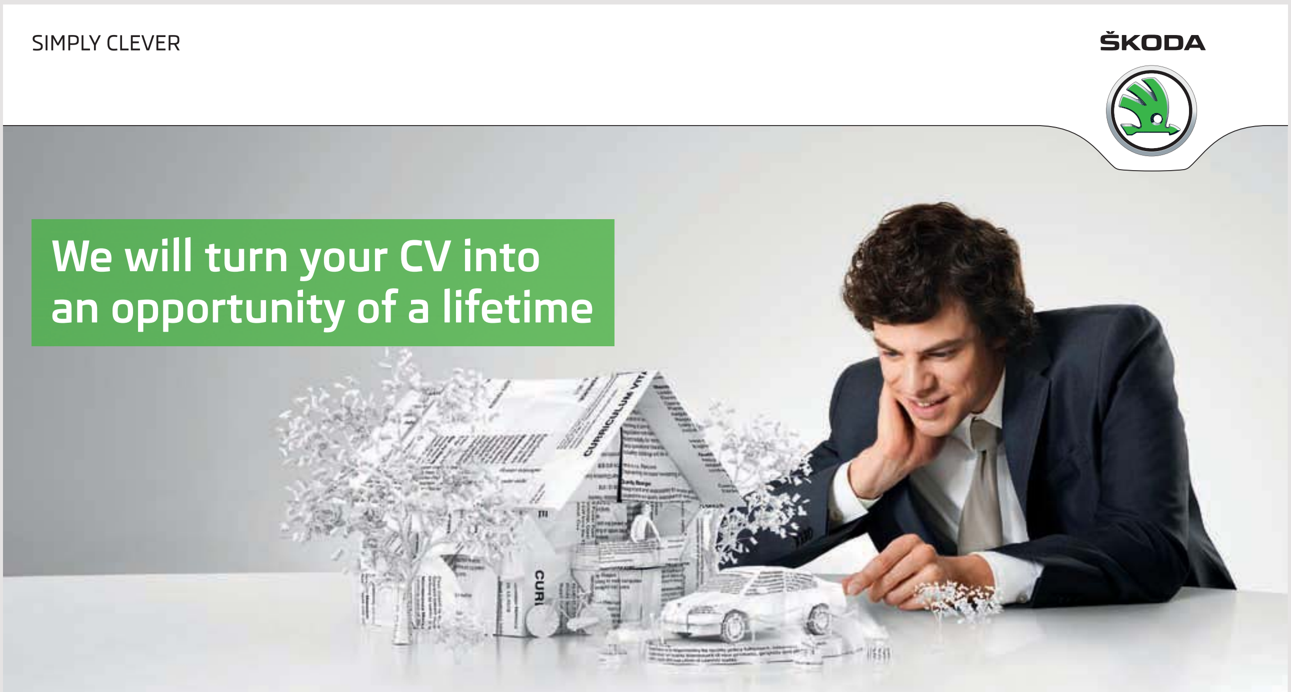
Download free eBooks at bookboon.com

In this new tableau, $\alpha = (4, 5, 3)$, $\overline{x}_\alpha = (-\frac{3}{2}, -12, \frac{15}{2})$ and $\overline{y} = (0, 0, \frac{3}{2})$. The most negative element of $\overline{x}_\alpha$ is to be found in the second row, and the least ratio z_k/a'_{2k} with a positive denominator a'_{2k} is obtained for $k = 2$. Pivoting around the element at location $(2, 2)$ leads to the following simplex tableau:

$$
\begin{array}{cccccc|c}
\frac{9}{4} & 0 & 0 & -1 & -\frac{1}{4} & \frac{1}{2} & \frac{9}{2} \\
\frac{1}{2} & 1 & 0 & 0 & -\frac{1}{2} & 0 & 6 \\
-\frac{1}{4} & 0 & 1 & 0 & \frac{1}{4} & -\frac{1}{2} & \frac{9}{2} \\
\hline
\frac{3}{4} & \underline{0} & \underline{0} & \underline{0} & \frac{1}{4} & \frac{3}{2} & -\frac{51}{2}
\end{array}
$$

Now, $\alpha = (4, 2, 3)$, $\overline{x}_\alpha = (-\frac{9}{2}, 6, \frac{9}{2})$ and $\overline{y} = (0, \frac{1}{4}, \frac{3}{2})$. This time, we should select the element in the first row and the first column as pivot element, which leads to the next tableau.

$$
\begin{array}{cccccc|c}
1 & 0 & 0 & -\frac{4}{9} & -\frac{1}{9} & \frac{2}{9} & 2 \\
0 & 1 & 0 & \frac{2}{9} & -\frac{4}{9} & -\frac{1}{9} & 5 \\
0 & 0 & 1 & -\frac{1}{9} & \frac{2}{9} & -\frac{4}{9} & 5 \\
\hline
\underline{0} & \underline{0} & \underline{0} & \frac{1}{3} & \frac{1}{3} & \frac{4}{3} & -27
\end{array}
$$



Do you like cars? Would you like to be a part of a successful brand?
We will appreciate and reward both your enthusiasm and talent.
Send us your CV. You will be surprised where it can take you.

Send us your CV on
www.employerforlife.com

Click on the ad to read more

Here, $\alpha = (1, 2, 3)$, $\overline{x}_\alpha = (2, 5, 5)$ and $\overline{y} = (\frac{1}{3}, \frac{1}{3}, \frac{4}{3})$, and the optimality criterion is met since $\overline{x}_\alpha \geq 0$. The optimal value is 27 and $(2, 5, 5, 0, 0, 0)$ is the optimal point. The dual maximization problem attains its maximum at $(\frac{1}{3}, \frac{1}{3}, \frac{4}{3})$. The optimal solution to our original minimization problem is of course $x = (2, 5, 5)$. □

13.9 Complexity

How many iterations are needed to solve an LP problem using the simplex algorithm? The answer will depend, of course, on the size of the problem. Experience shows that the number of iterations largely grows linearly with the number of rows m and sublinearly with the number of columns n for realistic problems, and in most real problems, n is a small multiple of m, usually not more than $10m$. The number of iterations is therefore usually somewhere between m and $4m$, which means that the simplex algorithm generally performs very well.

The worst case behavior of the algorithm is bad, however (for all known pivoting rules). Klee and Minty has constructed an example where the number of iterations grows exponentially with the size of the problem.

Example 13.9.1 (Klee and Minty, 1972). Consider the following LP problem in n variables and with n inequality constraints:

$$\begin{aligned} &\max\ 2^{n-1}x_1 + 2^{n-2}x_2 + \cdots + 2x_{n-1} + x_n \\ &\text{s.t.} \begin{cases} x_1 & \leq\ \ 5 \\ 4x_1 + \quad x_2 & \leq\ 25 \\ 8x_1 + \quad 4x_2 + \ x_3 & \leq 125 \\ \vdots & \quad \vdots \\ 2^n x_1 + 2^{n-1}x_2 + \ldots + 4x_{n-1} + x_n & \leq\ 5^n \end{cases} \end{aligned}$$

The polyhedron of feasible solutions has in this case 2^n extreme points.

Suppose that we apply the simplex algorithm to the equivalent standard problem, in each iteration choosing as pivot column the column with the most negative value of the reduced cost. If we start from the feasible basic solution that corresponds to $x = 0$, then we have to go through all the 2^n feasible basic solutions before we finally reach the optimal solution $(0, 0, \ldots, 5^n)$. The number of iterations is therefore equal to 2^n and thus increases exponentially with n. □

An algorithm for solving a problem in n variables is called *strictly polynomial* if there exists a positive integer k such that the number of elementary

Download free eBooks at bookboon.com

arithmetic operations in the algorithm grows with n as at most $O(n^k)$. In many algorithms, the number of operations also depends on the size of the input data. An algorithm is called *polynomial* if the number of arithmetic operations is growing as a polynomial in L, where L is the number of binary bits needed to represent all input (i.e. the matrices A, b and c in linear programming).

Gaussian elimination is a strictly polynomial algorithm, because a system of linear equations with n equations and n unknowns is solved with $O(n^3)$ arithmetic operations.

Klee–Minty's example and other similar examples demonstrate that the simplex algorithm is not strictly polynomial. But all experience shows that the simplex algorithm works very well, even if the worst case behavior is bad. This is also supported by probabilistic analyzes, made by Borgwardt (1987), Smale (1983), Adler and Megiddo (1985), among others. Such an analysis shows, for example, that (a variant of) the simplex algorithm, given a certain special probability distribution of the input data, on average converges after $O(m^2)$ iterations, where m is the number of constraints.

The existence of a polynomial algorithm that solves LP problems (with rational coefficients as input data) was first demonstrated in 1979 by Leonid Khachiyan. His so-called *ellipsoid algorithm* reduces LP problems to the problem of finding a solution to a system $Ax > b$ of strict inequalities with a bounded set of solutions, and the algorithm generates a sequence of shrinking ellipsoids, all guaranteed to contain all the solutions to the system. If the center of an ellipsoid satisfies all inequalities of the system, then a solution has been found, of course. Otherwise, the process stops when a generated ellipsoid has too small volume to contain all solutions, if there are any, with the conclusion that there are no solutions.

LP problems in standard form with n variables and input size L are solved by the ellipsoid method in $O(n^4L)$ arithmetic operations. However, in spite of this nice theoretical result, it was soon clear that the ellipsoid method could not compete with the simplex algorithm on real problems of moderate size due to slow convergence. (The reason for this is, of course, that the implicit constant in the O-estimate is very large.)

A new polynomial algorithm was discovered in 1984 by Narendra Karmarkar. His algorithm generates a sequence of points, which lie in the interior of the set of feasible points and converge towards an optimal point. The algorithm uses repeated centering of the generated points by a projective scaling transformation. The theoretical complexity bound of the original version of the algorithm is also $O(n^4L)$.

Karmarkar's algorithm turned out to be competitive with the simplex algorithm on practical problems, and his discovery was the starting point for

Download free eBooks at bookboon.com

an intensive development of alternative interior point methods for LP problems and more general convex problems. We will study such an algorithm in Chapter 18.

It is still an open problem whether there exists any strictly polynomial algorithm for solving LP problems.

Exercises

13.1 Write the following problems in standard form.

a) $$\begin{array}{ll} \min & 2x_1 - 2x_2 + x_3 \\ \text{s.t.} & \begin{cases} x_1 + x_2 - x_3 \geq 3 \\ x_1 + x_2 - x_3 \leq 2 \\ x_1, x_2, x_3 \geq 0 \end{cases} \end{array}$$

b) $$\begin{array}{ll} \min & x_1 + 2x_2 \\ \text{s.t.} & \begin{cases} x_1 + x_2 \geq 1 \\ \phantom{x_1 + {}} x_2 \geq -2 \\ x_1 \phantom{{} + x_2} \geq 0. \end{cases} \end{array}$$

13.2 Find all nonnegative basic solutions to the following systems of equations.

a) $$\begin{cases} 5x_1 + 3x_2 + x_3 = 40 \\ x_1 + x_2 + x_3 = 10 \end{cases}$$

b) $$\begin{cases} x_1 - 2x_2 - x_3 + x_4 = 3 \\ 2x_1 + 5x_2 - 3x_3 + 2x_4 = 6. \end{cases}$$

Click on the ad to read more

13.3 State the dual problem to

$$\begin{array}{ll} \min & x_1 + x_2 + 4x_3 \\ \text{s.t.} & \begin{cases} x_1 - x_3 = 1 \\ x_1 + 2x_2 + 7x_3 = 7, \ x \geq 0 \end{cases} \end{array}$$

and prove that $(1,3,0)$ is an optimal solution and that $(\frac{1}{2},\frac{1}{2})$ is an optimal solution to the dual problem.

13.4 Solve the following LP problems using the simplex algorithm.

a) $$\begin{array}{ll} \min & -x_4 \\ \text{s.t.} & \begin{cases} x_1 + x_4 = 1 \\ x_2 + 2x_4 = 2 \\ x_3 - x_4 = 3, \ x \geq 0 \end{cases} \end{array}$$

b) $$\begin{array}{ll} \max & 2x_1 - x_2 + x_3 - 3x_4 + x_5 \\ \text{s.t.} & \begin{cases} x_1 + 2x_4 - x_5 = 15 \\ x_2 + x_4 + x_5 = 12 \\ x_3 - 2x_4 + x_5 = 9, \ x \geq 0 \end{cases} \end{array}$$

c) $$\begin{array}{ll} \max & 15x_1 + 12x_2 + 14x_3 \\ \text{s.t.} & \begin{cases} 3x_1 + 2x_2 + 5x_3 \leq 6 \\ x_1 + 3x_2 + 3x_3 \leq 3 \\ 5x_3 \leq 4, \ x \geq 0 \end{cases} \end{array}$$

d) $$\begin{array}{ll} \max & 2x_1 + x_2 + 3x_3 + x_4 + 2x_5 \\ \text{s.t.} & \begin{cases} x_1 + 2x_2 + x_3 + x_5 \leq 10 \\ x_2 + x_3 + x_4 + x_5 \leq 8 \\ x_1 + x_3 + x_4 \leq 5, \ x \geq 0 \end{cases} \end{array}$$

e) $$\begin{array}{ll} \min & x_1 - 2x_2 + x_3 \\ \text{s.t.} & \begin{cases} x_1 + x_2 - 2x_3 \leq 3 \\ x_1 - x_2 + x_3 \leq 2 \\ -x_1 - x_2 + x_3 \leq 0, \ x \geq 0 \end{cases} \end{array}$$

f) $$\begin{array}{ll} \min & x_1 - x_2 + 2x_3 - 3x_4 \\ \text{s.t.} & \begin{cases} 2x_1 + 3x_2 + x_3 = 2 \\ x_1 + 3x_2 + x_3 + 5x_4 = 4, \ x \geq 0. \end{cases} \end{array}$$

13.5 Carry out in detail all the steps of the simplex algorithm for the problem

$$\begin{array}{ll} \min & -x_2 + x_4 \\ \text{s.t.} & \begin{cases} x_1 + x_4 + x_5 = 1 \\ x_2 - 2x_4 - x_5 = 1 \\ x_3 + 2x_4 + x_5 = 3, \ x \geq 0. \end{cases} \end{array}$$

Is the optimal solution unique?

Download free eBooks at bookboon.com

13.6 Use artificial variables to solve the LP problem

$$\begin{array}{ll} \max & x_1 + 2x_2 + 3x_3 - x_4 \\ \text{s.t.} & \begin{cases} x_1 + 2x_2 + 3x_3 & = 15 \\ 2x_1 + x_2 + 5x_3 & = 20 \\ x_1 + 2x_2 + x_3 + x_4 = 10, & x \geq 0. \end{cases} \end{array}$$

13.7 Use the simplex algorithm to show that the following systems of equalities and inequalities are consistent.

a) $\begin{cases} 3x_1 + x_2 + 2x_3 + x_4 + x_5 = 2 \\ 2x_1 - x_2 + x_3 + x_4 + 4x_5 = 3, \ x \geq 0 \end{cases}$

b) $\begin{cases} x_1 - x_2 + 2x_3 + x_4 \geq 6 \\ -2x_1 + x_2 - 2x_3 + 7x_4 \geq 1 \\ x_1 - x_2 + x_3 - 3x_4 \geq -1, \ x \geq 0. \end{cases}$

13.8 Solve the LP problem

$$\begin{array}{ll} \min & x_1 + 2x_2 + 3x_3 \\ \text{s.t.} & \begin{cases} 2x_1 + x_3 \geq 3 \\ x_1 + 2x_2 \geq 4 \\ x_2 + 2x_3 \geq 5, \ x \geq 0. \end{cases} \end{array}$$

13.9 Write the following problem in standard form and solve it using the simplex algorithm.

$$\begin{array}{ll} \min & 8x_1 - x_2 \\ \text{s.t.} & \begin{cases} 3x_1 + x_2 \geq 1 \\ x_1 - x_2 \leq 2 \\ x_1 + 2x_2 = 20, \ x \geq 0. \end{cases} \end{array}$$

13.10 Solve the following LP problems using the dual simplex algorithm.

a) $\begin{array}{ll} \min & 2x_1 + x_2 + 3x_3 \\ \text{s.t.} & \begin{cases} x_1 + x_2 + x_3 \geq 2 \\ 2x_1 - x_2 \geq 1 \\ x_2 + 2x_3 \geq 2, \ x \geq 0 \end{cases} \end{array}$

b) $\begin{array}{ll} \min & x_1 + 2x_2 \\ \text{s.t.} & \begin{cases} x_1 - 2x_3 \geq -5 \\ -2x_1 + 3x_2 - x_3 \geq -4 \\ -2x_1 + 5x_2 - x_3 \geq 2, \ x \geq 0 \end{cases} \end{array}$

c) $\begin{array}{ll} \min & 3x_1 + 2x_2 + 4x_3 \\ \text{s.t.} & \begin{cases} 4x_1 + 2x_3 \geq 5 \\ x_1 + 3x_2 + 2x_3 \geq 4, \ x \geq 0. \end{cases} \end{array}$

Download free eBooks at bookboon.com

13.11 Suppose $b_2 \geq b_1 \geq 0$. Show that $\overline{x} = \big(b_1, \frac{1}{2}(b_2 - b_1), 0\big)$ is an optimal solution to the problem

$$\begin{array}{ll} \min & x_1 + x_2 + 4x_3 \\ \text{s.t.} & \begin{cases} x_1 \phantom{{}+2x_2} - x_3 = b_1 \\ x_1 + 2x_2 + 7x_3 = b_2, \ x \geq 0. \end{cases} \end{array}$$

13.12 Investigate how the optimal solution to the LP problem

$$\begin{array}{ll} \max & 2x_1 + tx_2 \\ \text{s.t.} & \begin{cases} x_1 + x_2 \leq 5 \\ 2x_1 + x_2 \leq 7, \ x \geq 0 \end{cases} \end{array}$$

varies as the real parameter t varies.

13.13 A shoe manufacturer produces two shoe models A and B. Due to limited supply of leather, the manufactured number of pairs x_A and x_B of the two models must satisfy the inequalities

$$x_A \leq 1000, \quad 4x_A + 3x_B \leq 4100, \quad 3x_A + 5x_B \leq 5000.$$

Download free eBooks at bookboon.com

Click on the ad to read more

The sale price of A and B is 500 SEK and 350 SEK, respectively per pair. It costs 200 SEK to manufacture a pair of shoes of model B. However, the cost of producing a pair of shoes of model A is uncertain due to malfunctioning machines, and it can only be estimated to be between 300 SEK and 410 SEK. Show that the manufacturer may nevertheless decide how many pairs of shoes he shall manufacture of each model to maximize his profit.

13.14 Joe wants to meet his daily requirements of vitamins P, Q and R by only living on milk and bread. His daily requirement of vitamins is 6, 12 and 4 mg, respectively. A liter of milk costs 7.50 SEK and contains 2 mg of P, 2 mg of Q and nothing of R; a loaf of bread costs 20 SEK and contains 1 mg of P, 4 mg of Q and 4 mg of R. The vitamins are not toxic, so a possible overdose does not harm. Joe wants to get away as cheaply as possible. Which daily bill of fare should he choose? Suppose that the price of milk begins to rise. How high can it be without Joe having to change his bill of fare?

13.15 Using the assumptions of Lemma 13.4.1, show that the reduced cost z_k is equal to the direction derivative of the objective function $\langle c, x\rangle$ in the direction $-v$.

13.16 This exercise outlines an alternative method to prevent cycling in the simplex algorithm. Consider the problem

$$\text{(P)} \qquad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b,\ x \geq 0 \end{array}$$

and let α be an arbitrary feasible basic index set with corresponding basic solution $\overline{x}$. For each positive number ϵ, we define new vectors $\overline{x}(\epsilon) \in \mathbf{R}^n$ and $b(\epsilon) \in \mathbf{R}^m$ as follows:

$$\begin{aligned} &\overline{x}(\epsilon)_\alpha = \overline{x}_\alpha + (\epsilon, \epsilon^2, \dots, \epsilon^m) \quad \text{and} \quad \overline{x}(\epsilon)_j = 0 \text{ for all } j \notin \alpha, \\ &b(\epsilon) = A\overline{x}(\epsilon). \end{aligned}$$

Then $\overline{x}(\epsilon)$ is obviously a nonnegative basic solution to the system $Ax = b(\epsilon)$ with α as the corresponding basic index set, and the coordinates of the vector $b(\epsilon)$ are polynomials of degree m in the variable ϵ.

a) Prove that all basic solutions to the system $Ax = b(\epsilon)$ are non-degenerate except for finitely many numbers $\epsilon > 0$. Consequently, there is a number $\epsilon_0 > 0$ so that all basic solution are non-degenerate if $0 < \epsilon < \epsilon_0$.

b) Prove that if $0 < \epsilon < \epsilon_0$, then all feasible basic index sets for the problem

$$(\mathrm{P}_\epsilon) \qquad \begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b(\epsilon),\ x \geq 0 \end{array}$$

are also feasible basic index sets for the original problem (P).

Download free eBooks at bookboon.com

c) The simplex algorithm applied to problem (P_ϵ) will therefore stop at a feasible basic index set β, which is also feasible for problem (P), provided ϵ is a sufficiently small number. Prove that β also satisfies the stopping condition for problem (P).

Cycling can thus be avoided by the following method: Perturb the right-hand side by forming $\overline{x}(\epsilon)$ and the column matrix $b(\epsilon)$, where ϵ is a small positive number. Use the simplex algorithm on the perturbed problem. The algorithm stops at a basic index set β. The corresponding unperturbed problem stops at the same basic index set.

13.17 Suppose that $\mathcal{A}$ is a polynomial algorithm for solving systems $Cx \geq b$ of linear inequalities. When applied to a solvable system, the algorithm finds a solution $\overline{x}$ and stops with the output $\mathcal{A}(C, b) = \overline{x}$. For unsolvable systems, it stops with the output $\mathcal{A}(C, b) = \emptyset$. Use the algorithm $\mathcal{A}$ to construct a polynomial algorithm for solving arbitrary LP problems

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax \geq b,\ x \geq 0. \end{array}$$

13.18 Perform all the steps of the simplex algorithm for the example of Klee and Minty when $n = 3$.

Download free eBooks at bookboon.com

Click on the ad to read more

Bibliografical and historical notices

The theory of convex programs has its roots in a paper by Kuhn–Tucker [1], which deals with necessary and sufficient conditions for optimality in nonlinear problems. Kuhn–Tucker noted the connection between Lagrange multipliers and saddle points, and they focused on the role of convexity. A related result with Lagrange multiplier conditions had otherwise been shown before by John [1] for general differentiable constraints, and KKT conditions are present for the first time in an unpublished master's thesis by Karush [1]. Theorem 11.2.1 can be found in Uzawa [1].

The duality theorem in linear programming was known as a result of game theory by John von Neumann, but the first published proof of this theorem appears in Gale–Kuhn–Tucker [1].

The earliest known example of linear programming can be found in works by Fourier [1] from the 1820s and deals with the problem of determining the best, with respect to the maximum norm, fit to an overdetermined system of linear equations. Fourier reduced this problem to minimizing a linear form over a polyhedron, and he also hinted a method, equivalent to the simplex algorithm, to compute the minimum.

It was to take until the 1940s before practical problems on a larger scale began to be formulated as linear programming. The transportation problem was formulated by Hitchcock [1], who also gave a constructive solution method, and the diet problem was studied by Stigler [1], who, however, failed to compute the exact solution. The Russian mathematician and economist Kantorovich [1] had some years before formulated and solved LP problems in production planning, but his work was not noticed outside the USSR and would therefore not influence the subsequent development.

The need for mathematical methods for solving military planning problems had become apparent during the Second World War, and in 1947 a group of mathematicians led by George Dantzig and Marshall Wood worked at the U.S. Department of the Air Force with such problems. The group's

Download free eBooks at bookboon.com

work resulted in the realization of the importance of linear programming, and the first version of the simplex algorithm was described by Dantzig [1] and Wood–Dantzig [1].

The simplex algorithm is contemporary with the first computers, and this suddenly made it possible to treat large problems numerically and contributed to a breakthrough for linear programming. A conference on linear programming, arranged by Tjalling Koopmans 1949 in Chicago, was also an important step in the popularization of linear programming. During this conference, papers on linear programming were presented by economists, mathematicians, and statisticians. The papers were later published in Koopmans [1], and this book became the start for a rapidly growing literature on linear programming.

Dantzig's [2] basic article 1951 treated the non-degenerate case of the simplex algorithm, and the possibility of cycling in the degenerate case caused initially some concern. The first example with cycling was constructed by Hoffman [1], but even before this discovery Charnes [1] had proposed a method for avoiding cycling. Other such methods were then given by Dantzig–Orden–Wolfe [1] and Wolfe [2]. Bland's [1] simple pivoting rule is relatively recent.

It is easy to modify the simplex algorithm so that it is directly applicable to LP problems with bounded variables, which was first noted by Charnes–Lemke [1] and Dantzig [3].

The dual simplex algorithm was developed by Beale [1] and Lemke [1]. The currently most efficient variants of the simplex algorithm are primal-dual algorithms.

Convex quadratic programs can be solved by a variant of the simplex algorithm, formulated by Wolfe [1].

Khachiyan's [1] complexity results was based on the ellipsoid algorithm, which was first proposed by Shor [1] as a method in general convex optimization. See Bland–Goldfarb–Todd [1] for an overview of the ellipsoid method.

Many variants of Karmarkar's [1] algorithm were developed after his publication in 1984. Algorithms for LP problems with $O(n^3L)$ as complexity bound are described by Gonzaga [1] and Ye [1].

There are numerous textbooks on linear programming. Two early such books, written by pioneers in the field, are Dantzig [4], which in addition to the mathematical material also contains a thorough historical overview, many applications and an extensive bibliography, and Gale [1], which provides a concise but mathematically rigorous presentation of linear programming with an emphasis on economic applications. More recent books are Chvatal [1] and Luenberger [1].

Download free eBooks at bookboon.com

References

Adler, I. & Megiddo, N.

[1] A simplex algorithm whose average number of steps is bounded between two quadratic functions of the smaller dimension. *J. ACM* 32 (1985), 871–895.

Beale, E.M.L

[1] An alternative method of linear programming, *Proc. Cambridge Philos. Soc.* 50 (1954), 513–523.

Bland, R.G.

[1] New Finite Pivoting Rules for the Simplex Method, *Math. Oper. Res.* 2 (1977), 103–107.

Bland, R.G., Goldfarb, D. & Todd, M.J.

[1] The ellipsoid method: A survey, *Oper. Res.* 29 (1981), 1039–1091.

Borgwardt, K.H.

[1] *The Simplex Method – a probabilistic analysis,* Springer-Verlag, 1987.

Boyd, S. & Vandenberghe, L.

[1] *Convex Optimization*, Cambridge Univ. Press, Cambridge, UK, 2004.

Download free eBooks at bookboon.com
Click on the ad to read more

Charnes, A.
[1] Optimality and Degeneracy in Linear Programming, *Econometrica* 20 (1952), 160–170.

Charnes, A. & Lemke, C.E.
[1] *The bounded variable problem.* ONR memorandum 10, Carnegie Institute of Technology, 1954.

Chvátal, V.
[1] *Linear Programming.* W.H. Freeman, 1983.

Dantzig, G.B.
[1] Programming of Interdependent Activities. II. Mathematical Model, *Econometrica* 17 (1949), 200–211.
[2] Maximization of Linear Functions of Variables Subject to Linear Inequalities. Pages 339–347 in T.C. Koopmans (ed.), *Activity Analysis of Production and Allocation*, John Wiley, 1951.
[3] Upper Bounds, Secondary Constraints and Block Triangularity in Linear Programming, *Econometrica* 23 (1955), 174–183.
[4] *Linear Programming and Extensions.* Princeton University Press, 1963.

Dantzig, G.B., Orden, A. & Wolfe, P.
[1] The genereralized simplex method for minimizing a linear form under linear inequality constraints, *Pacific J. Math.* 5 (1955), 183–195.

Fourier, J.-B.
[1] Solution d'une question particulière du calcul des inégalités. Sid 317–328 i *Oeuvres de Fourier II*, 1890.

Gale, D.
[1] *The Theory of Linear Economic Models.* McGraw–Hill, 1960.

Gale, D., Kuhn, H.W. & Tucker, A.W.
[1] Linear programming and the theory of games. Pages 317–329 in Koopmans, T.C. (ed.), *Activity Analysis of Production and Allocation*, John Wiley & Sons, 1951.

Gonzaga, C.C.
[1] An algorithm for solving linear programming problems in $O(n^3L)$ operations. Pages 1–28 in Megiddo, N. (ed.), *Progress in Mathematical Programming: Interior-Point and Related Methods,* Springer-Verlag, 1988.

Hitchcock, F.L.
[1] The distribution of a product from several socurces to numerous localities, *J. Math. Phys.* 20 (1941), 224–230.

Hoffman, A. J.
[1] *Cycling in the Simplex Algorithm.* Report No. 2974, National Bureau of Standards, Gaithersburg, MD, USA, 1953.

Download free eBooks at bookboon.com

John, F.
[1] Extremum problems with inequalities as subsidiary conditions, 1948. Sid 543–560 i Moser J. (ed.), *Fritz John, Collected Papers*, Birkhäuser Verlag, 1985.

Kantorovich, L.V.
[1] Mathematical methods of organizing and planning production, Leningrad State Univ. (in Russian), 1939. Engl. translation in *Management Sci.* 6 (1960), 366-422.

Karmarkar, N.
[1] A new polynomial-time algorithm for linear programming, *Combinatorica* 4 (1984), 373–395.

Karush, W.
[1] *Minima of Functions of Several Variables with Inequalities as Side Constraints.* M. Sc. Disseratation. Dept. of Mathematics, Univ. of Chicago, Chicago, Illinois, 1939.

Khachiyan, L.G.
[1] A polynomial algorithm in linear programming, *Dokl. Akad. Nauk SSSR* 244 (1979), 1093–1096. Engl. translation in *Soviet Math. Dokl.* 20 (1979), 191-194.

Klee, V. & Minty, G.J.
[1] How Good is the Simplex Algorithm? Pages 159–175 in Shisha, O. (ed.), *Inequalities, III,* Academic Press, 1972.

Koopmans, T.C., ed.
[1] *Activity Analysis of Production and Allocation.* John Wiley & Sons, 1951.

Kuhn, H.W. & Tucker, A.W.
[1] Nonlinear programming. Pages 481–492 in *Proc. of the second Berkeley Symposium on Mathematical Statistics and Probability.* Univ. of California Press, 1951.

Lemke, C.E.
[1] The dual method of solving the linear programming problem, *Naval Res. Logist. Quart.* 1 (1954), 36–47.

Luenberger, D.G.
[1] *Linear and Nonlinear Programming.* Addison–Wesley, 1984

Shor, N.Z.
[1] Utilization of the operation of space dilation in the minimization of convex functions, *Cybernet. System Anal.* 6 (1970), 7–15.

Download free eBooks at bookboon.com

Smale, S.

[1] On the average number of steps in the simplex method of linear programming, *Math. Program.* 27 (1983), 241–262.

Stigler, G.J.

[1] The Cost of Subsistence, *J. Farm Econ.* 27 (1945), 303–314.

Uzawa, H.

[1] The Kuhn–Tucker theorem in concave programming. In Arrow, K.J., Hurwicz, L. & H. Uzawa, H. (eds.), *Studies in Linear and Non-linear Programming,* Stanford Univ. Press, 1958.

Wolfe, P.

[1] The Simplex Method for Quadratic Programming, *Econometrica* 27 (1959), 382–398.

[2] A Technique for Resolving Degeneracy in Linear Programming, *J. of the Soc. for Industrial and Applied Mathematics* 11 (1963), 205–211.

Wood, M.K. & Dantzig, G.B.

[1] Programming of Interdependent Activities. I. General discussion, *Econometrica* 17 (1949), 193–199.

Ye, Y.

[1] An $O(n^3L)$ potential reduction algorithm for linear programming. *Math. Program.* 50 (1991), 239–258.

Click on the ad to read more

Download free eBooks at bookboon.com

Answers and solutions to the exercises

Chapter 9

9.1 $$\begin{aligned} &\min\ 5000x_1 + 4000x_2 + 3000x_3 + 4000x_4 \\ &\text{s.t.} \begin{cases} -x_1 + 2x_2 + 2x_3 + x_4 \geq 16 \\ 4x_1 + x_2 + 2x_4 \geq 40 \\ 3x_1 + x_2 + 2x_3 + x_4 \geq 24,\ x \geq 0 \end{cases} \end{aligned}$$

9.2 $$\begin{aligned} &\max\ v \\ &\text{s.t.} \begin{cases} 2x_1 + x_2 - 4x_3 \geq v \\ x_1 + 2x_2 - 2x_3 \geq v \\ -2x_1 - x_2 + 2x_3 \geq v \\ x_1 + x_2 + x_3 = 1,\ x \geq 0 \end{cases} \end{aligned}$$

9.3 The row player should choose row number 2 and the column player column number 1.

9.4 Payoff matrix:

	Sp E	Ru E	Ru 2
Sp E	−1	1	−1
Ru E	1	−1	−2
Sp 2	−1	2	2

The column players problem can be formulated as

$$\begin{aligned} &\min\ u \\ &\text{s.t.} \begin{cases} -y_1 + y_2 + y_3 \leq u \\ y_1 - y_2 - 2y_3 \leq u \\ -y_1 + 2y_2 + 2y_3 \leq u \\ y_1 + y_2 + y_3 = 1,\ y \geq 0 \end{cases} \end{aligned}$$

9.5 a) $(\frac{4}{5}, \frac{13}{15})$

Download free eBooks at bookboon.com

9.6 a) $\max\ r$
$$\text{s.t.}\begin{cases}-x_1+\ \ x_2+r\sqrt{2}\le 0\\ \ \ x_1-2x_2+r\sqrt{5}\le 0\\ \ \ x_1+\ \ x_2+r\sqrt{2}\le 1\end{cases}$$

b) $\max\ r$
$$\text{s.t.}\begin{cases}-x_1+\ \ x_2+2r\le 0\\ \ \ x_1-2x_2+3r\le 0\\ \ \ x_1+\ \ x_2+2r\le 1\end{cases}$$

Chapter 10

10.1 $\phi(\lambda)=2\lambda-\frac{1}{2}\lambda^2$

10.2 The dual functions ϕ_a and ϕ_b of the two problems are given by:
$$\phi_a(\lambda)=0 \text{ for all } \lambda\ge 0 \ \text{ and } \ \phi_b(\lambda)=\begin{cases}0 & \text{if } \lambda=0,\\ \lambda-\lambda\ln\lambda & \text{if } 0<\lambda<1,\\ 1 & \text{if } \lambda\ge 1.\end{cases}$$

10.5 The inequality $g_i(x_0)\ge g_i(\hat{x})+\langle g_i'(\hat{x}),x_0-\hat{x}\rangle=\langle g_i'(\hat{x}),x_0-\hat{x}\rangle$ holds for all $i\in I(\hat{x})$. It follows that $\langle g_i'(\hat{x}),\hat{x}-x_0\rangle\ge -g_i(x_0)>0$ for $i\in I_{\text{oth}}(\hat{x})$, and $\langle g_i'(\hat{x}),\hat{x}-x_0\rangle\ge -g_i(x_0)\ge 0$ for $i\in I_{\text{aff}}(\hat{x})$.

10.6 a) $v_{\min}=-1$ for $x=(-1,0)$ b) $v_{\max}=2+\frac{\pi}{4}$ for $x=(1,1)$
c) $v_{\min}=-\frac{1}{3}$ for $x=\pm(\frac{2}{\sqrt{6}},-\frac{1}{\sqrt{6}})$ d) $v_{\max}=\frac{1}{54}$ for $x=(\frac{1}{6},\,2,\,\frac{1}{3})$

Chapter 11

11.1 $\hat{\lambda}=2b$

11.3 b) Let $L\colon\Omega\times\Lambda\to\mathbf{R}$ and $L_1\colon(\mathbf{R}\times\Omega)\times(\mathbf{R}_+\times\Lambda)\to\mathbf{R}$ be the Lagrange functions of the problems (P) and (P′), respectively, and let ϕ and ϕ_1 be the corresponding dual functions. The two Lagrange functions are related as follows:
$$L_1(t,x,\lambda_0,\lambda)=(1-\lambda_0)(t-f(x))+L(x,\lambda).$$
The Lagrange function L_1 is for fixed $(\lambda_0,\lambda)\in\mathbf{R}_+\times\Lambda$ bounded below if and only if $\lambda_0=1$ and $\lambda\in\operatorname{dom}\phi$. Hence, $\operatorname{dom}\phi_1=\{1\}\times\operatorname{dom}\phi$. Moreover, $\phi_1(1,\lambda)=\phi(\lambda)$ for all $\lambda\in\operatorname{dom}\phi$.

11.4 Let I be the index set of all non-affine constraints, and let k be the number of elements of I. Slater's condition is satisfied by the point $\overline{x}=k^{-1}\sum_{i\in I}\overline{x}_i$.

11.5 Let $b^{(1)}$ and $b^{(2)}$ be two points in U, and let $0<\lambda<1$. Choose, given $\epsilon>0$, feasible points $x^{(i)}$ for the problems $(\mathrm{P}_{b^{(i)}})$ so that $f(x^{(i)})<v_{\min}(b^{(i)})+\epsilon$. The point $x=\lambda x^{(1)}+(1-\lambda)x^{(2)}$ is feasible for the problem (P_b), where $b=\lambda b^{(1)}+(1-\lambda)b^{(2)}$. Therefore,
$$\begin{aligned}v_{\min}(\lambda b^{(1)}+(1-\lambda)b^{(2)})&\le f(x)\le\lambda f(x^{(1)})+(1-\lambda)f(x^{(2)})\\&<\lambda v_{\min}(b^{(1)})+(1-\lambda)v_{\min}(b^{(2)})+\epsilon,\end{aligned}$$

Download free eBooks at bookboon.com

and since $\epsilon > 0$ is arbitrary, this shows that the function $v_{\min}$ is convex on U.

11.6 a) $v_{\min} = 2$ for $x = (0,0)$ b) $v_{\min} = 2$ for $x = (0,0)$
c) $v_{\min} = \ln 2 - 1$ for $x = (-\ln 2, \frac{1}{2})$ d) $v_{\min} = -5$ for $x = (-1,-2)$
e) $v_{\min} = 1$ for $x = (1,0)$ f) $v_{\min} = 2\,\mathrm{e}^{1/2} + \frac{1}{4}$ for $x = (\frac{1}{2}, \frac{1}{2})$

11.7 $v_{\min} = 2 - \ln 2$ for $x = (1,1)$

11.9 min $50x_1^2 + 80x_1x_2 + 40x_2^2 + 10x_3^2$

$$\text{s.t.}\begin{cases} 0.2x_1 + 0.12x_2 + 0.04x_3 \geq 0.12 \\ \;\; x_1 + \quad x_2 + \quad x_3 = 1, \quad x \geq 0 \end{cases}$$

Optimum for $x_1 = x_3 = 0.5$ miljon dollars.

Chapter 12

12.1 All nonempty sets $X(b) = \{x \mid Ax \geq b\}$ of feasible points have the same recession cone, since $\operatorname{recc} X(b) = \{x \mid Ax \geq 0\}$ if $X(b) \neq \emptyset$. Therefore, it follows from Theorem 12.1.1 that the optimal value $v(b)$ is finite if $X(b) \neq \emptyset$. The convexity of the optimal value function v is a consequence of the same theorem, because

$$v(b) = \min\{\langle -b, y\rangle \mid A^T y \leq c,\, y \geq 0\},$$

according to the duality theorem.

12.2 E.g. min $x_1 - x_2$ and max $y_1 + y_2$

$$\text{s.t.}\begin{cases} -x_1 \geq 1 \\ x_2 \geq 1,\; x \geq 0 \end{cases} \qquad \text{s.t.}\begin{cases} -y_1 \leq 1 \\ y_2 \leq -1,\; y \geq 0 \end{cases}$$

12.5 $$v_{\max} = \begin{cases} \dfrac{t-3}{t+1} & \text{for } x = \Big(-\dfrac{2}{t+1}, \dfrac{t-1}{t+1}\Big) \text{ if } t < -2, \\ 5 & \text{for } x = (2,3) \text{ if } t \geq -2. \end{cases}$$

Chapter 13

13.1 a) min $2x_1 - 2x_2 + x_3$

$$\text{s.t.}\begin{cases} x_1 + x_2 - x_3 - s_1 = 3 \\ x_1 + x_2 - x_3 + s_2 = 2 \\ x_1, x_2, x_3, s_1, s_2 \geq 0 \end{cases}$$

b) min $x_1 + 2x_2' - 2x_2''$

$$\text{s.t.}\begin{cases} x_1 + x_2' - x_2'' - s_1 = 1 \\ x_2' - x_2'' - s_2 = -2 \\ x_1, x_2', x_2'', s_1, s_2 \geq 0 \end{cases}$$

13.2 a) $(5,5,0)$ and $(7\frac{1}{2}, 0, 2\frac{1}{2})$ b) $(3,0,0,0)$ and $(0,0,0,3)$

Download free eBooks at bookboon.com

13.3 $\max\ y_1 + 7y_2$

$$\text{s.t.} \begin{cases} y_1 + \ \ y_2 \le 1 \\ \qquad 2y_2 \le 1 \\ -y_1 + 7y_2 \le 4 \end{cases}$$

13.4 a) $v_{\min} = -1$ for $x = (0,0,4,1)$ b) $v_{\max} = 56$ for $x = (24,0,0,1,11)$
c) $v_{\max} = 30\frac{6}{7}$ for $x = (1\frac{5}{7}, \frac{3}{7}, 0)$ d) $v_{\max} = 23$ for $x = (2,0,3,0,5)$
e) $v_{\min} = -\infty$ f) $v_{\min} = -1\frac{13}{15}$ for $x = (0, \frac{2}{3}, 0, \frac{2}{5})$

13.5 $v_{\min} = -2$ is attained at all points on the line segment between the points $(0,3,1,1,0)$ and $(0,2,2,0,1)$.

13.6 $v_{\max} = 15$ for $x = (2\frac{1}{2}, 2\frac{1}{2}, 2\frac{1}{2}, 0)$

13.8 $v_{\min} = 9$ for $x = (\frac{2}{3}, 1\frac{2}{3}, 1\frac{2}{3})$

13.9 $v_{\min} = -40\frac{3}{5}$ for $x = (-3\frac{3}{5}, 11\frac{4}{5})$

13.10 a) $v_{\min} = 4\frac{1}{4}$ for $x = (\frac{3}{4}, \frac{1}{2}, \frac{3}{4})$ b) $v_{\min} = \frac{4}{5}$ for $x = (0, \frac{2}{5}, 0)$
c) $v_{\min} = 5\frac{7}{12}$ for $x = (1\frac{1}{4}, \frac{11}{12}, 0)$

13.12 $$v_{\max} = \begin{cases} 7 & \text{for } x = (3\frac{1}{2}, 0) \text{ if } t \le 1, \\ 4 + 3t & \text{for } x = (2,3) \text{ if } 1 < t < 2, \\ 5t & \text{for } x = (0,5) \text{ if } t \ge 2. \end{cases}$$

13.13 500 pairs of model A and 700 pairs of model B.

13.14 4 liters of milk and 1 loaf. The milk price could rise to 10 SEK/l.

13.17 First, use the algorithm $\mathcal{A}$ on the system consisting of the linear inequalities $Ax \ge b$, $x \ge 0$, $A^T y \le c$, $y \ge 0$, $\langle c, x\rangle \le \langle b, y\rangle$. If the algorithm delivers a solution $(\overline{x}, \overline{y})$, then $\overline{x}$ is an optimal solution to the minimization problem because of the complementarity theorem.
If the algorithm instead shows that the system has no solution, then we use the algorithm on the system $Ax \ge b$, $x \ge 0$ to determine whether the minimization problem has feasible points or not. If this latter system has feasible points, then it follows from our first investigation that the dual problem has no feasible points, and we conclude that the objective function is unbounded below, because of the duality theorem.

Download free eBooks at bookboon.com

Index

Download free eBooks at bookboon.com

Download free eBooks at bookboon.com
Click on the ad to read more

To see Part III, download:
Descent and Interior-point Methods:
Convexity and Optimization – Part III

Download free eBooks at bookboon.com