UNIVERSITY OF PASSAU

Faculty of Computer Science and Mathematics

Lecture

Foundations of Optimization

given in WS 2022/23, 2023/2024

Prof. Dr. Tobias Harks

Universität Passau
Institut für Mathematik
Dr.-Hans-Kapfinger-Straße 30
94032 Passau

Email: tobias.harks@uni-passau.de

January 23, 2024

Preface

This scriptum has been developed during the optimization lectures given at Augsburg University 2015-2022. The contents of Chapters 2,3,5-12 are based on hand-written notes of myself taken as a student at a lecture given by Helmut Maurer (University of Münster) in 2003. Chapter 4 is based on lecture notes of Stefan Ulbrich (TU Darmstadt, as of WS 2013), which in turn are based on lecture notes of Martin Grötschel (ZIB and TU Berlin). Chapter 13 is loosely based on lecture notes of Hans-Joachim Oberle (Hamburg University).

Special thanks go to Lukas Graf and Georg Kraus who suggested several corrections to the scriptum during the course held in SS 2020.

Passau, October 2023
Prof. Dr. Tobias Harks

Current Version: January 23, 2024.

Contents

Chapter 1

Introduction

1.1 Motivation and Terminology

We consider a normed vector space $(V, \|\cdot\|)$ for $K \subset V$ nonempty. We are given a function

$$f : K \to \mathbb{R}.$$

Our goal is to solve $\min\{f(x) \mid x \in K\}$. We sometimes write:

$$\begin{aligned} &\text{minimize } f(x) \\ &\quad \text{s.t.: } x \in K. \end{aligned} \tag{1.1}$$

An optimal solution of (1.1) is called global minimum.

Definition 1.1. $x^* \in K$ is a global minimum of f over K, if

$$f(x) \geq f(x^*) \text{ for all } x \in K.$$

If

$$f(x) > f(x^*) \text{ for all } x \in K, x \neq x^*,$$

we speak of a strict global minimum. Global maxima are defined analogously.

- infinite dimensional optimization (e.g., V is a function space, L_1 or L_p space.),
- finite dimensional optimization ($V = \mathbb{R}^n$),
- continuous optimization ($\text{int}(K) \neq \emptyset$),
- discrete optimization ($K \subseteq \mathbb{Z}^n$).

1.2 Examples and Applications

Example 1.2 (Optimal Supply). The goal is to buy an amount M of a certain commodity. We have the offers of n suppliers, where every supplier $i \in M$ has a maximum supply of

M_i units of the commodity. The prices of the i-th supplier are given by a function $f_i(x_i)$.

$$\min \sum_{i=1}^{n} f_i(x_i)$$
$$\text{s.t.}: \sum_{i=1}^{n} x_i = M$$
$$0 \leq x_i \leq M_i, i = 1, \dots, n.$$

Example 1.3 (Regression). An experiment shows the following data $(t_i, y_i), i = 1, \dots, m$. The hypothesis class under which this data is generated is given by a parameterized function $f(t, p)$. The goal is to choose parameters p in order to minimize the resulting error measured as:

$$\sum_{i=1}^{m} (y_i - f(t_i, p))^2$$

A more general least-squares problem is given as:

$$\min \sum_{i=1}^{q} \Phi_i(x)^2$$
$$\text{s.t.}: \; g_j(x) \leq 0, j = 1, \dots, r$$
$$h_j(x) = 0, j = 1, \dots, s.$$

Example 1.4 (Optimal control). We search for a control function that steers a car with minimal "effort" in a given time frame $[0, t_f]$ from A to B. We use Newton's laws describing where $s(t)$ denotes the location at time t, $v(t)$ the speed at time t and $a(t)$ denotes the acceleration:

$$\dot{s}(t) = v(t), \dot{v}(t) = a(t).$$

Suppose we are in dimension 1 and there is a straight line between A nach B of length d. The coordinate of A is normalized to $s = 0$ and $s = d$ for B. We need to satisfy $s(0) = 0, v(0) = 0, s(t_f) = d, v(t_f) = 0$.

The control function corresponds to $a(t)$, where a positive sign is acceleration and negative sign is slow down. The effort accumulates quadratically in a:

$$\int_0^{t_f} a(t)^2 dt$$

We obtain the following optimal control problem:

$$\min \int_0^{t_f} a(t)^2 dt$$
$$\dot{s}(t) = v(t)$$
$$\dot{v}(t) = a(t)$$
$$s(0) = 0, s(t_f) = d$$

$$v(0) = 0, v(t_f) = 0.$$

1.3 Finite Dimensional Optimization

In this lecture, we consider only finite dimensional problems, that is, $V := \mathbb{R}^n$.
The set of points in K that attain the minimum is denoted by $\arg\min(f, K)$. We get

$$\alpha = \min\{f(x) : x \in K\} \Leftrightarrow \arg\min(f, K) = \{x \in K \mid f(x) = \alpha\}.$$

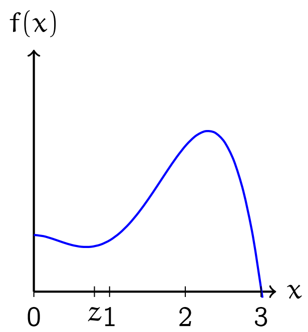


Figure 1.1: f attains on $K = [0, 3]$ its global minimum at $x^* = 3$. There is a further local minimum at z.

Definition 1.5. A point $x^* \in K$ is a local Minimum of f over K, if there is $\rho > 0$ such that

$$f(x) \geq f(x^*) \text{ for all } x \in K \cap B_\rho(x^*),$$

where

$$B_\rho(x) = \left\{y \in \mathbb{R}^n \;\middle|\; \|x - y\| < \rho\right\}$$

denotes the open ball around x with radius $\rho > 0$. If

$$f(x) > f(x^*) \text{ for all } x \in K \cap B_\rho(x^*), x \neq x^*,$$

we speak of a strict local minimum.

Usually K is represented via functional inequalities or equalities. In this case, we obtain:

$$K = \left\{x \in \mathbb{R}^n \;\middle|\; h_i(x) = 0,\;\; i \in I_1 = \{1, \ldots, m\},\; g_j(x) \leq 0,\;\; j \in I_2 = \{1, \ldots, p\}\right\}, \qquad (1.2)$$

where all functions satisfy $f, h_i, g_j \in C^2$ for all $i \in I_1, j \in I_2$.
We obtain the following classes of optimization problems:

- unrestricted optimization: $m = p = 0$
- restricted optimization: $m > 0$ oder $p > 0$
- linear optimization: f linear, g_j, h_i affin linear

- quadratic optimization: $f(x) = x^T A x + b^T x + c$, g_j, h_i affin linear
- convex optimization: f konvex, g_j konvex, h_i affin linear
- L_2-problems: the function f has the form

$$f(x) = \sum_{i=1}^{n} w_i \left(f_i(x)\right)^2,$$

  with smooth functions f_i and weights $w_i > 0$.
- minimax problems: f has the form

$$f(x) = \max\{f_i(x), i = 1, ..., m\}$$

  with smooth functions f_i.

The following questions are the key drivers for the content of this lecture:

- When do optimal solutions exist?
- Are they unique?
- Can we derive useful necessary and sufficient optimality conditions?
- What about algorithms for solving such problems?
- What is the dependence of optimal solutions on problem parameters?

We recap a fundamental result due to Weierstrass.

Theorem 1.6 (Weierstrass). Let $K \subset \mathbb{R}^n$ be nonempty and compact and $f : K \to \mathbb{R}$ continuous. Then, there is $x^* \in K$ with

$$f(x^*) \le f(x) \text{ for all } x \in K.$$

Proof. As f is continuous, the image $f(K)$ of the compactum K is bounded in $\mathbb{R}$ and the infimum

$$A := \inf\{f(x) | x \in K\} \in \mathbb{R}$$

exists. Hence, there is a sequence $x_n \in K, n \in \mathbb{N}$ with

$$\lim_{n\to\infty} f(x_n) = A.$$

As $x_n, n \in \mathbb{N}$ is bounded, we can use the theorem of Bolzano/Weierstrass giving a convergent subsequence $x_{n_k}, k \in \mathbb{N}$ with

$$\lim_{k\to\infty} x_{n_k} =: x^* \in K.$$

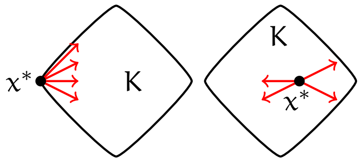


Figure 1.2: The red arcs represent feasible directions in $D_K(x^*)$.

With continuity of f we get

$$f(x^*) = \lim_{k\to\infty} f(x_{n_k}) = A.$$

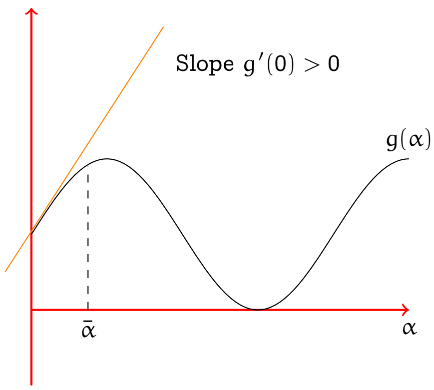
Thus, f attains at x^* its minimum over K . □

1.4 Differentiable Classic Optimization

1.4.1 Variational Inequalities

Definition 1.7 (Feasible Directions). Let $x \in K \subseteq \mathbb{R}^n$ with $K \neq \emptyset$. The vector $d \in \mathbb{R}^n$ is a feasible direction at x, if there is $\bar{\alpha} > 0$ such that $x + \alpha d \in K$ for all $0 \leq \alpha \leq \bar{\alpha}$.

We denote by $D_K(x)$ the set of feasible directions at x. It is easy to see that $D_K(x)$ is a pointed cone containing 0 (cf. 2.1), hence, $D_K(x)$ is known as the cone of feasible directions.



For continuous optimization problems (1.1) we obtain the following necessary optimality conditions.

Theorem 1.8 (Variational Inequality). Let $K \subseteq \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable. Let x^* be a local minimum of f over K and $d \in D_K(x^*)$. Then

$$\nabla f(x^*)^T d \geq 0.$$

Proof. Since $d \in D_K(x^*)$, there is $\bar{\alpha} > 0$ such that $x^*(\alpha) := x^* + \alpha d \in K$ for all $0 \leq \alpha \leq \bar{\alpha}$. We define a 1-dimensional function $g(\alpha) := f(x^*(\alpha))$. For a local minimum x^* (w.r.t. $B_\rho(x^*)$), we have

$$g(\alpha) \geq g(0) \text{ for all } \alpha \in [0, \min\{\bar{\rho}, \bar{\alpha}\}],$$

where

$$\bar{\rho} := \sup\{\alpha \geq 0 | x^* + \alpha d \in B_{\rho/2}(x^*)\}.$$

Thus,

$$\lim_{\alpha \to +0} \frac{g(\alpha) - g(0)}{\alpha} \geq 0.$$

With the differentiability of f we get

$$0 \leq \lim_{\alpha \to +0} \frac{g(\alpha) - g(0)}{\alpha} = g'(0) = \nabla f(x^*)^T d.$$

□

Theorem 1.9. Let $K \subseteq \mathbb{R}^n$ and $f : K \to \mathbb{R}$ be continuously differentiable. If $x^* \in \mathrm{int}(K)$ is a local minimum of f over K, then:

$$\nabla f(x^*) = 0. \tag{1.3}$$

In particular, we have (1.3) for every local minimum of an unconstrained optimization problem.

Proof. With Theorem 1.8 we get for every $d \in D_K(x^*)$: $\nabla f(x^*)^T d \geq 0$. With $x^* \in \mathrm{int}(K)$, we get $D_K(x^*) = \mathbb{R}^n$. □

Remark 1.10. Note that the concept of feasible directions of Definition 1.7 is not useful for sets given my algebraic manifolds. Here, we need curved directions leading to concepts of the tangent cone and linearized cone that we will see later.

1.4.2 Convex Optimization

We consider now a differentiable convex function f over a convex set $K \subset \mathbb{R}^n$.

Definition 1.11. A set $K \subset \mathbb{R}^n$ is convex, if for all $x, y \in K$ the segment between x and y lies in K, that is,

$$\lambda x + (1 - \lambda) y \in K \text{ for all } \lambda \in [0, 1].$$

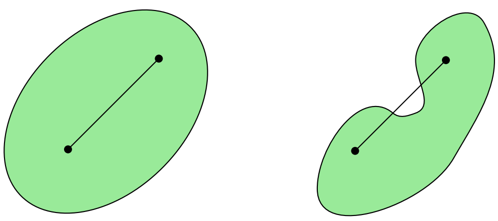


Figure 1.3: Left: convex set. Right: non-convex set.

Definition 1.12. Let $K \subset \mathbb{R}^n$ be convex. A function $f : K \to \mathbb{R}$ is convex, if for all $x, y \in K$ and $\lambda \in [0, 1]$ we have:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y). \tag{1.4}$$

f is <u>strict convex</u>, if for all $x \neq y$ and $\lambda \in (0, 1)$ the above inequality is strict. The function f is called (strictly) concave, if $-f$ is (strictly) convex.

Theorem 1.13. Let $K \subset \mathbb{R}^n$ be convex, and let $f_1, f_2 : K \to \mathbb{R}$ be convex functions and let $\alpha > 0$. Then, the functions $\alpha f_1, f_1 + f_2$ and $\max\{f_1, f_2\}$ are convex over K.

Proof. Exercise. □

Differences, products and minima of convex functions are not always convex!

Definition 1.14. Let $K \subset \mathbb{R}^n$ be convex, and $f : K \to \mathbb{R}$. The set

$$\mathrm{Epi}(f) = \{(x, \alpha) \in K \times \mathbb{R} : f(x) \leq \alpha\}$$

is the <u>epigraph</u> vof f. For $\beta \in \mathbb{R}$, we term the set

$$L(f, \beta) = \{x \in K : f(x) \leq \beta\}$$

<u>lower level set</u> of f with level β.

Theorem 1.15. Let $K \subseteq \mathbb{R}^n$ and $f : K \to \mathbb{R}$. Then:

1. f is convex $\Leftrightarrow$ $\mathrm{Epi}(f)$ is convex.
2. f is convex $\Rightarrow$ $L(f, \beta)$ is convex for all $\beta \in \mathbb{R}$. The reverse need not be true.

Proof. Exercise. □

For convex <u>differentiable functions</u> we obtain the following characterization:

Theorem 1.16. Let $f : \mathbb{R}^n \to \mathbb{R}$ and $f \in C^1$. Then:

1. f is konvex over the convex set $K \subseteq \mathbb{R}^n$ iff for all $x, y \in K$:

$$f(y) \geq f(x) + \nabla f(x)^T (y - x). \tag{1.5}$$

2. f is strict convex $\Rightarrow$ (1.5) is strict for all $x \neq y \in K$.

Proof. We first show $\Leftarrow$ for the first statement. Assume (1.5) holds for all $x, y \in K$. Choose arbitrary $x, y \in K$ and $\lambda \in (0, 1)$. With convexity of K we get

$$z = \lambda x + (1 - \lambda)y \in K. \tag{1.6}$$

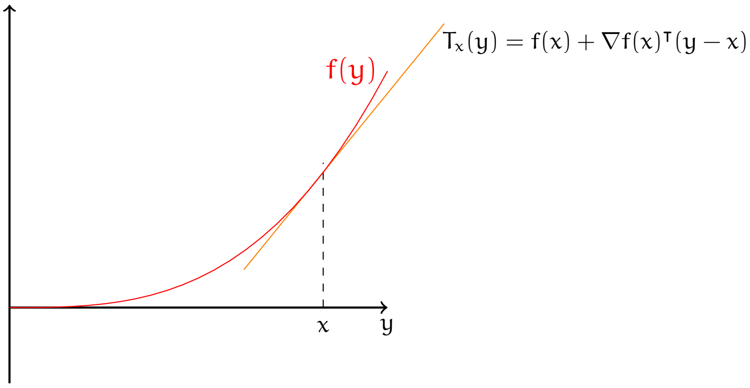


Figure 1.4: Illustration of inequality 1.5. $T_x(y)$ represents the tangent plane of f in x and we have $T_x(y) \leq f(y)$ for all $y \in K$.

With (1.5), we get for $x, y, z \in K$:

$$f(x) \geq f(z) + (x-z)^T \nabla f(z) \quad (1.7)$$
$$f(y) \geq f(z) + (y-z)^T \nabla f(z). \quad (1.8)$$

Multiply (1.7) with λ and (1.8) with $(1-\lambda)$, add both inequalities and obtain:

$$\begin{aligned} \lambda f(x) + (1-\lambda) f(y) &\geq f(z) + \Big((\lambda(x-z) + (1-\lambda)(y-z)\Big)^T \nabla f(z) \\ &= f(z) + \Big(\lambda x + (1-\lambda)y - z)\Big)^T \nabla f(z) \\ &= f(z). \end{aligned}$$

With (1.6), the second expression of the second equation is 0. Thus, f is convex.
$\Rightarrow$: Let f be convex. We choose $x, y \in K$ and define $\psi : \mathbb{R} \to \mathbb{R}$ as

$$\psi(\lambda) = (1-\lambda) f(x) + \lambda f(y) - f((1-\lambda)x + \lambda y).$$

With convexity of f we get for all $\lambda \in [0,1]$ that $\psi(\lambda) \geq 0$. Moreover $\psi(0) = 0$. We compute the derivative of ψ at 0 and get

$$0 \leq \lim_{t \to 0+} \frac{\psi(t) - \psi(0)}{t} = \dot{\Psi}(0) = -f(x) + f(y) - \nabla f(x)^T (y-x).$$

The second statement is easy.

□

We obtain a <u>sufficient</u> optimality criterion for convex optimization problems.

Theorem 1.17. Let $K \subset \mathbb{R}^n$ be convex and let $f : K \to \mathbb{R}$ be a differentiable konvex function. Then, every local minimum of f over K is also a <u>global</u> Minimum.

Proof. Let x^* be a local minimum. With Theorem 1.8, we get for every $d \in D_K(x^*)$ the condition $\nabla f(x^*)^T d \geq 0$. Because K is convex, for any $y \in K$, we get $x^* + \lambda(y - x^*) =$

$\lambda y + (1-\lambda)x^* \in K$ for all $\lambda \in [0,1]$. Hence, $y - x^* \in D_K(x^*)$. We get

$$f(y) \geq f(x^*) + \nabla f(x^*)^T (y - x^*) \geq f(x^*),$$

where the first inequality follows from Theorem 1.16 and the second one from the variational inequality. □

For unrestricted convex problems, we get the following implication.

Corollary 1.18. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable konvex function. Then, every $x^* \in \mathbb{R}^n$ with $\nabla f(x^*) = 0$ is a global minimum of the associated unrestricted optimization problem.

Chapter 2

Convexity and Separating Hyperplanes

2.1 Convex Sets and Cones

Definition 2.1.

1. For $M \subset \mathbb{R}^n$ we define
$$co(M) := \cap\{K \supset M | K \text{ convex}\}$$
as the convex hull of M. For $x^0, \dots, x^k \in \mathbb{R}^n$ we define
$$co(x^0, \dots, x^k) = co(\{x^0, \dots, x^k\}).$$
This set is known as the simplex spanned by the points $x^0, \dots, x^k$. If $x^1 - x^0, \dots, x^k - x^0$ are linearly independent, then the simplex is non-degenerate.

2. A subset $K \subset \mathbb{R}^n$ is a cone (pointed at 0), if for all $x \in K$ the half-ray through x lies in K, i.e.
$$\alpha x \in K \text{ for all } \alpha \geq 0.$$

3. Let $K \subset \mathbb{R}^n$ and $x \in K$. The cone
$$K(x) := \{\alpha(y - x) | y \in K, \alpha > 0\} = \bigcup_{\alpha > 0} \alpha(K - x)$$
is termed the conic hull of K w.r.t. x.

See Fig. 2.1 for an illustration.

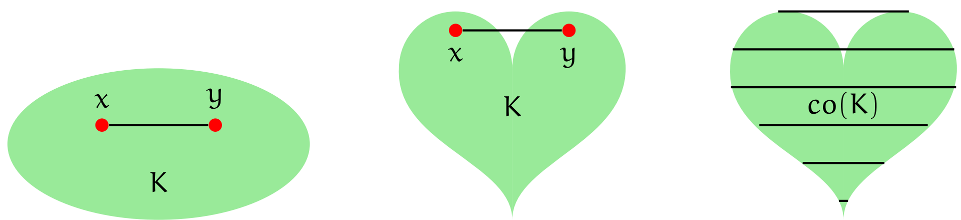


Figure 2.1: The first set is convex. The second heart is non-convex and the dashed set represents the convex hull.

Definition 2.2. Let $a \in \mathbb{R}^n \setminus \{0\}$ and $b \in \mathbb{R}$.

1. The set $H = \{x \in \mathbb{R}^n : a^\intercal x = b\}$ is called hyperplane.
2. The sets $H^- = \{x \in \mathbb{R}^n : a^\intercal x \leq b\}$ and $H^+ = \{x \in \mathbb{R}^n : a^\intercal x \geq b\}$ are Halfspaces.
3. Let A be a real-valued $m \times n$ matrix and $b \in \mathbb{R}^m$.
   $K = \{x \in \mathbb{R}^n | Ax \leq b\}$ is a polyhedron and $K = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$ is a polyhedron in standard form.

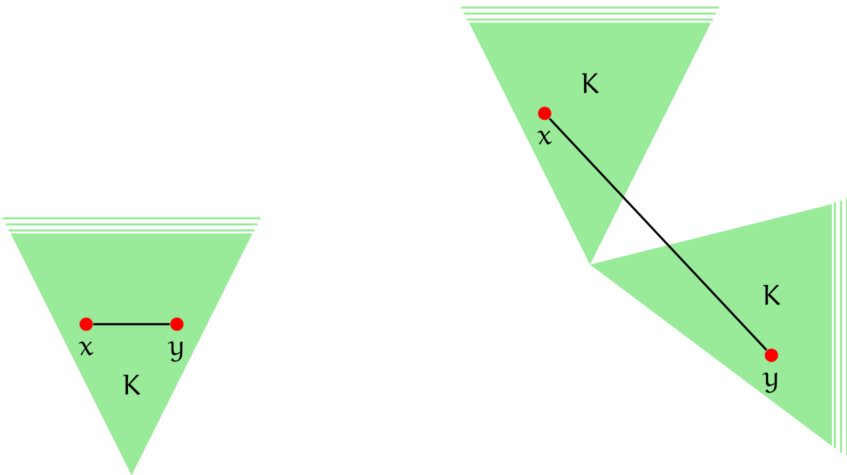


Figure 2.2: Left: convex cone. Right: non-convex cone.

2.2 Convex Combinations

Definition 2.3. Let $x^1, \dots, x^k \in \mathbb{R}^n$ und $\lambda_1, \dots, \lambda_k \in \mathbb{R}_{\geq 0}$ with $\lambda_1 + \cdots + \lambda_k = 1$.

- The vector $\sum_{i=1}^k \lambda_i x^i$ is called convex combination of $x^1, \dots, x^k$.

Theorem 2.4.
1. The intersection of convex sets is convex.
2. Every polyhedron is convex.
3. A convex combination of a finite points of a convex set lies in the respective set.
4. The convex hull of a set $K \subset \mathbb{R}^n$ is the set of all convex combinations of points in K. The set of convex combinations of a finite point set is convex.

Proof. (1): Let $X_i, i \in I$ be convex sets and define $X := \cap_{i \in I} X_i$. For $x, y \in X$ we have $x, y \in X_i$ for all $i \in I$, hence $\lambda x + (1-\lambda)y \in X_i$ for all $i \in I$ and, hence, $\lambda x + (1-\lambda)y \in X$.
(2): A polyhedron is the intersection of finitely many convex halfspaces.

(3): We prove via induction over k, that every convex combination of k points in X lies in X. For $k = 1$ the statement is trivial and for $k = 2$ the statement follows from the convexity of X. For the step $k-1 \to k$ consider a convex combination $\mu_1 x^1 + \cdots \mu_k x^k$. If $\mu_i = 0$ for some $i \in \{1, \dots, k\}$ we can use the induction hypothesis, hence, we can assume w.l.o.g. that $\mu_k \in (0,1)$. Define

$$\nu_1 = \frac{\mu_1}{1-\mu_k}, \dots, \nu_{k-1} = \frac{\mu_{k-1}}{1-\mu_k} \geq 0, \sum_{l=1}^{k-1} \nu_l = 1.$$

Set

$$y := \sum_{l=1}^{k-1} \nu_l x^l$$

and observe that $y \in X$ follows by the induction hypothesis. We get

$$x = (1-\mu_k)y + \mu_k x^k \in X,$$

because the case $k = 2$ was shown already.

(4): Let L be the set of convex combinations of points in K.

$L \subseteq \operatorname{co}(K)$: With (1) the set $\operatorname{co}(K)$ is convex, hence, all convex combinations of points in $\operatorname{co}(K)$ lie again in $\operatorname{co}(K)$. With the definition of the convex hull, we get $K \subseteq \operatorname{co}(K)$, thus, $L \subseteq \operatorname{co}(K)$.

$\operatorname{co}(K) \subseteq L$: Let $x, y \in L$ with

$$x = \sum_{i=1}^{k} \alpha_i x^i \text{ and } y = \sum_{j=1}^{l} \beta_j y^j, \text{ where } \alpha_i, \beta_j \geq 0, \sum_{i=1}^{k} \alpha_i = 1, \sum_{j=1}^{l} \beta_j = 1.$$

For $\lambda \in [0,1]$ we get

$$z = \lambda x + (1-\lambda)y = \sum_{i=1}^{k} \lambda \alpha_i x^i + \sum_{j=1}^{l} (1-\lambda)\beta_j y^j,$$

and thus $z \in L$ because

$$0 \leq \lambda\alpha_i \leq 1 \,\forall i, \; 0 \leq (1-\lambda)\beta_j \leq 1 \,\forall j \text{ and}$$

$$\sum_{i=1}^{k} \lambda\alpha_i + \sum_{j=1}^{l}(1-\lambda)\beta_j = \lambda \sum_{i=1}^{k} \alpha_i + (1-\lambda) \sum_{j=1}^{l} \beta_j = \lambda + 1 - \lambda = 1,$$

which shows that z is a convex combination of points in K. Thus L is convex. Obviously $K \subseteq L$, since every $x^p \in K$ can be written as

$$x^p = \sum_{j \in J} \lambda_j x^j \text{ with } \lambda_p = 1 \text{ and } \lambda_j = 0 \text{ for } i \neq p.$$

Per definition we get $\operatorname{co}(K) \subseteq L$, because L is a convex set containing K. □

Corollary 2.5. The set of convex combinations of $x^1, \ldots, x^k \in \mathbb{R}^n$ is the smallest (w.r.t. inclusion) convex subset of $\mathbb{R}^n$, which contains $x^1, \ldots, x^k$.

Proof. Let X be the set of convex combinations of $x^1, \ldots, x^k \in \mathbb{R}^n$. Define

$$Y := \operatorname{co}(x^1, \ldots, x^k) = \bigcap_{\substack{C \subseteq \mathbb{R}^n \\ C \text{ convex} \\ \{x^1,\ldots,x^k\} \subset C}} C. \tag{2.1}$$

Y ist well-defined because $\mathbb{R}^n$ is one candidate C. With Theorem 2.4(1.) Y is convex as intersection of convex sets. Y is also the smallest convex set containing $x^1, \ldots, x^k$. Since X is convex (see Theorem 2.4(4.) we get $Y \subseteq X$. Let $x \in X$. With the definition of X we get $x = \sum_{i=1}^k \lambda_i x^i$. As by assumption $x^1, \ldots, x^k \in Y$ we get with Theorem 2.4(3.), that $x \in Y$. □

Theorem 2.6 (Charathéodory)**.** For $K \subset \mathbb{R}^n$, $\operatorname{co}(K)$ is equal to the set of all convex combinations which require at most $(n+1)$ points of K.

Proof. Let $x \in \operatorname{co}(K)$. With Theorem 2.4 (4) there are $x^1, \ldots, x^k \in K$ with

$$x = \sum_{i=1}^k \lambda_i x^i \text{ mit } \lambda_i \geq 0 \text{ for all } i = 1, \ldots, k \text{ and } \sum_{i=1}^k \lambda_i = 1.$$

If $k \leq n+1$ we are done. If $k > n+1$, then we show that for the representation of x we can ignore one of the k points: Define the $(k-1)$ vectors $y^i = x^i - x^k, i = 1, ..., k-1$. For $k > n+1$, the points y^i are linearly dependent, i.e., there are $\alpha_1, \ldots, \alpha_{k-1}$ with $\alpha_j \neq 0$ for at least one $j \in \{1, \ldots, k-1\}$ with

$$\begin{aligned} &\sum_{i=1}^{k-1} \alpha_i y^i = 0 \\ \Leftrightarrow &\sum_{i=1}^{k-1} \alpha_i (x^i - x^k) = 0 \\ \Leftrightarrow &\sum_{i=1}^{k-1} \alpha_i x^i + (-\sum_{i=1}^{k-1} \alpha_i) x^k = 0. \end{aligned}$$

With $\alpha_k = -\sum_{i=1}^{k-1} \alpha_i$ we get

$$\sum_{i=1}^k \alpha_i x^i = 0 \text{ und } \sum_{i=1}^k \alpha_i = 0.$$

Because $\alpha_j \neq 0$ for at least one $j \in \{1, \ldots, k-1\}$, the following value is well-defined:

$$i_0 = \arg \min_{i \in \{1,\ldots,k\}} \{\frac{\lambda_i}{\alpha_i} | \alpha_i > 0\} = \frac{\lambda_{i_0}}{\alpha_{i_0}}.$$

We get

$$\lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}}\alpha_i \geq 0 \ \forall i \ \text{ and } \sum_{i=1}^{k} \lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}}\alpha_i = 1.$$

Moreover

$$x = \sum_{i=1}^{k} \lambda_i x^i = \sum_{i=1}^{k} \left(\lambda_i x^i - \frac{\lambda_{i_0}}{\alpha_{i_0}}\alpha_i x^i \right) = \sum_{i=1}^{k} (\lambda_i - \alpha_i \frac{\lambda_{i_0}}{\alpha_{i_0}}) x^i.$$

Here, we have $\lambda_{i_0} - \alpha_{i_0}\frac{\lambda_{i_0}}{\alpha_{i_0}} = 0$, and, hence, x can represented as a convex combination of at most $k-1$ points. □

2.3 Separating Hyperplanes

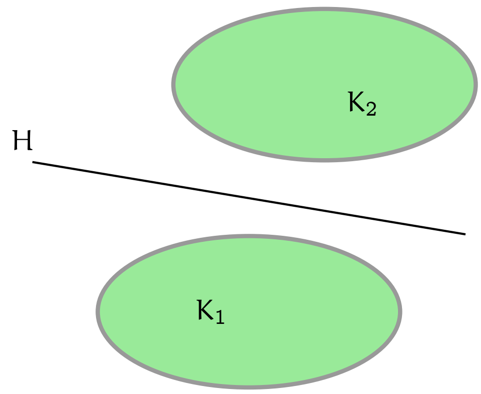


Figure 2.3: Illustration of the (strict) separation of two disjoint convex sets.

Definition 2.7. Two sets $K_1, K_2 \subset \mathbb{R}^n$ are called separable, if there are $c \in \mathbb{R}$ and row vector $\lambda \in \mathbb{R}^n, \lambda \neq 0$ with

$$\lambda x \leq c \leq \lambda y \text{ for all } x \in K_1, y \in K_2.$$

The hyperplane $H = \{x \in \mathbb{R}^n | \lambda x = c\}$ is called separating hyperplane (cf. Fig. 2.3); The sets K_1, K_2 are strictly separable via H, if $K_1 \cup K_2$ is not contained in H. The hyperplane H defines two halfspaces

$$H^+ = \{x \in \mathbb{R}^n | \lambda x \geq c\}, \ H^- = \{x \in \mathbb{R}^n | \lambda x \leq c\}.$$

The sets K_1, K_2 are separable, if either $K_1 \subseteq H^+, K_2 \subseteq H^-$ or $K_1 \subseteq H^-, K_2 \subseteq H^+$.

Theorem 2.8. Let $K \subset \mathbb{R}^n$ be a non-empty, convex set an let $y \notin \overset{o}{K} := \mathrm{int}(K)$. Then, $\{y\}$ and K are separable, i.e., there is a row vector $\lambda \in \mathbb{R}^n \setminus \{0\}$ with

$$\lambda y \leq \lambda x \text{ for all } x \in K.$$

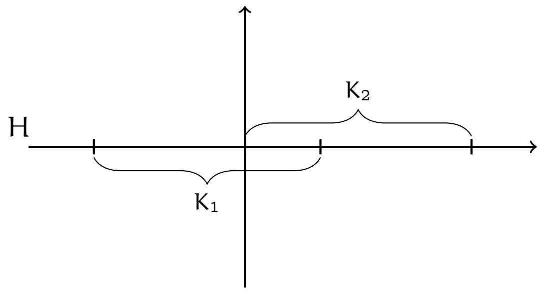


Figure 2.4: The two sets $K_1, K_2 \subset \mathbb{R}^2$ are separable but not strictly. The separating hyperplane is given as $H = \{x \in \mathbb{R}^2 | x_2 = 0\}$.

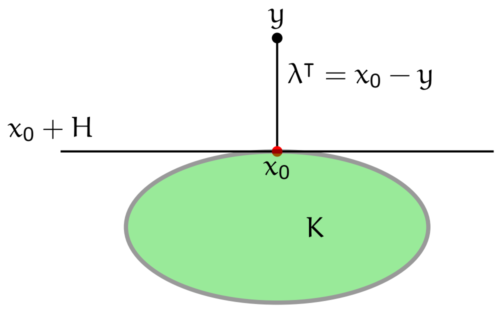


Figure 2.5: Illustration of the proof of Theorem 2.8. If $x_0 = y$ then $H + y$ is a separating supporting hyperplane.

> If $\overset{o}{K} \neq \emptyset$ then $\{y\}$ and $\overset{o}{K}$ are strictly separable and we get
>
> $$\lambda y < \lambda x \text{ for all } x \in \overset{o}{K}.$$

Proof. 1. Case: $y \notin \bar{K}$, where $\bar{K}$ denotes the topological closure of K. With $\|x\|$ we denote as usual the Euklidian norm. Set

$$d := \inf_{x \in \bar{K}} \|x - y\| > 0.$$

The function $f(x) := \|y - x\|$ is continuous and attains on $\bar{K} \cap \{x \in \mathbb{R}^n | \, \|x - y\| \leq 2d\}$ its minimum (Theorem of Weierstrass). As $\bar{K}$ is closed, there is $x_0 \in \bar{K}$ with $d = \|y - x_0\|$. With convexity of $\bar{K}$ one can further show that the point x_0 is unique (cf. Fig. 2.5).

We show $\lambda := (x_0 - y)^T \neq 0$ satisfies the conditions of the theorem. Let $x \in K$. With convexity of $\bar{K}$ we get

$$x_0 + \alpha(x - x_0) \in \bar{K} \text{ für } 0 \leq \alpha \leq 1.$$

Hence,

$$\|x_0 + \alpha(x - x_0) - y\|^2 \geq \|x_0 - y\|^2$$

and therefore

$$2\alpha(x_0 - y)^T(x - x_0) + \alpha^2 \|x - x_0\|^2 \geq 0.$$

Division by$\alpha > 0$ yields for $\alpha \downarrow 0$

$$(x_0 - y)^T(x - x_0) \geq 0$$

and therefore we get using $\lambda^T = (x_0 - y)$ and $d = \|\lambda\|$

$$\lambda x \geq \lambda x_0 = \lambda y + d^2 > \lambda y.$$

Thus, $H := \{x \in \mathbb{R}^n | \lambda x = \lambda x_0\}$ is a hyperplane separating $\{y\}$ and K.

2. Case: $y \in \partial K = \bar{K} - \overset{o}{K}$.

For $y \in \partial K$ there is a sequence $\{y_k\}, y_k \notin \bar{K}$, with $y = \lim_{k\to\infty} y_k$. For y_k we can choose according to Case 1. a row vector $\lambda_k \neq 0$ with

$$\lambda_k y_k \leq \lambda_k x \text{ for all } x \in K.$$

W.l.o.g. we can set $\|\lambda_k\| = 1$ and hence we can assume that the bounded sequence $\{\lambda_k\}$ converges with $\lambda = \lim_{k\to\infty} \lambda_k, \|\lambda\| = 1$. Taking the limit on both sides yields

$$\lambda y \leq \lambda x \text{ for all } x \in K.$$

The statement

$$\lambda y < \lambda x \text{ for all } x \in \overset{o}{K}$$

follows immediately. □

In case $y \in \partial K$ we call the hyperplane <u>supporting</u>. From the separating hyperplane theorem 2.8 we get:

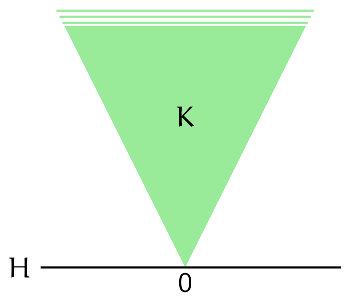


Figure 2.6: Separation of a convex cone via a hyperplane through 0.

Theorem 2.9. Let $K \subset \mathbb{R}^n$ be a non-empty convex and closed cone and suppose $y \notin K$. Then, there is $\lambda \in \mathbb{R}^n \setminus \{0\}$ with

$$\lambda y < 0 \leq \lambda x \text{ for all } x \in K.$$

Proof. With $\bar{K} = K$ there is – using the first statement of Theorem 2.8 – a row vector $\lambda \in \mathbb{R}^n \setminus \{0\}$ such that for all $x \in K$ we have $\lambda y < \lambda x$. From $0 \in K$ we get $\lambda y < \lambda 0 = 0$. Suppose there is $x \in K$ with $\lambda x < 0$. Then,

$$\lambda(\alpha x) = \alpha(\lambda x) \to_{\alpha \to \infty} -\infty,$$

contradicting boundedness of λK via λy from below. Thus, we get

$$\forall x \in K : \lambda y < 0 \leq \lambda x.$$

□

We get as an implication the following theorem of the alternatives:

Theorem 2.10 (Lemma of Farkas). Let B a $k \times n$ matrix and $d \in \mathbb{R}^k$. Then, exactly one of the following statements is true

1. $Bx = d, x \geq 0$ admits a solution $x \in \mathbb{R}^n$.
2. $\lambda B \geq 0, \lambda d < 0$ admits a solution $\lambda \in \mathbb{R}^k$.

Proof. The cone

$$K := \{Bx | x \geq 0\} \subset \mathbb{R}^k$$

non-empty convex and closed. Exactly one of the statements is true

(a) $d \in K$.

(b) $d \notin K$.

Statement (a) is statement (1) of the theorem. In case (b) we get with Theorem 2.8 the existence of some $\lambda \in \mathbb{R}^k$ with

$$\lambda d < 0 \leq \lambda z \text{ for all } z \in K.$$

Also $\lambda Bx \geq 0$ for all $x \geq 0$, i.e., $\lambda B \geq 0$. This is Statement (2) of the theorem. Note that (a) and (1) (and (b) and (2)) are equivalent, and, thus, (1) and (2) cannot be true simultaneously. □

Chapter 3

Introduction to Linear Optimization

3.1 Examples

3.1.1 Production Modells

A company produces n products $P_1, \dots, P_n$, and for the production process, m activities $A_1, \dots, A_m$ (workers, materials, etc.) are needed. Product P_j requires a_{ij} shares of the activity A_i yields a net-gain of c_j Euro. For activity A_i there is an upper bound of b_i. The production amount x_j of product P_j should be determined in order maximize net gain:

$$z(x) = \sum_{j=1}^{n} c_j x_j$$

subject to:

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i, \text{for all } i = 1, \dots, m$$

$$x_j \geq 0, \text{for all } j = 1, \dots, n.$$

Example 3.1. A shoe fabric produces two types of products. There are 40 employees and 10 machines. The working time budget and material budget is depicted in Table 3.1: With the decision variables

Table 3.1: Parameters.

	Type 1	Type 2	available
Production time [h]	20	10	8000
Machine hours [h]	4	5	2000
Material supply [dm^2]	6	15	4500
Net gain [EUR]	16	32	—

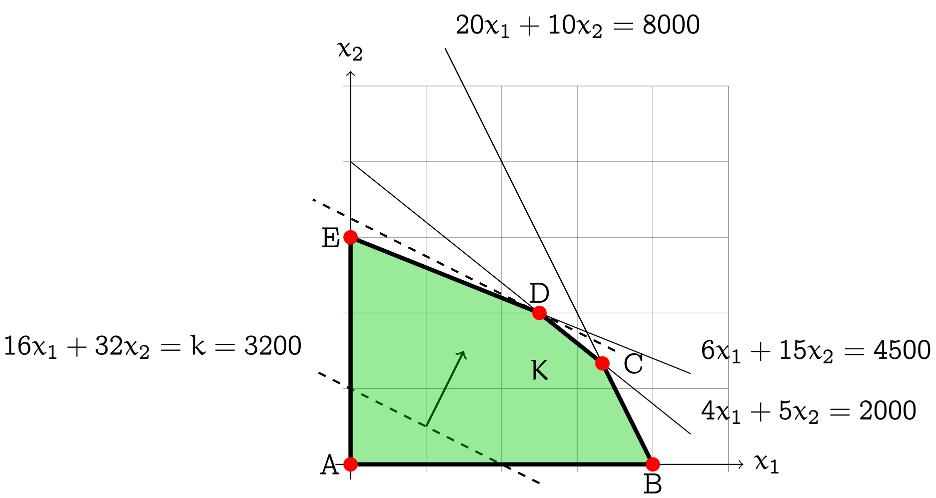
x_1 :amount type 1

$$x_2 : \text{amount type } 2$$

we get:

$$\text{maximize } z(x_1, x_2) = 16x_1 + 32x_2$$
$$\text{s.t.:}$$
$$20x_1 + 10x_2 \leq 8000$$
$$4x_1 + 5x_2 \leq 2000$$
$$6x_1 + 15x_2 \leq 4500$$
$$x_1 \geq 0$$
$$x_2 \geq 0$$

The level plane with level k is given by $z(x_1, x_2) = 16x_1 + 32x_2 = k$. One way of solving the problem is to determine the maximal k such that at least one point on $z(x_1, x_2) = k$ is feasible.



The function z attains its maximal value at <u>vertex</u> D of the feasible set K. The vertices of K are A, B, C, D, E:

Table 3.2: Vertices of the feasible region of the production problem

Vertex	z	x_1	x_2	y_1	y_2	y_3
B	6400	400	0	0	400	2100
C	9600	1000/3	400/3	0	0	500
D	10400	250	200	100	0	0

In every vertex, exactly two variables are 0 (basic solution). If we walk along vertices $B \to C \to D$ the value of z increases. The optimal vertex is D with objective value:

$$x_1 = 250, x_2 = 200, z(x_1, x_2) = 10400.$$

3.2 Mathematical Formulation of Linear Optimization Problems

The linear optimization (LP) in standard form is given by

$$\begin{aligned} \text{minimize } z(x_1, \ldots, x_n) = \sum_{j=1}^{n} c_j x_j \\ \text{s.t.:} \\ \sum_{j=1}^{n} a_{ij} x_j = b_i, 1 \leq i \leq m \\ x_j \geq 0, j = 1, \ldots, n \end{aligned}$$

In vector notation

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, c = (c_1, \ldots, c_n) \in \mathbb{R}^n, b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m, A = (a_{ik}) \; m \times n \text{ matrix}$$

we get

$$\begin{aligned} \text{minimize } z(x) = cx \\ \text{unter d.N.} \\ Ax = b \\ x \geq 0. \end{aligned} \tag{3.1}$$

3.3 Reduction of other LP's to Standardform

Suppose we have an LP that contains inequalities or free variables.

3.3.1 Inequalities

Let A be an $m \times n$ matrix (not necessarily $m < n$).

$$\begin{aligned} \text{minimize } z(x) &= cx \\ \text{s.t.:} & \\ Ax &\leq b \\ x &\geq 0. \end{aligned} \tag{3.2}$$

Define slack-variables

$$y := b - Ax \in \mathbb{R}^m.$$

Then, $Ax \leq b$ is equivalent to

$$Ax + y = b, \; y \geq 0.$$

With $\tilde{c} := (c, 0) \in \mathbb{R}^{n+m}$, $\tilde{x} := \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n+m}$, I being the $m \times m$ identity matrix and the $m \times (n+m)$-matrix $\tilde{A} := (A|I)$ we get that (3.2) is equivalent to the standard form formulation

$$\begin{aligned} \text{minimize } \tilde{z}(x) &= \tilde{c}\tilde{x} \\ \text{s.t.:} & \\ \tilde{A}\tilde{x} &= \tilde{b} \\ \tilde{x} &\geq 0. \end{aligned} \tag{3.3}$$

3.3.2 Free Variable

In standard form (3.1) let some x_i without sign constraint, e.g. x_1.

$$\begin{aligned} \text{minimize } z(x) &= cx \\ \text{s.t.:} & \\ Ax &= b \\ x_2 \geq 0, \ldots, x_n &\geq 0. \end{aligned} \tag{3.4}$$

1. Method: Elimination of x_1

Choose index i with $a_{i1} \neq 0$ and eliminate from

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$$

the variable x_1 as linear combination of $x_2, \ldots, x_n$. This yields a reduced linear equation system of

$$\tilde{A} \begin{pmatrix} x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \tilde{b} \in \mathbb{R}^{m-1}.$$

2. Method:

Set $x_1 = u_1 - v_1, u_1 := \max\{0, x_1\} \geq 0, v_1 := \max\{0, -x_1\} \geq 0$. This yields an LP with $n+1$ variables $u_1, v_1, x_2, \ldots, x_n$.

Chapter 4

Theory of Polyhedra

We will study fundamental elements of the theory of polyhedra including vertices, faces and valid inequalities.
The set

$$K := \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$$

is a polyhedron in standard form. The description of K is given via matrices A, b, hence, we write $P^=(A, b) := K$. Similarly, for

$$K := \{x \in \mathbb{R}^n | Ax \leq b\}$$

we write $P(A, b) := K$.

Definition 4.1. A bounded polyhedron is a polytope.

4.1 Faces of Polyhedra

We consider in the following polyedra of the form $P(A, b)$, where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.

Definition 4.2. Let $K \subseteq \mathbb{R}^n, a \in \mathbb{R}^n, \alpha \in \mathbb{R}$.

1. The inequality $a^T x \leq \alpha$ is called valid w.r.t. K, if

$$K \subseteq \{x \in \mathbb{R}^n | a^T x \leq \alpha\}.$$

2. The hyperplane $H = \{x \in \mathbb{R}^n | a^T x = \alpha\}, a \neq 0$, is called supporting hyperplane of K, if $a^T x \leq \alpha$ is valid w.r.t. K and $K \cap H \neq \emptyset$.

Note that $a = 0$ in the first statement of Definition 4.2 is feasible.

Definition 4.3. Let $K \subset \mathbb{R}^n$. A set $F \subseteq K$ is called a face of K, if there is a valid inequality w.r.t. K of the form $d^T x \leq \delta$ such that

$$F = K \cap \{x \in \mathbb{R}^n | d^T x = \delta\}.$$

A face is proper, if $F \neq K$. F is called non-trivial, is $\emptyset \neq F \neq K$.

Is $d^T x \leq \delta$ valid w.r.t. K, then $K \cap \{x \in \mathbb{R}^n | d^T x = \delta\}$ is called induced face by $d^T x \leq \delta$.

Note again that $d = 0$ in Definition 4.3 is allowed.

Example 4.4. We give an example.

$$\begin{aligned} 2x_1 + x_2 &\leq 8 \\ 4x_1 + 5x_2 &\leq 20 \\ 2x_1 + 5x_2 &\leq 15 \\ x_1, x_2 &\geq 0 \end{aligned}$$

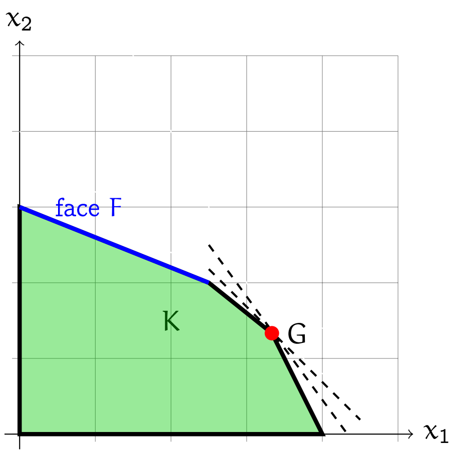


Figure 4.1: Example of a face.

The face F is induced by

$$2x_1 + 5x_2 \leq 15,$$

because

$$F = K \cap \{x \in \mathbb{R}^2_+ | 2x_1 + 5x_2 = 15\}.$$

Moreover, $G = (\frac{10}{3}, \frac{4}{3})$ is a face indued by the valid inequalities (in Proposition 4.8 it becomes clear that a face can be induced by several valid inequalities).

$$\begin{aligned} 2x_1 + x_2 &\leq 8 \\ 4x_1 + 5x_2 &\leq 20. \end{aligned}$$

Note that $G = (\frac{10}{3}, \frac{4}{3})$ as a face of K can also be induced via (see Fig. 4.1)

$$\begin{aligned} 4x_1 + 3x_2 &\leq 52/3 \\ x_1 + x_2 &\leq 14/3 \end{aligned}$$

Proposition 4.5. Let $K = P(A, b) \subset \mathbb{R}^n$be a polyhedron. Then, the following statements hold:

1. K is a face of itself.
2. $\emptyset$ is a face of K.
3. If $F = \{x | d^T x = \delta\} \cap K$ is a non-trivial face of K, then $d \neq 0$.

Proof. (1): $K = K \cap \{x \in \mathbb{R}^n | 0^T x = 0\}$.
(2): Let $\delta > 0$ be arbitrary. We have $0^T x \leq \delta$ is a valid inequality for K and we get $\emptyset = K \cap \{x \in \mathbb{R}^n | 0^T x = \delta\}$.
(3): For $d = 0$ the case $\delta \geq 0$ yields a valid inequality of the form $d^T x \leq \delta$ and, hence, one of the first two cases applies. □

Theorem 4.6. Let $K = P(A, b)$ be a non-empty polyhedron and $c^T \in \mathbb{R}^n$. Consider the LP

$$\min\{cx | x \in K\}.$$

Let F^* be the solution set and in case $F^* \neq \emptyset$ let $z^* = \min\{cx | x \in K\}$. Then:

1. If $F^* \neq \emptyset$, then $F^* = \{x \in K | cx = z^*\}$ is a non-empty face of K and if $c \neq 0$, the set $\{x \in \mathbb{R}^n | cx = z^*\}$ is a supporting hyperplane of K.
2. The set of optimal solutions of $\min\{cx | x \in K\}$ is a face of K.

Proof. For (1):
Let $F^* \neq \emptyset$. With the definition of z^* we get $cx \geq z^*$ for all $x \in K$, thus the inequality $-cx \leq -z^*$ is valid for K. Moreover, $F^* = \{x \in \mathbb{R}^n | cx = z^*\} \cap K$, implying that F^* is a non-empty face of K. We get immediately that $\{x \in \mathbb{R}^n | cx = z^*\}$ is a supporting hyperplane in case $c \neq 0$.
For (2): For $F^* = \emptyset$ the statement is clear, otherwise we get that (2) is just a reformulation of (1). □

An illustration of the above theorem appears in Fig. 4.2.
For $A \in \mathbb{R}^{m \times n}$ with row index set $M := \{1, \ldots, m\}$ and column index set $N = \{1, \ldots, n\}$ we denote by $a_i, i \in M$ the rows and by $a^j, j \in N$ the columns of A. For $I \subseteq M$ we denote by A_I the submatrix consisting of the rows $a_i, i \in I$.

Definition 4.7. Let $K = P(A, b) \subset \mathbb{R}^n$ and M the row index set of A. For $F \subseteq K$ let

$$eq(F) := \{i \in M | a_i x = b_i \forall x \in F\},$$

i.e. $eq(F)$ is the set of all active inequalities at $x \in F$. For $I \subseteq M$ denote by

$$fa(I) := \{x \in K | A_I x = b_I\}$$

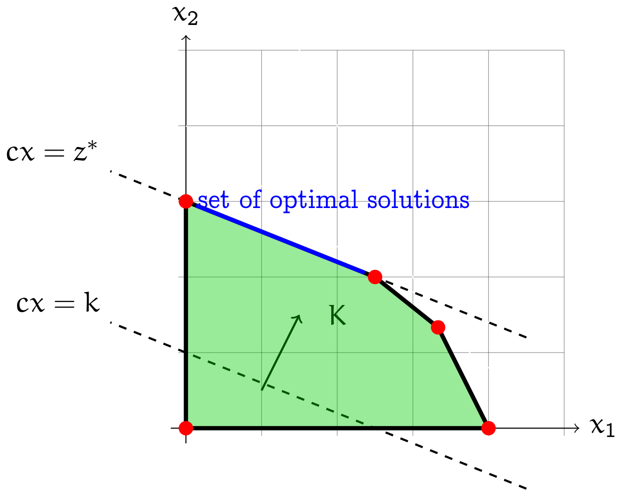


Figure 4.2: Graphical illustration of Theorem (4.6).

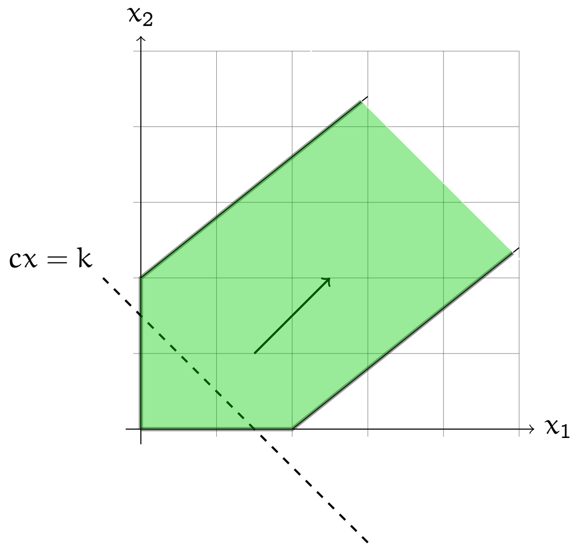


Figure 4.3: In this example, the set of optimal solutions is empty.

Figure 4.4: The black points of the polytopes are vertices. For a circle, every point on the boundary is a vertex.

the induced face by valid inequalities corresponding to index set I.

We verify that indeed $fa(I)$ is a face of K.

Proposition 4.8. The set $F := fa(I)$ defined in Definition 4.7 is a face of K.

Proof. If $I = \emptyset$, then $F = K$ is a trivial face of K. Let $|I| \geq 1$. Define

$$a^T := \sum_{i \in I} a_i \text{ and } \gamma := \sum_{i \in I} b_i.$$

We have that $a^T x \leq \gamma$ is a valid inequality and for $x \in K \setminus F$, at least ine inequality is strict, hence,

$$a^T x \begin{cases} = \gamma, & \text{for } x \in F, \\ < \gamma, & \text{else.} \end{cases}$$

We get $F = \{x \in K | a^T x = \gamma\} = K \cap \{x \in \mathbb{R}^n | a^T x = \gamma\}$. □

We consider again the Example 4.4. Here, we have $M = \{1, 2, 3, 4, 5\}$, $fa(\{1, 2\}) = G$ and $eq(G) = \{1, 2\}$.

4.2 Vertices and Extreme Points

Definition 4.9. Let $K \subset \mathbb{R}^n$.

1. $x \in K$ is called extreme point of K, if there are no two distinct points $y, z \in K$ with

$$x = \alpha y + (1 - \alpha) z \text{ for some } \alpha \in (0, 1).$$

2. $x \in K$ is called vertex of K, if $\{x\}$ is a 0-dimensional face of K.

Definition 4.9 works for any set $K \subset \mathbb{R}^n$.

Theorem 4.10. Let $K = P(A, b) \subset \mathbb{R}^n$ be a polyhedron and $x \in K$. Then, the following is equivalent:

1. x is a vertex of K.

2. $\{x\}$ is a 0-dimensional face of K.
3. x is an extreme point of K.
4. $\mathrm{rank}(A_{eq(\{x\})}) = n$.
5. There is $c^T \in \mathbb{R}^n \setminus \{0\}$, such that x is the unique solution to the LP $\min\{cy : y \in K\}$.

Proof. The statements (1) and (2) just correspond to the Definition 4.9 of a vertex. The proof works as follows: (2)⇒(5), (5)⇒(3), (3)⇒(4) and (4)⇒(2).
(2)⇒(5):
Per definition, x is a face, hence there is a valid inequality w.r.t. K $d^T x \leq \gamma$, such that $\{y \in K | d^T x = \gamma\} = \{x\}$.
Thus, x is the unique optimal solution to $\min\{cy : y \in K\}$ for $c := -d^T$.
If $K \neq \{x\}$, then $c \neq 0$ because of Proposition 4.5(3), otherwise we can choose $c \neq 0$ arbitrarily.
(5)⇒(3):
Let x be the unique optimal solution to $\min\{cy | y \in K\}$ with value γ. If $x = \lambda w + (1-\lambda)z$ for $w, z \in K, w \neq z, 0 < \lambda < 1$, then

$$\begin{aligned}\gamma = cx &= c(\lambda w + (1-\lambda)z)\\ &= \lambda cw + (1-\lambda)cz\\ &> \lambda\gamma + (1-\lambda)\gamma = \gamma,\end{aligned}$$

contradiction.
(3)⇒(4):
Suppose (4) does not hold. Then, there is $d \neq 0$ with $A_{eq(\{x\})} d = 0$. For small $\epsilon > 0$ we get

$$A(x \pm \epsilon d) \leq b.$$

With $y = x - \epsilon d$, and $z = x + \epsilon d$ we get $y, z \in K$ and $x = 1/2y + 1/2z$. Hence, we also get that (3) is not valid.
(4)⇒(2):
We know that $fa(eq(\{x\}))$ is a face and with $\mathrm{rank}(A_{eq(\{x\})}) = n$ we get

$$fa(eq(\{x\})) = \{y \in K | A_{eq(\{x\})} y = b_{eq(\{x\})}\} = \{x\}.$$

Thus, $\{x\}$ is a face and its dimension is 0. □

We get a similar result for polyeder in standard form.

Theorem 4.11. Let $K = P^=(A, b)$, i.e.

$$K := \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}.$$

$x \in K$ is an extreme point or equivalently a vertex, if and only if the column vectors a^k of A that correspond to indices k with $x_k > 0$ are linearly independent.

Proof. $\Rightarrow$: Let $x \in K$ be a vertex. W.l.o.g.

$$x = (x_1, \dots, x_r, 0, \dots, 0)^T, x_i > 0, i = 1, \dots, r.$$

If $r = 0$, the column set is empty, and the set is linearly independent. For $r > 0$ we have

$$\sum_{i=1}^{r} a^i x_i = b.$$

Contra-position: suppose $a^1, \dots, a^r$ are linearly dependent. Then, there are scalars $d_1, \dots, d_r, d_i \neq 0$ for at least one i, with

$$\sum_{i=1}^{r} a^i d_i = 0.$$

If $x_i > 0$, then for $\epsilon > 0$ small enough, we get

$$x_i \pm \epsilon d_i > 0, \text{ for } i = 1, \dots, r.$$

We set

$$d := (d_1, \dots, d_r, 0, \dots, 0)^T, y := x + \epsilon d, z := x - \epsilon d.$$

Then, $y, z \geq 0$ and with

$$\sum_{i=1}^{r} a^i (x_i \pm \epsilon d_i) = \sum_{i=1}^{r} a^i x_i \pm \epsilon \underbrace{\sum_{i=1}^{r} a^i d_i}_{=0} = b,$$

we have $y, z \in K$. With $x \neq y, z$, and $x = \frac{y+z}{2}$ we get a contradiction that x is an extreme point of K. Hence, $a^1, \dots, a^r$ are linearly independent.

$\Leftarrow$: W.l.o.g. assume that the first r components of x are positive, and assume that $a^1, \dots, a^r$ are linearly independent.

1. Case: $r = 0 :\Rightarrow x = 0$. If $x = 0$ is no extreme point, there are $y, z \in K, y \neq z$ and $0 < \alpha < 1$ with

$$0 = x = \alpha y + (1 - \alpha) z.$$

With $y, z \geq 0$ and $\alpha \neq 0$, we get $y = 0, z = 0$, contradiction.

2. Case: $r > 0$: Per definition we have $\sum_{i=1}^{r} a^i x_i = b$.

Contra-position: x is no extreme point of K. Then, there are $y, z \in K$, $y \neq z$, and $0 < \alpha < 1$ with

$$x = \alpha y + (1 - \alpha) z.$$

As in Case 1., we get

$$y_{r+1} = \dots = y_n = 0, z_{r+1} = \dots = z_n = 0.$$

Moreover,

$$Ay = Az = b \text{ hence } A(y-z) = 0 \Rightarrow \sum_{i=1}^{r} a^i(y_i - z_i) = 0$$

As $a^1, \dots, a^r$ are linearly independent, we get

$$y_i = z_i \text{ für } i = 1, \dots, r \Rightarrow y = z \Rightarrow x = y = z, \text{ contradiction.}$$ □

Definition 4.12. A polyhedron is called pointed, if it contains a vertex.

We define terms like edge and line of a polyhedron.

Definition 4.13. A polyhedron $K \subset \mathbb{R}^n$ contains a line, if $x \in K$ and there is $d \in \mathbb{R}^n$, such that

$$x + \lambda d \in K \text{ for all } \lambda \in \mathbb{R}.$$

An edge of K is a face of dimension 1 connecting two vertices of K.

Theorem 4.14. Let $K = P(A, b) \subset \mathbb{R}^n$ be non-empty. The following statements are equivalent:

1. K is pointed.
2. $\operatorname{rank}(A) = n$.
3. Every non-empty face of K is pointed.

Proof. (1)⇒(2): Is x a vertex of K, then with Theorem 4.10 we get

$$n = \operatorname{rank}(A_{eq(\{x\})}) \leq \operatorname{rang}(A) \leq n.$$

Hence, $\operatorname{rank}(A) = n$.
$(2) \Rightarrow (1)$: We choose $x \in K$ such that the set $I = eq(\{x\})$ is inclusion maximal. Let

$$F = \{x \in K | A_I x = b_I\}.$$

If $\operatorname{rank}(A_I) = n$, then we get with Thm. 4.10 that x is a vertex, hence, assume $\operatorname{rank}(A_I) < n$. Then, the kernel of A_I contains $d \neq 0$ with

$$x \pm \epsilon d \in K, \text{ for } \epsilon \text{ small enough.}$$

The line $\{x + \lambda d | \lambda \in \mathbb{R}\}$ hits at least one of the hyperplanes $H_j = \{x \in \mathbb{R}^n | a_j x = b_j\}$ for some $j \notin I$. (Suppose not, then all hyperplanes lie completely in K. Then,

$$a_i(x + \lambda d) \leq b_i, \text{ for all row indices } i, \text{ and all } \lambda \in \mathbb{R}.$$

This implies $Ad = 0$ and with $\mathrm{rank}(A) = n$ we get $d = 0$, contradiction.) Hence, there is $\delta \in \mathbb{R}$ such that $x + \delta d \in K$ and $eq(\{x + \delta d\}) \supset I$, in contradiction to the maximality of I.

(3)⇒(1): Is trivial as K is a face of itself.

(2)⇒(3): For every non-empty F of K we have

$$F = \{x \in \mathbb{R}^n | Ax \leq b, A_{eq(F)} x \leq b_{eq(F)}, -A_{eq(F)} x \leq -b_{eq(F)}\}.$$

From (2) and the equivalence of (2) and (1), we get that F must be pointed. □

We get an important corollary for polyhedra in standard form.

Corollary 4.15. Let $K = P^{=}(A, b)$, then

$$K \neq \emptyset \Leftrightarrow K \text{ is pointed.}$$

Proof. We obtain a representation of K via

$$K = P^{=}(A, b) = P(D, d) \text{ with } D = \begin{pmatrix} A \\ -A \\ -I \end{pmatrix}, d = \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}.$$

D has rank n. From Thm. 4.14 the statement follows. □

For polytopes we get a similar result.

Corollary 4.16. Let $K = P(A, b)$ be a polytope. Then,

$$K \neq \emptyset \Leftrightarrow K \text{ is pointed.}$$

Proof. As K is bounded, there is u with $K \subseteq \{x | x \leq u\}$. Thus we get a representation of K via

$$K = P(A, b) = P(D, d) \text{ with } D = \begin{pmatrix} A \\ I \end{pmatrix}, d = \begin{pmatrix} b \\ u \end{pmatrix}.$$

D has rank n. From Thm. 4.14 the statement follows. □

Corollary 4.17. Let $K = P(A, b)$ be a pointed polyhedron and suppose the LP

$$\min cx \text{ s.t. } x \in K$$

has a finite optimal solution. Then, the LP has an optimal solution that is a vertex.

Proof. $F = \{x \in K | cx = \min\{cy | y \in K\}\}$ is a non-empty face of K and contains using Thm. 4.14 a vertex. □

Corollary 4.18. If K is a non-empty polytope, then, every LP of the form

$$\min cx \text{ s.t. } x \in K$$

has an optimal vertex solution.

We collect these results in the main theorem of linear programming.

Theorem 4.19. The LP

$$\min cx \text{ s.t. } Ax = b, x \geq 0$$

admits a finite optimal solution if and only if admits an optimal vertex solution.

Proof. With Cor. 4.15 we have that K is pointed (if K non-empty) and with Cor 4.17 we get that LP has an optimal vertex solution (if there is a finite optimal solution). □

Thm. 4.19 can be used to solve an LP in standard form by trying out all vertices.

Example 4.20.

$$\begin{aligned} \min x_1 + 2x_2 + 3x_3 \\ \text{s.t.:} \\ 2x_1 + x_2 + 5x_3 = 5 \\ x_1 + 2x_2 + x_3 = 4 \\ x_i \geq 0, i = 1, 2, 3 \end{aligned}$$

Solution: For a vertex, one component is 0.
1. Possibility: $x_1 = 0$. Solve

$$\begin{bmatrix} 1 & 5 \\ 2 & 1 \end{bmatrix} \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \Rightarrow x_2 = 5/3, \ x_3 = 2/3$$

Vertex: $(0, 5/3, 2/3)^T$, objective $16/3$.

2. Possibility: $x_2 = 0$. Solve

$$\begin{bmatrix} 2 & 5 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \Rightarrow x_1 = 5,\ x_3 = -1$$

Infeasible.

3. Possibility: $x_3 = 0$. Solve

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \Rightarrow x_1 = 2,\ x_2 = 1$$

Vertex: $(2, 1, 0)^\top$, objective 4.

The set K is bounded, hence the objective function attains at $(2, 1, 0)^\top$ its minimum.

In contrast to exhaustive search (cf. Chapter 5), the simplex algorithm is a far more efficient algorithm for finding an optimal vertex.

4.3 Basic Solutions

We consider an LP in standard form:

$$\begin{aligned} \min\ z(x) &= cx \\ \text{s.t.:} & \\ Ax &= b \\ x &\geq 0 \end{aligned}$$

Notation: $K := P^{=}(A, b) = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$.

We assume $\mathrm{rank}(A) = m < n$. (Later we will see that this holds w.l.o.g.). We consider the linear equation system $Ax = b$, whose solution space is an $(n - m)$-dimensional affine subspace. The columns of A are denoted by $a^j, j = 1, \dots, n$.

Definition 4.21. An index vector $B = (i_1, \dots, i_m)$ with m distinct indices $i_j \in \{1, \dots, n\}$ is called basis, if the corresponding column vectors are linearly independent. The complement vector to B is denoted by $N = (j_1, \dots, j_{n-m})$, $j_k \in \{1, \dots, n\}$ and is called non-basis. We have $B \oplus N = \{1, \dots, n\}$. With A_B and A_N we denote the submatrices, defined via column vectors corresponding to B and N:

$$\begin{aligned} A_B &: m \times m \text{ matrix with column vectors } a^i, i \in B \\ A_N &: m \times (n - m) \text{ matrix with column vectors } a^j, j \in N \end{aligned}$$

For such a subdivision of the set $\{1, \dots, n\}$ we write the set K as

$$(A_B, A_N) \begin{pmatrix} x_B \\ x_N \end{pmatrix} = b, x_B, x_N \geq 0.$$

For B, we denote A_B as basis matrix and A_N as non-basis matrix. The variables $x_i, i \in B$ are called basic variables and the variables $x_j, j \in N$ are called non-basic variables. $x = (x_B, x_N)$ with

$$x_B := A_B^{-1} b \text{ und } x_N := 0$$

is termed basic solution w.r.t. basis B. A basic solution is feasible, if $x_B \geq 0$. A feasible basic solution is non-degenerate, if $x_B > 0$; if more than $n - m$ components of x are equal 0, we speak of a degenerate basic solution.

We obtain a characterization of basic solutions via Thm. 4.11:

Theorem 4.22. Let $K = P^{=}(A, b)$. Then, the following is equivalent:

1. $x \in K$ is an extreme point.
2. $x \in K$ is a vertex.
3. $x \in K$ is a feasible basic solution w.r.t. a basis B.

Proof. (1) $\Leftrightarrow$ (2) follows from Thm. 4.10.
(1) $\Rightarrow$ (3) : Let $x \in K$ be an extreme point. With Thm. 4.11 we obtain that all column vectors $a^j, j \in J$ mit $J := \{i \in \{1, \dots, n\} : x_i > 0\}$ are linearly independent and they can be extended to a basis B. Per definition of J we get $x_N = 0$, where N denotes the non-basis w.r.t. B. Hence,

$$b = Ax = A_B x_B,$$

and using that A_B is invertible

$$x_B = A_B^{-1} b.$$

Thus, x is a feasible basic solution w.r.t. basis B.
(3) $\Rightarrow$ (1) : Let x be a feasible basic solution to basis B, i.e., $x = (x_B, x_N)$ with $x_N = 0$ and $x_B := A_B^{-1} b \geq 0$. The set of indices with positive entries of x is a subset of B. Aas the vectors $a^i, i \in B$ are linearly independent, the statement follows by Thm. 4.11 . ☐

$$\text{vertex} \overset{\text{unique}}{\leftrightarrow} \text{feasible basic solution} \overset{\text{non-unique}}{\leftrightarrow} \text{basis.}$$

We obtain a corollary on the number of vertices of a polyhedron K with $\text{rank}(A) = m < n$.

Corollary 4.23. There are at most $\binom{n}{m}$ distinct vertices of K.

Proof. Every vertex leads to a feasible basic solution with a basic matrix. Every basic matrix has m linearly independent columns of A and there are $\binom{n}{m}$ different possibilities to choose m linearly independent columns from A. ☐

We close this chapter with showing that $\text{rank}(A) = m$ is w.l.o.g.

Theorem 4.24. Let $K = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$ be a non-empty polyhedron in standard form with matrix $A \in \mathbb{R}^{m \times n}$. Let $\text{rank}(A) = k < m$ and suppose that the row vectors $a_{i_1}, \dots, a_{i_k}$ are linearly independent. Consdider

$$P := \{x \in \mathbb{R}^n | a_{i_1} x = b_{i_1}, \dots, a_{i_k} x = b_{i_k}, x \geq 0\}.$$

Then, $K = P$.

Proof. W.l.o.g. $i_1 = 1, \dots, i_k = k$. Trivially $K \subseteq P$.

We show only $P \subseteq K$. With $\mathrm{rank}(A) = k$ we that the row space of A has dimension k and the vectors $a_1, \dots, a_k$ form a basis of that space. Hence, every row a_i of A can be represented as $a_i = \sum_{j=1}^{k} \lambda_{ij} a_j$ for scalars λ_{ij}. Let $x \in K$. We have

$$b_i = a_i x = \sum_{j=1}^{k} \lambda_{ij} a_j x = \sum_{j=1}^{k} \lambda_{ij} b_j, \; i = 1, \dots, m. \tag{4.1}$$

Let $y \in P$. For all $i = 1, \dots, m$ we get with (4.1)

$$a_i y = \sum_{j=1}^{k} \lambda_{ij} a_j y = \sum_{j=1}^{k} \lambda_{ij} b_j = b_i,$$

hence $y \in K$. □

4.4 Degeneracy of Basic Solutions

For a feasible basic solution, there are m inequalities active and usually also $n-m$ variables 0 and thus define additional $n-m$ active inequalities. If more than $n-m$ variables are equal 0, we speak of a degenerate basic solution, see Definition 4.21.
We give an example.

Example 4.25.
1. Redundant variables:

$$\begin{aligned} x_1 + x_2 &= 2 \\ x_3 &= 0 \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

2. Redundant inequalities:

$$\begin{array}{rrrrrl} 2x_1 & +x_2 & +y_1 & & & = 3 \\ x_1 & +2x_2 & & +y_2 & & = 3 \\ x_1 & +x_2 & & & +y_3 & = 2 \\ x_1, & x_2, & y_1, & y_2, & y_3 & \geq 0. \end{array}$$

In $x = (1, 1, 0, 0, 0)^T$ the ineq. $y_3 \geq 0$ is redundant.

3. geometric reasons (see Oktahedron in Fig. 4.5).

Remark 4.26. By disturbing a linear equation system with random noise, with high probability we get a non-degenerate problem.

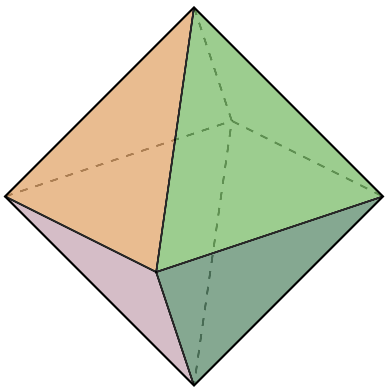


Figure 4.5: The vertices of the Oktahedron are degenerate.

Chapter 5

The Simplex Method

The most important method for solving an LP in standard form

$$\begin{aligned} \text{minimize } z(x) &= cx \\ \text{s.t.} & \\ Ax &= b \\ x &\geq 0. \end{aligned} \tag{5.1}$$

is the Simplex-Method. We assume $rang(A) = m < n$. With Thm. (4.19) we know that an optimal solution (if it exists) is a vertex of

$$K = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}.$$

The simplex-method consists of executing the following steps:

Geometric form:

1. Find vertex x of $P^=(A, b)$
2. Computing adjacent vertex y of x with smaller objective value. Replace x with y and repeat (2).
3. If (2) is not possible, there are three exclusive possibilities:
   - x is optimal.
   - LP is unbounded.
   - The iterate leads to a basis describing the same vertex, in this case repeat (2).

5.1 Parametrization of the Solution Space

For basis B we get that the linear equation system $Ax = b$ can be represented as

$$Ax = A_B x_B + A_N x_N = b.$$

This way, we obtain a parametrization of the $n-m$ dimensional solution space of $Ax = b$ via

$$x_B = A_B^{-1} b - A_B^{-1} A_N x_N,\ x_N \in \mathbb{R}^{n-m}, \tag{5.2}$$

where x_B is the dependent variable and x_N denotes the independent variable. Subdivide c in $c_B \in \mathbb{R}^m$ and $c_N \in \mathbb{R}^{n-m}$. Inserting (5.2) into the objective leads to

$$\begin{aligned} z(x) &= cx = c_B x_B + c_N x_N \\ &= c_B A_B^{-1} b - \left(c_B A_B^{-1} A_N - c_N\right) x_N \\ &=: z_0 - r_N x_N, \end{aligned} \tag{5.3}$$

where

$$\begin{aligned} z_0 &:= c_B A_B^{-1} b \\ r_N &:= c_B A_B^{-1} A_N - c_N = (r_j)_{j \in N} \in \mathbb{R}^{n-m} \text{ (vector of reduced costs)} \end{aligned}$$

With Thm. 4.22 the basic solutions describe the vertices of K. The representation (5.3) of the objective yields the following optimality criterion:

Theorem 5.1. Let B be a basis with

1. the corresponding solution x is feasible, i.e. $x_B \geq 0$,
2. $r_N = c_B A_B^{-1} A_N - c_N \leq 0$.

Then, x is optimal for the LP (5.1) and the optimal value is $z_0 = c_B A_B^{-1} b$.

Proof. For every feasible $\tilde{x}$ we get $\tilde{x}_B \geq 0$ and with (5.3) and using $r_N \leq 0$, we get

$$z(\tilde{x}) = c\tilde{x} = z_0 - r_N \tilde{x}_N \geq z_0 = z(x).$$

□

For a special case of linear inequality restrictions of the form

$$\min cx \text{ u.d.N. } Ax \leq b, x \geq 0, \text{ where } b \geq 0,$$

we get an equivalent reformulation

$$\min cx + 0y \text{ s.t. } Ax + y = \tilde{A}\begin{pmatrix} x \\ y \end{pmatrix} = b, x \geq 0, y \geq 0,$$

where $\tilde{A} = (A, I)$. Choose basis $B = \{n+1, \dots, n+m\}, N = \{1, \dots, n\}$ and we get a basic solution $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix} \in \mathbb{R}^{n+m}$. The reduced cost are $\tilde{A}_B = I, \tilde{c}_B = 0, \tilde{c}_N = c$ with

$$r_N = \tilde{c}_B \tilde{A}_B^{-1} \tilde{A}_N - \tilde{c}_N = -c \in \mathbb{R}^n$$
$$z_0 = 0.$$

5.2 Basis Exchange

The Simplex-Method is based on the sufficient optimality conditions of Thm. 5.1. We search for a basis B with

1. $x_B \geq 0$
2. $r_N \leq 0$.

We start with B satisfying 1. If 2. is violated, we go to an adjacent basis B' via a basis exchange step so that the objective value goes down.

Definition 5.2. Two basic solutions $x = (x_B, x_N)$ and $x' = (x'_{B'}, x'_{N'})$ are called adjacent, if $|B \cap B'| = m-1$, i.e., B and B' differ by the exchange of one basic and non-basic variable, respectively.

For B, we use (5.2)

$$x_B = A_B^{-1} b - A_B^{-1} A_N x_N, \; x_N \in \mathbb{R}^{n-m}.$$

Computing A_B^{-1} can be done with elementary row multiplication (Gauss-Jordan-Elimination). The following example demonstrate one exchange-step (pivot step):

Example 5.3.

$$\text{minimize } z(x_1, x_2, x_3) = -x_1 - 2x_2 - 3x_3$$
$$\text{s.t.}$$
$$2x_1 + x_2 + 5x_3 = 5$$
$$x_1 + 2x_2 + x_3 = 4$$
$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0.$$

We choose basis $B = (1, 2)$ and we use a tableau form $b|A$:

b	x_1	x_2	x_3	
5	2	1	5	$\cdot 1/2$
4	1	2	1	

b	x_1	x_2	x_3	
5/2	1	1/2	5/2	
4	1	2	1	-1. row

b	x_1	x_2	x_3	
5/2	1	1/2	5/2	
3/2	0	3/2	$-3/2$	$\cdot 2/3$

b	x_1	x_2	x_3	
5/2	1	1/2	5/2	$-1/2 \cdot$ 2. row
1	0	1	-1	

b	x_1	x_2	x_3
2	1	0	3
1	0	1	-1

We get:

$$x_B = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 3 \\ -1 \end{pmatrix} x_3 = A_B^{-1} b - A_B^{-1} A_N x_N.$$

The basic solution

$$x_1 = 2, x_2 = 1, x_3 = 0$$

is feasible but not optimal, as we have

$$r_N = r_3 = c_B A_B^{-1} A_N - c_N = (-1, -2) \begin{pmatrix} 3 \\ -1 \end{pmatrix} - (-3) = 2 > 0.$$

For $0 \leq x_3 \leq 2/3$ we get that $x_B \geq 0$ is feasible: geometrically, we follow an edge of K. For $x_3 = 2/3$ we get a new basic solution $x = (0, 5/3, 2/3)^\mathsf{T}$. The transition $x_3 = 0 \to x_3 = 2/3$

corresponds to a basic exchange step

$$B = (1, 2) \to B' = (3, 2), N = (3) \to N' = (1).$$

The non-basic variable x_3 is exchanged with the basic variable x_1: $x_3 = 2/3 - 1/3x_1$. For the new basis B' we get

$$x_{B'} = \begin{pmatrix} x_3 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2/3 \\ 5/3 \end{pmatrix} - \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix} x_1$$

with

$$r_{N'} = r'_1 = (-3, -2) \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix} - (-1) = -2/3 < 0.$$

Hence, the found solution is optimal.

We will formalize this idea now. Suppose

$$A_B = I, B = (1, \dots, m) \text{ and } N = (m+1, \dots, n)$$

and $x_B = b \geq 0$, i.e. the starting x is feasible. If $r_N \leq 0$, x is optimal with Thm. 5.1. Let $r_s > 0$ for some $s \in N$. In

$$x_B = A_B^{-1} b - A_B^{-1} A_N x_N, z(x) = z_0 - r_N x_N,$$

we insert

$$A_B = I, x_j = 0 \text{ for } j \in N \setminus \{s\}$$

and get

$$\begin{aligned} x_B &= b - a^s x_s \\ z(x) &= z_0 - r_s x_s. \end{aligned} \tag{5.4}$$

1. Case: $a^s \leq 0$.
In this case, we get

$$x_B = b - a^s x_s \geq 0 \text{ for } x_s \to \infty.$$

We get $z(x) = z_0 - r_s x_s \to -\infty$ for $x_s \to \infty$ and hence the problem has no finite solution; the polyhdron K is unbounded in a descent direction of the objective.
2. Case: $a_{is} > 0$ for some $i \in \{1, \dots, m\}$. Define $p \in B$ via

$$\frac{b_p}{a_{ps}} = \min \left\{ \frac{b_i}{a_{is}} \middle| a_{is} > 0, i = 1, \dots, m \right\}.$$

In order to ensure $x_B \geq 0$, the value of x_s can be at most

$$x_s = \frac{b_p}{a_{ps}} \Rightarrow x_p = 0.$$

This corresponds to an exchange of the non-basic variable x_s with the basic variable x_p (recall $B = (1, \dots, m)$). The value $a_{ps} > 0$ is called pivot element. we get

$$z(x) = z_0 - r_s \frac{b_p}{a_{ps}}. \tag{5.5}$$

We can differentiate the following cases:

(a) $b_p = 0$: basic solution is degenerate, $z(x) = z_0$

(b) $b_p > 0$: We get a strict improvement $z(x) < z_0$

We get the following necessary optimality condition:

Corollary 5.4. Let the basic solution x with $x_B \geq 0$ and $x_N = 0$ optimal. If $x_B > 0$, then $r_N \leq 0$.

The index $s \in N$ (pivot column) can be computed as follows:

1. (Rule of Dantzig): Choose smallest $s \in N$ with:

$$r_s = \max_{j \in N} r_j. \tag{5.6}$$

   Choose smallest $p \in B$ with

$$\frac{b_p}{a_{ps}} = \min \left\{ \frac{b_i}{a_{is}} \middle| a_{is} > 0, i = 1, \dots, m \right\}. \tag{5.7}$$

2. (Rule of Bland): Choose smallest $s \in N$ with $r_s > 0$. Choose smallest $p \in B$, so that b_p is smallest and $\frac{b_p}{a_{ps}} = \min \left\{ \frac{b_i}{a_{is}} \middle| a_{is} > 0, i = 1, \dots, m \right\}$.

Let us describe a basis exchange after executing the Gauss-Jordan elimination. That is, we start with

$$A_B = I, B = (1, \dots, m), N = (m + 1, \dots, n),$$

and thus

$$x_i = b_i - \sum_{j \in N} a_{ij} x_j, \; i \in B. \tag{5.8}$$

The p-th equation can be solved for x_s using $a_{ps} > 0$:

$$x_p = b_p - \sum_{j \in N} a_{pj} x_j = b_p - \sum_{j \in N, j \neq s} a_{pj} x_j - a_{ps} x_s$$

Solving for x_s yields:

$$x_s = \frac{b_p}{a_{ps}} - \sum_{j \in N, j \neq s} \frac{a_{pj}}{a_{ps}} x_j - \frac{1}{a_{ps}} x_p. \tag{5.9}$$

Insert x_s in the other equations (5.8) with $i \neq p$:

$$x_i = \underbrace{b_i - a_{is}\frac{b_p}{a_{ps}}}_{=b_i'} - \sum_{j\in N, j\neq s}\left(a_{ij} - a_{is}\frac{a_{pj}}{a_{ps}}\right)x_j + \frac{a_{is}}{a_{ps}}x_p. \tag{5.10}$$

Define new basis: Transition

$$B = (1, \ldots, m) \to B' = (1, \ldots, p-1, s, p+1, \ldots, m).$$

New non-basis:

$$N = (m+1, \ldots, n) \to N' = (m+1, \ldots, s-1, p, s+1, \ldots, n).$$

have the form

$$x_{B'} = \begin{pmatrix} x_1 \\ \vdots \\ x_{p-1} \\ x_s \\ x_{p+1} \\ \vdots \\ x_m \end{pmatrix} = b' - A'_{N'}x_{N'},$$

and hence have the form (5.2). The elements of the matrix

b'	$A'_{N'}$

are given via

Pivot element: reciprocal value :

$$a'_{ps} = \frac{1}{a_{ps}}$$

Other row p: divide by pivot element:

$$a'_{pj} = \frac{a_{pj}}{a_{ps}}, j \neq s$$
$$b'_p = \frac{b_p}{a_{ps}}$$

Other column s: divide by negative pivot element:

$$a'_{is} = -\frac{a_{is}}{a_{ps}}, i \neq p$$

Other elements: subtract the a_{is}-multiple of the new row p from the i-th row:

$$a'_{ij} = a_{ij} - a_{is}\frac{a_{pj}}{a_{ps}} = a_{ij} - a_{is}a'_{pj}, i \neq p, j \neq s$$

$$b'_i = b_i - a_{is}\frac{b_p}{a_{ps}} = b_i - a_{is}b'_p, i \neq p$$

We give an example.

	b	x_1	x_2	x_3	x_4	x_5	x_6
x_1	5	1	0	0	(1)	1	−1
x_2	3	0	1	0	2	−3	1
x_3	−1	0	0	1	−1	2	−1

Note that the current basic solution is infeasible. Change non-basic variable x_4 with basic variable x_1:

	b'	x_1	x_2	x_3	x_4	x_5	x_6
x_4	5	1	0	0	1	1	−1
x_2	−7	−2	1	0	0	−5	3
x_3	4	1	0	1	0	3	−2

With the choice of the pivot element, we did not follow any of the rules of Dantzig or Bland and indeed: x_2 becomes negative.

As the identity matrix $A_B = I$ contains no new information if we know B, N, A_N, b, we will consider in the following only the reduced tableau.

	b'	x_1	x_5	x_6
x_4	5	1	1	−1
x_2	−7	−2	−5	3
x_3	4	1	3	−2

We can verify:

$$b'_2 = 3 - 2 \cdot 5 = -7$$
$$b'_3 = -1 - (-1) \cdot 5 = 4$$
$$x_{B'} = \begin{pmatrix} x_4 \\ x_2 \\ x_3 \end{pmatrix} = b' = \begin{pmatrix} 5 \\ -7 \\ 4 \end{pmatrix}$$
$$B' = (4, 2, 3)$$

5.3 Objective Function and Reduced Costs

Via elimination of x_s in (5.9), we get

$$\begin{aligned} z(x) &= z_0 - \sum_{j\in N} r_j x_j \\ &= z_0 - \sum_{j\in N, j\neq s} r_j x_j - r_s\left(\frac{b_p}{a_{ps}} - \sum_{j\in N, j\neq s} \frac{a_{pj}}{a_{ps}} x_j - \frac{1}{a_{ps}} x_p\right) \\ &= z_0 - r_s\frac{b_p}{a_{ps}} - \sum_{j\in N, j\neq s}\left(r_j - r_s\frac{a_{pj}}{a_{ps}}\right) x_j + \frac{r_s}{a_{ps}} x_p \\ &=: z_0' - \sum_{j\in N'} r_j' x_j \end{aligned}$$

with

$$\begin{aligned} z_0' &:= z_0 - r_s\frac{b_p}{a_{ps}} \\ r_p' &:= -\frac{r_s}{a_{ps}} \\ r_j' &:= r_j - r_s\frac{a_{pj}}{a_{ps}} = r_j - r_s a_{pj}',\ j \neq p. \end{aligned} \tag{5.11}$$

The row (z_0, r_N) will be handled as the rows in

b	A_N

This leads to the extended tableau form:

		$x_j, j \in N$
$z(x)$	z_0	r_N
$x_i, i \in B$	b	A_N

(5.12)

Figure 5.1: Extended Simplex-Tableau.

We give an example.

Example 5.5.

$$\begin{gathered} \text{minimize } z(x_1, x_2, x_3) = -x_1 - 2x_2 - 3x_3 \\ \text{s.t.} \\ 2x_1 + x_2 + 5x_3 = 5 \\ x_1 + 2x_2 + x_3 = 4 \end{gathered}$$

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0.$$

For $B = (1,2), N = (3)$ we can transform the linear equation system to $A_B = I$:

$$\begin{array}{|c|c|}\hline b & A_N \\ \hline\end{array} = \begin{array}{|c|c|}\hline 2 & 3 \\ 1 & -1 \\ \hline\end{array}, \quad x_B = b = \begin{pmatrix} 2 \\ 1 \end{pmatrix} > 0$$

$$c = (-1,-2,-3)$$

$$z_0 = c_B b = (-1,-2)\begin{pmatrix} 2 \\ 1 \end{pmatrix} = -4$$

$$r_3 = c_B A_B^{-1} A_N - c_3 = 2 > 0.$$

We get the Simplex-Tableau:

		x_N
$z(x), r_N$	-4	2
x_1	2	3
x_2	1	-1

Exchange x_3 with x_1: $r_1' = -2/3 < 0$, hence the basic solution $x_1 = 0, x_2 = 5/3, x_3 = 2/3$

		x_N
$z(x), r_N$	$-16/3$	$-2/3$
x_3	$2/3$	$1/3$
x_2	$5/3$	$1/3$

is optimal.
The algorithmic execution of the Simplex-Method is illustrated in Fig..5.1 with a corresponding tableau-matrix.

$$T = (t_{ij}) = \begin{array}{|c|c|}\hline z_0 & r_N \\ \hline b & A_N \\ \hline\end{array} \qquad \begin{array}{c} 0 \leq i \leq m \\ 0 \leq j \leq n-m \end{array} \tag{5.13}$$

Now, we formally describe the simplex-method (with the rule of Dantzig).

1. **Start**: Let B be a basis with $A_B = I$. Suppose the corresponding basic solution is feasible, i.e., $x_B = b \geq 0$. Compute with N the tableau-matrix $T = (t_{ij})$ in (5.13). Denote with $B(i)$ the basis index corresponding to the i-th row of T for $1 \leq i \leq m$. Analogously denote with $N(j)$ the non-basis index corresponding to the j-th column of T for $1 \leq j \leq n - m$.

2. If $t_{0s} > 0$ for all $1 \leq s \leq n - m$, go to 3. Otherwise, the current basis-solution is optimal. Set

$$\begin{aligned} x_{B(i)} &:= t_{i0}, 1 \leq i \leq m \\ x_{N(j)} &:= 0, 1 \leq j \leq n - m \\ z &:= t_{00} \end{aligned}$$

3. **Compute the Exchange-Column**: Choose index $1 \leq s \leq n - m$ (with $N(s)$ smallest) with

$$t_{0s} := \max_{1 \leq j \leq n-m} t_{0j}.$$

4. If $t_{is} \leq 0$, for all $1 \leq i \leq m$, there is no finite solution. **Stop.**
   If there is $t_{is} > 0$ for some $1 \leq i \leq m$, go to step 5.

5. **Compute the Exchange-Row**: Choose index $1 \leq p \leq m$ (with $N(p)$ smallest) with

$$\frac{t_{p0}}{t_{ps}} = \min \left\{ \frac{t_{i0}}{t_{is}} \middle| t_{is} > 0, i = 1, \ldots, m \right\}.$$

   Go to step 6.

6. Exchange the s-th element of N with the p-th element of B:
   $B \leftarrow \big(B(1), \ldots, B(p-1), N(s), B(p+1) \ldots, B(m)\big)$
   $N \leftarrow \big(N(1), \ldots, N(s-1), B(p), N(s+1) \ldots, N(n-m)\big)$
   Execute pivot operation with pivot element $t_{ps} > 0$:

   - pivot elemet: $t'_{ps} := 1/t_{ps}$
   - pivot row: $t'_{pj} := \frac{t_{pj}}{t_{ps}}, j = 0, 1, \ldots, n-m, j \neq s$
   - pivot column: $t'_{is} := -\frac{t_{is}}{t_{ps}}, i = 0, 1, \ldots, m, i \neq p$
   - other elements: $t'_{ij} := t_{ij} - t_{is}\frac{t_{pj}}{t_{ps}}, i \neq p, j \neq s$
   - set $t_{ij} := t'_{ij}$ and go to step 2.

Figure 5.2: Formal Execution of the Simplex-Method.

5.4 Application to LPs with Inequalities

The LP

$$\min\{cx | Ax \leq b, x \geq 0\}, x \in \mathbb{R}^n, b \in \mathbb{R}^n, b \geq 0$$

is equivalent to

$$\min\{cx + 0 \cdot y | Ax + y = b, x \geq 0, y \geq 0\}, \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n+m}, b \in \mathbb{R}^n$$

Start: $B = (n+1, \dots, n+m), N = (1, \dots, n)$

Basic solution: $x = 0, y = b \geq 0, z_0 = 0, r_N = -c$. The corresponding $(m+1) \times (n+1)$ dimensional Simplex-Tableau:

z	0	$-c$
y	b	A

We solve the example from Subsec. 3.1.1.

$$\text{maximize } z(x_1, x_2) = 16x_1 + 32x_2$$
$$\text{s.t.}$$
$$20x_1 + 10x_2 \leq 8000$$
$$4x_1 + 5x_2 \leq 2000$$
$$6x_1 + 15x_2 \leq 4500$$
$$x_1 \geq 0$$
$$x_2 \geq 0$$

The geometric sequence $A \to B \to C \to D$ is executed algebraically via the Tableau-Form.

		x_1	x_2
z	0	16	32
y_1	8000	(20)	10
y_2	2000	4	5
y_3	4500	6	15

Let us choose 20 as pivot element and we exchange x_1 with y_1: This corresponds to $A \to B$.

		y_1	x_2
z	-6400	$-4/5$	24
x_1	400	$1/20$	$1/2$
y_2	400	$-1/5$	$\textcircled{3}$
y_3	2100	$-3/10$	12

Then, we exchange x_2 with y_2: This leads to $B \to C$.

		y_1	y_2
z	-9600	$4/5$	-8
x_1	$1000/3$	$1/12$	$-1/6$
x_2	$400/3$	$-1/15$	$1/3$
y_3	500	$\textcircled{1/2}$	-4

Exchange y_1 with y_3: This is $C \to D$.

		y_3	y_2
z	-10400	$-8/5$	$-8/5$
x_1	250	$-1/6$	$1/2$
x_2	200	$2/15$	$-1/5$
y_1	1000	2	-8

Here $r_N = (-8/5, -8/5) < 0$, hence $x_1 = 250, x_2 = 200, z_{min} = -10400$, i.e. $z_{max} = 10400$, is optimal.

We show next that for non-degenerate problems the simplex-method always terminates.

Theorem 5.6. If, during the execution of the simplex-method, all computed basic-solutions are non-degenerate, the method terminates.

Proof. For every pivot step j we get

$$z_0^j = z_0^{j-1} - r_s^{j-1} \frac{b_p^{j-1}}{a_{ps}^{j-1}}.$$

As for every j per assumption $b_p^j > 0$, we get a strictly monotone sequence

$$z_0^1 > z_0^2 > \dots$$

Hence, no basis is visited twice and since there are only finitely many solutions, the algorithm terminates. □

5.5 Cycling of the Simplex-Method and Lexicographical Pivoting

For a pivot operation with $b_p = 0$ the objective value does not change

$$\begin{aligned} z(x) &= z_0 - r_s \frac{b_p}{a_{ps}} = t_{00} - t_{0s} \frac{t_{p0}}{t_{ps}} \\ &= z_0 \end{aligned}$$

There are examples showing that the simplex-method (using a "wrong" pivot-rule) may cycle forever (see exercise).
In order to avoid cycling, we will now introduce a lexicographical variant.

Definition 5.7. Let $k \in \mathbb{N}$. Then, we can define a total order on $\mathbb{R}^k$ via

$$x \succeq y :\Leftrightarrow x = y \text{ or } x_i > y_i \text{ for } i = \min\{s | x_s \neq y_s\}$$

that is compatible with addition on $\mathbb{R}^k$. This order is called lexicographical order.

If $0 \in \mathbb{R}^k$ and $x \in \mathbb{R}^k$ with $x \succ 0$ (that means $x \succeq 0$ and $x \neq 0$), then x is called lexikographically positive.
We consider now the complete tableau in standard form:

$$T = (t_{ij})_{\substack{i=0,\dots,m \\ j=0,\dots,n}} = \begin{array}{|c|c|c|} \hline z_0 & 0 & r_N \\ \hline b & A_B = I & A_N \\ \hline \end{array}$$

The rows $t_i, i = 1, \dots, m$ are vectors in $\mathbb{R}^{n+1}$. For basis B, we denote by $B(i)$ the i-th entry of the current basis B. Initially $B = (1, \dots, m)$ and $N = (m+1, \dots, n)$.

Definition 5.8 (Lexicographical Rule (LEX)). Let B be the current basis and denote the corresponding tableau by T.

- Choose arbitrary column index $s \in N$ with $r_s = t_{0s} > 0$.
- Condsider $I = \{i \in \{1, \dots, m\} | t_{is} > 0\}$. Choose $p \in I$ with

$$\frac{t_p}{t_{ps}} = \text{lexmin}\left\{ \frac{t_i}{t_{is}} \middle| i \in I \right\}.$$

This lex. minimum satisfies

$$\frac{t_i}{t_{is}} \succ \frac{t_p}{t_{ps}} \text{ for all } i \in I \setminus \{p\}.$$

Then, $B(p)$ leaves the current basis B and s enters the new basis, that is, $B'(p) = s$.

By choice of p, we get

$$\frac{t_{p0}}{t_{ps}} = \min \left\{ \frac{t_{i0}}{t_{is}} \middle| i \in I \right\}.$$

Hence, the pivot step leads to a feasible new basis.

Theorem 5.9. Suppose we start the simplex-method using LEX with a tableau that contains only lex. positive rows (except the reduced cost row). Then, the method terminates.

Proof. We first claim that the assumption of lex. positive rows of the theorem are easily satisfiable via applying the Gauss-Jordan transformation, i.e., $(b|I|A_N)$ (except first row). Because of feasibility $x_B \geq 0$, the first column contains only numbers greater equal than 0 (except first row). As the first sub-matrix is lex. positive, the claim is shown.
We show inductively over the execution of LEX, that this property is preserved. Consider a basis B with the required property. Let s and p be chosen according to LEX. Denote the new basis by B'. For the p-th row t'_p of the new tableau T' we get

$$t'_p = \frac{t_p}{t_{ps}} \succ 0,$$

since by assumption $t_p \succ 0$ and $t_{ps} > 0$. For $i \in \{1, \dots, m\}, i \neq p$, the i-th row t'_i of T' is given as:

$$t'_i = t_i - \frac{t_{is}}{t_{ps}} t_p.$$

For $t_{is} > 0$, we get according to LEX

$$t'_i = t_{is} \left(\frac{t_i}{t_{is}} - \frac{t_p}{t_{ps}} \right) \succ 0.$$

For $t_{is} \leq 0$, we get

$$t'_i = t_i + \frac{|t_{is}|}{t_{ps}} t_p \succeq t_i \succ 0.$$

Overall, the rows remain lex. positive.
Let us consider the first row t_0 and t'_0. We have

$$t'_0 = t_0 - \frac{t_{0s}}{t_{ps}} t_p.$$

With $t_{0s} = r_s > 0$ we get

$$t_0 = t'_0 + \frac{t_{0s}}{t_{ps}} t_p \succ t'_0.$$

Thus, after executing a pivot step using LEX, the 0-th row decreases strictly lexicographically and hence no tableau can appear twice. □

5.6 Computing a Feasible Basic Solution - Two-Phase Method

Consider the LP in standard form:

$$\begin{aligned} &\text{minimize } z(x) = cx \\ &\text{s.t.} \\ &Ax = b \\ &x \geq 0. \end{aligned} \tag{5.14}$$

W.l.o.g. assume $b \geq 0$ ist. If there is some row with negative b_i just multiply with -1. We define the following auxiliary problem:

$$\begin{aligned} &\text{minimize } z(x, y) = \sum_{i=n+1}^{n+m} y_i \\ &\text{s.t.} \\ &Ax + y = b \\ &x, y \geq 0. \end{aligned} \tag{5.15}$$

We use here additional artificial variables y_i, where $i \in \{n+1, \dots, n+m\}$. One of the following statements holds true.

1. The auxiliary problem has an optimal solution with value 0. Then, the system $Ax = b, x \geq 0$ is feasible.
2. The auxiliary problem has an optimal solution with value > 0. Then, $Ax = b, x \geq 0$ is infeasible.

The auxiliary problem can be solved with the simplex-method, where we use as start basis $y_i, i \in \{n+1, \dots, n+m\}$ (recall $b \geq 0$). We write

$$B = (n+1, \dots, n+m), N = (1, \dots, n), c_B = (1, 1, \dots, 1) \in \mathbb{R}^m \text{ and } c_N = 0 \in \mathbb{R}^n.$$

This yields

$$\begin{aligned} z_0 &= c_B b = \sum_{i=1}^{m} b_i \geq 0, \\ r_N &= c_B A_N, \end{aligned} \tag{5.16}$$

where $r_j = \sum_{i=1}^{m} a_{ij}, j = 1, \dots, n$. The start tableau reads

$$\begin{array}{c|c|c|} & \multicolumn{2}{c}{\qquad x} \\ \cline{2-3} z(x,y), r_N & z_0 & r_N \\ \cline{2-3} y & b & A_N \\ \cline{2-3} \end{array} \tag{5.17}$$

The two-phase method for LP (5.14) consists of the following steps:

Phase I: Solve (5.15). The basic variables $y_{n+1}, \dots, y_{n+m}$ need to be non-basic variables in the optimal tableau. Then, we get a feasible solution for the original problem.

Phase II: Compute starting tableau for LP (5.14) via

- Delete the columns belonging to $y_{n+1}, \dots, y_{n+m}$,
- Compute z_0 and r_N.

Example 5.10. We apply the two-phase method to the following problem:

$$\begin{array}{rrcrcrcl} \min & 4x_1 & + & x_2 & + & x_3 & & \\ \text{u.d.N.} & 2x_1 & + & x_2 & + & 2x_3 & = & 4 \\ & 3x_1 & + & 3x_2 & + & x_3 & = & 3 \\ & x_1 & , & x_2 & , & x_3 & \geq & 0 \end{array}$$

The auxiliary problem reads as:

$$\begin{array}{rrcrcrcrcrcl} \min & & & & & & & y_4 & + & y_5 & & \\ \text{u.d.N.} & 2x_1 & + & x_2 & + & 2x_3 & + & y_4 & & & = & 4 \\ & 3x_1 & + & 3x_2 & + & x_3 & & & + & y_5 & = & 3 \\ & x_1 & , & x_2 & , & x_3 & , & y_4 & , & y_5 & \geq & 0 \end{array}$$

We start with basis $B = (4, 5)$ and $N = (1, 2, 3)$.

$$z_0 = b_1 + b_2 = 4 + 3 = 7 \quad \text{and} \quad r_N = \left(\sum_{i=1}^{2} a_{ij} \right)_{j=1,2,3} = (5, 4, 3).$$

The start tableau reads as:

		x_1	x_2	x_3
$z(x,y), r_N$	7	5	4	3
y_4	4	2	1	2
y_5	3	**3**	3	1

With the rule of Dantzig we get:

		y_5	x_2	x_3
$z(x,y), r_N$	2	-5/3	-1	4/3
y_4	2	-2/3	-1	**4/3**
x_1	1	1/3	1	1/3

		y_5	x_2	y_4
$z(x,y), r_N$	0	-1	0	-1
x_3	3/2	-1/2	-3/4	3/4
x_1	1/2	1/2	5/4	-1/4

We see that $r_N \leq 0$ and $z(x,y) = 0$, where y_4 and y_5 non-basic.Hence, a feasible solution to $Ax = b, x \geq 0$ is given via basis $B = (3,1)$ with

$$x_3 = 3/2, \quad x_1 = 1/2, \quad x_2 = 0.$$

Note that the basis matrix A_B is already given as the unit-matrix and hence phase II can be easily started. We delete the column vectors corresponding to y_4 and y_5. We get $c = (4,1,1)$, $B = (3,1)$, $N = (2)$, $c_B = (1,4)$ and $b = \binom{3/2}{1/2}$ (see tableau). For the new costs and the reduced costs we get:

$$z_0 = c_B b = (1,4)\begin{pmatrix} 3/2 \\ 1/2 \end{pmatrix} = 3/2 + 4/2 = 7/2 \text{ and}$$
$$r_N = r_2 = c_B a^2 - c_2 = (1,4)\begin{pmatrix} -3/4 \\ 5/4 \end{pmatrix} - 1 = 13/4,$$

where a^2 needs to be taken from the tableau.

The start tableau for phase II reads as:

		x_2
$z(x), r_N$	7/2	13/4
x_3	3/2	-3/4
x_1	1/2	5/4

After one pivot step we get

		x_1
$z(x), r_N$	11/5	-13/5
x_3	9/5	3/5
x_2	2/5	4/5

As $r_1 = -13/5 < 0$, the solution is optimal

$$x_1 = 0, \quad x_2 = 2/5, \quad x_3 = 9/5 \quad \text{with objective } 11/5.$$

Let us turn to problems with inequality constraints.

$$\begin{aligned} \text{minimize } z(x) &= cx \\ \text{s.t.} & \\ a_i x = b_i, & \quad i = 1, \ldots, k \\ a_i x \leq b_i, & \quad i = k+1, \ldots, m \\ x \geq 0. & \end{aligned} \tag{5.18}$$

We introduce artificial variables $y_{n+1}, \ldots, y_{n+k}$ for the first k equations and slack-variables $x_{n+k+1}, \ldots, x_{n+m}$ for the inequalities. Let us assume $b \geq 0$. Then, we need to solve in phase I the following auxiliary problem (as slack-variables can be ≥ 0, they don't appear in the objective):

$$\begin{aligned} \text{minimize } & \sum_{i=n+1}^{n+k} y_i \\ \text{s.t.} & \\ a_i x + y_{n+i} = b_i, & \quad i = 1, \ldots, k \\ a_i x + x_{n+i} = b_i & \quad i = k+1, \ldots, m \\ x, y \geq 0. & \end{aligned} \tag{5.19}$$

The start tableau reads as in 5.17, where

$$z_0 = \sum_{i=1}^{k} b_i \text{ and } r_j = \sum_{i=1}^{k} a_{ij}, j = 1, \dots, n.$$

Note that we assumed $b \geq 0$ which is this time not w.l.o.g. Let us consider the general case now:

$$\begin{aligned} \text{minimize } z(x) &= cx \\ \text{s.t.} & \\ a_i x \geq b_i, &\quad i = 1, \dots, k \\ a_i x \leq b_i, &\quad i = k+1, \dots, m \\ x \geq 0. & \end{aligned} \tag{5.20}$$

Here we can assume w.l.o.g. $b \geq 0$ ist. We introduce slack-variables $x_{n+1}, \dots, x_{n+m}$ and obtain a problem in standard form:

$$\begin{aligned} \text{minimize } z(x) &= cx \\ \text{s.t.} & \\ a_i x - x_{n+i} = b_i, &\quad i = 1, \dots, k \\ a_i x + x_{n+i} = b_i, &\quad i = k+1, \dots, m \\ x_i \geq 0, &\quad i = 1, \dots, n+m. \end{aligned} \tag{5.21}$$

We use the artificial variables $y_{n+m+1}, \dots, y_{n+m+k}$ for the first k equations (since here the slack-variables have a negative sign).
The auxiliary problem reads as

$$\begin{aligned} \text{minimize } & \sum_{i=n+m+1}^{n+m+k} y_i \\ \text{s.t.} & \\ a_i x - x_{n+i} + y_{n+m+i} = b_i, &\quad i = 1, \dots, k \\ a_i x + x_{n+i} = b_i, &\quad i = k+1, \dots, m \\ x_i \geq 0, &\quad i = 1, \dots, n+m \\ y_i \geq 0, &\quad i = n+m+1, \dots, n+m+k. \end{aligned} \tag{5.22}$$

Let us sort the $n + m + k$ variables as

$$\underbrace{x_1, \dots, x_n, x_{n+1}, \dots, x_{n+k}}_{N}, \underbrace{y_{n+m+1}, \dots, y_{n+m+k}, x_{n+k+1}, \dots, x_{n+m}}_{B}.$$

Then, we get

$$c_B = (\underbrace{1, 1, \ldots, 1}_{k}, \underbrace{0, 0, \ldots, 0}_{m-k}) \in \mathbb{R}^m,$$

$$z_0 = c_B b = \sum_{i=1}^{k} b_i,$$

$$r_j = c_B a^j - c_j = \begin{cases} \sum_{i=1}^{k} a_{ij}, & j = 1, \ldots, n \\ -1, & j = n+1, \ldots, n+k \end{cases}$$

Finally, we can compute the start tableau or phase I:

		$x_1 \cdots x_n$	$x_{n+1} \cdots x_{n+k}$
$z(x), r_N$	$\sum_{i=1}^{k} b_i$	$\sum_{i=1}^{k} a_{ij}$	$-1 \cdots -1$
y_{n+m+1}	b_1	a_1	
$\vdots$	$\vdots$	$\vdots$	$-I_k$
y_{n+m+k}	b_k	a_k	
x_{n+k+1}	b_{k+1}	a_{k+1}	
$\vdots$	$\vdots$	$\vdots$	0
x_{n+m}	b_m	a_m	

Remark 5.11. If the found vertex after phase 1 is degenerate, that is, some artificial variables are still in the basis, then we need to perform additional pivots until we reach a feasible solution for $Ax = b, x \geq 0$. Such pivot steps just change the basis, not the vertex itself.

Chapter 6

LP Duality

We consider an LP in standard form

$$\begin{aligned} \text{minimize } z(x) &= cx \\ Ax &= b \\ x &\geq 0, \end{aligned} \tag{P}$$

for $A \in \mathbb{R}^{m \times n}, c^T \in \mathbb{R}^n, b \in \mathbb{R}^m$. Denote

$$K = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}.$$

Let $x \in K$ and consider $\lambda^T \in \mathbb{R}^m$. We obtain

$$b = Ax \Leftrightarrow \lambda b = \lambda Ax \ \forall \lambda^T \in \mathbb{R}^m.$$

We get a lower bound on cx via:

$$\lambda b \leq cx \ \forall \lambda^T \in \mathbb{R}^m \text{ with } \lambda A \leq c.$$

Let us conmpute the best lower bound on cx. This leads to the dual problem:

$$\begin{aligned} \text{maximize } z^*(\lambda) &= \lambda b \\ \lambda A &\leq c. \end{aligned} \tag{D}$$

We denote by K^* the polyhedron for the dual problem.

$$K^* := \{\lambda^T \in \mathbb{R}^m | \lambda A \leq c\}.$$

Note that λ^T has no sign constraints.

6.1 Dual Programms

Every LP has a corresponding dual LP. One can differentiate between a symmetric and asymmetric form.

Symmetric Form:

Primal problem (P-sym):

$$\begin{aligned} \text{minimize } & cx \\ Ax &\geq b \\ x &\geq 0, \end{aligned} \tag{P-sym}$$

Dual problem (D-sym):

$$\begin{aligned} \text{maximize } & \lambda b \\ \lambda A &\leq c \\ \lambda &\geq 0. \end{aligned} \tag{D-sym}$$

Here, A is an $m \times n$ matrix, $x \in \mathbb{R}^n$ is a column vektor, and $\lambda \in \mathbb{R}^m$ a row vector.

Remark 6.1. The dual problem of the dual problem is the primal problem. Proof: Exercise.

We give an example.

Primal problem (P):

$$\begin{aligned} \text{minimize } & 3x_1 + 4x_2 + 5x_3 \\ x_1 + 2x_2 + 3x_3 &\geq 5 \\ 2x_1 + 2x_2 + x_3 &\geq 6 \\ x_i &\geq 0, i = 1, 2, 3 \end{aligned}$$

Dual problem (D):

$$\begin{aligned} \text{maximize } & 5\lambda_1 + 6\lambda_2 \\ \lambda_1 + 2\lambda_2 &\leq 3 \\ 2\lambda_1 + 2\lambda_2 &\leq 4 \\ 3\lambda_1 + \lambda_2 &\leq 5 \\ \lambda_i &\geq 0, i = 1, 2. \end{aligned}$$

The LP in standard form can be reduced to the symmetric form. The equation system

$$Ax = b$$

corresponds to

$$\begin{aligned} Ax &\geq b \\ -Ax &\geq -b. \end{aligned}$$

The dual problem for $w \in \mathbb{R}^{2m}$ is then using (P-sym)

$$\begin{aligned} \text{maximize } w & \begin{pmatrix} b \\ -b \end{pmatrix} \\ w & \begin{pmatrix} A \\ -A \end{pmatrix} \leq c \\ & w \geq 0. \end{aligned}$$

Insert $w := (u, v), u, v \in \mathbb{R}^m$, and we obtain

$$\begin{aligned} \text{maximize } ub - vb \\ uA - vA \leq c \\ u \geq 0, v \geq 0. \end{aligned}$$

With $\lambda := u - v$ we obtain the dual problem (D). Note that $\lambda = u - v$ has no sign constraints.

Here are some computing rules for dualizing.

Table 6.1: Dualizing rules.

Primal LP	Dual LP
(P_1) $\max cx, Ax \leq b, x \geq 0$	(D_1) $\min \lambda b, \lambda A \geq c, \lambda \geq 0$
(P_2) $\min cx, Ax \geq b, x \geq 0$	(D_2) $\max \lambda b, \lambda A \leq c, \lambda \geq 0$
(P_3) $\max cx, Ax = b, x \geq 0$	(D_3) $\min \lambda b, \lambda A \geq c$
(P_4) $\min cx, Ax = b, x \geq 0$	(D_4) $\max \lambda b, \lambda A \leq c$
(P_5) $\max cx, Ax \leq b$	(D_5) $\min \lambda b, \lambda A = c, \lambda \geq 0$
(P_6) $\min cx, Ax \geq b$	(D_6) $\max \lambda b, \lambda A = c, \lambda \geq 0$

We give a general version now.

Lemma 6.2. Consider matrices A, B, C, D and vectors a, b, c, d. For the primal LP

$$\begin{aligned} \text{maximize } cx + dy \\ Ax + By \leq a \\ Cx + Dy = b \\ x \geq 0, \end{aligned} \tag{LP}$$

we get the dual LP

$$\begin{aligned} \text{minimize } ua + vb & \\ uA + vC &\geq c \\ uB + vD &= d \\ u &\geq 0. \end{aligned} \tag{DP}$$

Proof. Exercise. □

6.2 The Strong Duality Theorem

$$\begin{aligned} \text{minimize } z(x) &= cx \\ Ax &= b \\ x &\geq 0, \end{aligned} \tag{P}$$

$$\begin{aligned} \text{maximize } z^*(\lambda) &= \lambda b \\ \lambda A &\leq c. \end{aligned} \tag{D}$$

The feasible sets of the LPs are denoted by K and K^*.
We get the following relationshship of (P) and (D).

Theorem 6.3 (Weak-Duality). For $x \in K, \lambda^T \in K^*$ we have $\lambda b \leq cx$.

Proof. Let $x \in K$ and $\lambda^T \in K^*$. We get

$$b = Ax \Rightarrow \lambda b = \lambda Ax \underset{\lambda^T \in K^*}{\leq} cx.$$

□

Corollary 6.4. Let $x^* \in K$ and $(\lambda^*)^T \in K^*$ with $\lambda^* b = cx^*$. Then, x^* and $(\lambda^*)^T$ are optimal for the respective problems (P) and (D).

Proof. For arbitrary $x \in K$ we get

$$cx \geq \lambda^* b = cx^*.$$

For arbitrary $\lambda \in K^*$

$$\lambda b \leq cx^* = \lambda^* b.$$

□

Theorem 6.5 (Strong Duality for Linear Optimization). If one of the problems (P) or (D) admits a finite optimal solution, then also the other admits a finite optimal solution and their objective values are equal:

$$\min\{cx | x \in K\} = \max\{\lambda b | \lambda^T \in K^*\}.$$

Proof. W.l.o.g. $x^* \in K$ is optimal with $cx^* = z_0$ finite. Define

$$C := \{(r, w) = t(cx - z_0, b - Ax)^T | x \geq 0, t > 0\} \subset \mathbb{R}^{m+1}.$$

C is a convex and closed cone containing 0. We get the following alternative representation of C:

$$C = \{(r, w) = (cx - tz_0, tb - Ax)^T | x \geq 0, t > 0\} \subset \mathbb{R}^{m+1}.$$

(Define $\tilde{x} = x/t$).

We claim that $(-1, 0) \notin C$. This follows since $t > 0$ and $z_0 \leq cx$ for all $x \in K$. Hence, we can us the separation theorem for convex and closed cones (cf. Thm. 2.9). There is $(\lambda_0, \lambda)^T \in \mathbb{R}^{m+1} \setminus \{0\}$ with $\lambda_0 \in \mathbb{R}, \lambda^T \in \mathbb{R}^m$ such that

$$(\lambda_0, \lambda) \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -\lambda_0 < 0 \leq (\lambda_0, \lambda) \begin{pmatrix} cx - tz_0 \\ tb - Ax \end{pmatrix}$$

for all $t > 0, x \geq 0$. We get $\lambda_0 > 0$ and w.l.o.g. $\lambda_0 = 1$ (multiply with $1/\lambda_0$). Computing the scalar products, we get for $t > 0, x \geq 0$

$$\begin{aligned} 0 &\leq (cx - tz_0) + \lambda(tb - Ax) \\ \Leftrightarrow tz_0 - cx &\leq \lambda(tb - Ax). \end{aligned} \tag{6.1}$$

Take limit $t \downarrow 0$:

$$cx \geq \lambda Ax \text{ for all } x \geq 0$$

and hence

$$c \geq \lambda A$$

implying $\lambda \in K^*$. Insert in (6.1) the values $t = 1, x = 0$ to obtain

$$z_0 \leq \lambda b.$$

With weak duality (cf. Thm. 6.3) we get $z_0 = \lambda b$ and using Corollary 6.4 we get that $x^* \in K$ and $\lambda \in K^*$ are optimal. □

Remark 6.6. The proof does not use $\text{rank}(A) = m$.

With Theorems (6.3) and (6.5) we obtain the following existence- and optimality criteria.

Theorem 6.7 (Existence). The following statements are equivalent.

1. (P) and (D) admit feasible solutions.
2. (P) has a finite optimal solution.
3. (P) and (D) have finite optimal solutions with equal optimal value.

4. (D) admits a finite optimal solution.

5. (P) admits feasible solutions and the objective is bounded from below or (D) admits feasible solutions and the objective is bounded from above.

Proof. (1) $\Rightarrow$ (2) : With Thm. (6.3) we get that the objective of (P) is bounded from below. Hence, (since (P) admits feasible solutions by assumption) with the Thm. of Weierstrass we get the existence of an optimal finite solution.
(2) $\Rightarrow$ (3) : The is the statement of Thm. (6.5).
(3) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Rightarrow$ (1) : Trivial. □

Theorem 6.8 (Complementarity). Let $x \in K$ and $\lambda \in K^*$ be feasible for (P) and (D). Then, the following statements are equivalent.

1. x and λ are optimal.

2. $(\lambda A - c)x = 0$

3. For all $i = 1, \ldots, n$, we get

$$
\begin{aligned}
x_i > 0 &\Rightarrow \lambda a^i = c_i \\
\lambda a^i < c_i &\Rightarrow x_i = 0.
\end{aligned}
$$

Proof: Exercise (trivial).

6.3 The Dual Simplex-Method

Consider the primal-dual programms

$$
\begin{aligned}
&\min\{cx | Ax = b, x \geq 0\} && \text{(P)} \\
&\max\{\lambda b | \lambda A \leq c\}. && \text{(D)}
\end{aligned}
$$

We assume $\mathrm{rank}(A) = m$. The basic solution w.r.t. a basis B of (P) is given as

$$x_B = A_B^{-1} b, x_N = 0.$$

The row vector

$$\lambda := c_B A_B^{-1} \in \mathbb{R}^m \tag{6.2}$$

is called <u>Simplex-</u> or <u>Lagrange-Multiplier</u>. Hence, the reduced cost read as

$$r_N = c_B A_B^{-1} A_N - c_N = \lambda A_N - c_N.$$

We obtain:

Theorem 6.9.
1. $\lambda = c_B A_B^{-1}$ is feasible for (D) $\Leftrightarrow r_N \le 0$.
2. Let the basic solution x be optimal and non-degenerate, i.e.,

$$x_B = A_B^{-1} b > 0, x_N = 0.$$

Then, $\lambda = c_B A_B^{-1}$ is optimal for (D) and we get

$$\lambda b = cx.$$

Proof. For (1): With $\lambda = c_B A_B^{-1}$ we get

$$\begin{aligned} \lambda A \le c &\Leftrightarrow (\lambda A_B, \lambda A_N) \le (c_B, c_N) \\ &\Leftrightarrow \lambda A_N \le c_N \\ &\Leftrightarrow r_N = \lambda A_N - c_N \le 0. \end{aligned}$$

For (2): Is x optimal and non-degenerate, then $r_N \le 0$ (see Corollary 5.4). Thus, $\lambda = c_B A_B^{-1}$ is feasible for (D) (using (1)) and we have

$$\lambda b = c_B A_B^{-1} b = c_B x_B = cx.$$

Because of Corollary (6.4) we get that $\lambda = c_B A_B^{-1}$ is optimal for D. □

The connection of primal optimality and dual feasibility is visible in the following definition.

Definition 6.10. A basic-solution x with

$$x_B = A_B^{-1} b, x_N = 0$$

is called dual feasible for (P), if and only if

$$\lambda = c_B A_B^{-1}$$

is feasible for (D), i.e., if $r_N \le 0$.

Let us now explain the idea of the dual simplex-method. Suppose we have a basis B and a basic solution x that is

- primal infeasible, i.e., $x_B = A_B^{-1} b \ge 0$ is not valid
- dual feasible, i.e., $r_N \le 0$.

We consider the primal tableau (P) but solve it from the dual point of view:

1. maintain $r_N \le 0$
2. until we have feasibility of x, i.e. $x_B = A_B^{-1} b \ge 0$

3. The objective of (P) grows.

Assuming $A_B = I$ the tableau for (P) is given via (5.13)

$$\begin{array}{r|c|c|} & \multicolumn{2}{c}{x_j, j \in N} \\ \hline z(x) & z_0 & r_N \\ \hline x_i, i \in B & b & A_N \\ \hline \end{array} \tag{6.3}$$

Figure 6.1: Reduced Simplex-Tableau.

with tableau-matrix:

$$T = (t_{ij}) = \begin{array}{|c|c|} \hline z_0 & r_N \\ \hline b & A_N \\ \hline \end{array} \quad \begin{array}{c} 0 \leq i \leq m \\ 0 \leq j \leq n-m \end{array} \tag{6.4}$$

The tableau (6.4) is

1. primal feasible, if $b \geq 0$,
2. dual feasible, if $r_N \leq 0$,
3. optimal, if $b \geq 0, r_N \leq 0$.

Let us write down the dual Simplex-Algorithms formally (see Fig.. 6.2).
We give an application. The primal problem is given as:

$$\min\{cx | Ax \geq b, x \geq 0\}, c \geq 0 \tag{P}$$

$Ax \geq b$ is equivalent to

$$Ax - y = b, y \geq 0.$$

Use Phase I of the simplex in order to compute a feasible start solution. Because of $c \geq 0$, we can solve problem (P) easily with the dual simplex. Consider

$$-Ax + y = -b, y \geq 0, x \geq 0.$$

We obtain the start-tableau

$$\begin{aligned} B &= (n+1, \ldots, n+m), N = (1, \ldots, n) \\ r_N &= -c \leq 0. \end{aligned}$$

Hence $x = 0, y = -b$ is dual feasible and we get the tableau:

1. **Start**: Let x be a dual feasible basis B.Compute $T = (t_{ij})$, as in (6.4). Denote by $B(i)$ the basis index for $1 \leq i \leq m$ that corresponds to the i-th row of T. Analogously we denote by $N(j)$ the index $1 \leq j \leq n-m$ that belongs to the j-th column of T.

2. If $t_{i0} \geq 0 (1 \leq i \leq m)$, then the current solution is optimal.
   Set
   $$\begin{aligned} x_{B(i)} &:= t_{i0}, 1 \leq i \leq m \\ x_{N(j)} &:= 0, 1 \leq j \leq n-m \\ z &:= t_{00} \end{aligned}$$
   Otherwise go to 3.

3. **Compute exchange row**: Choose $p \in [m]$ (with $B(p)$ smallest) such that
   $$t_{p0} = \min_{1 \leq i \leq m} t_{i0} \quad (t_{i0} = b_i).$$

4. If $t_{pj} \geq 0, 1 \leq j \leq n-m$, then (D) has no finite solution. **Stop.**
   If there is some $t_{pj} < 0$,go to Step 5.

5. **Compute exchange column**: Choose $s \in [n-m]$ (with $N(s)$ smallest) such that
   $$\frac{t_{0s}}{t_{ps}} = \min \left\{ \frac{t_{0j}}{t_{pj}} \middle| t_{pj} < 0, j = 1, \ldots, n-m \right\} \quad (t_{0j} = r_{N(j)} \leq 0).$$
   Go to Step 6.

6. Exchange the s-th element of N with the p-th element of B:
   $B \leftarrow \big(B(1), \ldots, B(p-1), N(s), B(p+1) \ldots, B(m)\big)$
   $N \leftarrow \big(N(1), \ldots, N(s-1), B(p), N(s+1) \ldots, N(n-m)\big)$
   Execute in T the pivot step with pivot element $t_{ps} < 0$ as in Fig. 5.2, point (6):
   - Pivot elemet: $t'_{ps} := 1/t_{ps}$
   - Pivot column: $t'_{is} := -\frac{t_{is}}{t_{ps}}, i = 0, 1, \ldots, m, i \neq p$
   - Pivot row: $t'_{pj} := \frac{t_{pj}}{t_{ps}}, j = 0, 1, \ldots, n-m, j \neq s$
   - other elements: $t'_{ij} := t_{ij} - t_{is}\frac{t_{pj}}{t_{ps}}, i \neq p, j \neq s$
   - Set $t_{ij} := t'_{ij}$ and go to Step 2.

Figure 6.2: Dual Simplex-Method

We solve the example from the last section:

$$\begin{aligned} \text{minimize } 3x_1 + 4x_2 + 5x_3 & \\ x_1 + 2x_2 + 3x_3 &\geq 5 \\ 2x_1 + 2x_2 + x_3 &\geq 6 \\ x_i &\geq 0, i = 1, 2, 3 \end{aligned}$$

		x
z	z_0	$r_N = -c \leq 0$
y	$-b$	$-A$

(6.5)

The start-tableau has the form:

		x_1	x_2	x_3
z	0	-3	-4	-5
y_1	-5	-1	-2	-3
y_2	-6	(-2)	-2	-1

We get

$$\min\{-5, -6\} = -6 \Rightarrow p = 2$$
$$\min\{3/2, 4/2, 5/1\} = 3/2 \Rightarrow s = 1.$$

Exchange y_2 and x_1:

		y_2	x_2	x_3
z	9	$-3/2$	-1	$-7/2$
y_1	-2	$-1/2$	(-1)	$-5/2$
x_1	3	$-1/2$	1	$1/2$

Exchange y_1 and x_2:

		y_2	y_1	x_3
z	11	-1	-1	-1
x_2	2	$1/2$	-1	$5/2$
x_1	1	-1	1	-2

The tableau is optimal; the optimal solution is

$$x_1 = 1, x_2 = 2, x_3 = 0.$$

The optimal dual solution is

$$\lambda = (1, 1).$$

6.4 Sensitivity and Shadow Prices

Suppose the LP

$$\min\{cx | Ax = b, x \geq 0\} \tag{6.6}$$

has an optimal basis B with non-degenerate basic solution

$$x_B = A_B^{-1} b > 0$$

An optimal dual solution is

$$\lambda = c_B A_B^{-1}.$$

For small changes

$$b \to b + \Delta b$$

of the right hand side of (6.6) the basis B remains feasible, as x_B is non-degenerate. Moreover, for B, the reduced costs

$$r_N = c_B A_B^{-1} A_N - c_N \leq 0$$

do not depend on b and therefore x remains optimal as long as it stays feasible.
For $b + \Delta b$ we have the optimal solution

$$x = A_B^{-1}(b + \Delta b) = x_B + \Delta x_B, \text{ with } \Delta x_B = A_B^{-1} \Delta b.$$

The objective value z changes to

$$\begin{aligned} \Delta z &= z(x) - z(x_B) \\ &= c_B \Delta x_B = \lambda \Delta b. \end{aligned} \tag{6.7}$$

The dual λ measures the sensitivity of the optimal value wrt. to changes of the right-hand side b; in particular we get the shadow price formula:

$$\frac{\partial z}{\partial b_j} = \lambda_j, j = 1, \ldots, m \tag{6.8}$$

Hence, λ_j has the interpretation of the marginal price wrt. b_j.

Chapter 7

Nonlinear Optimization under Constraints

7.1 The Problem

We are given $f : \mathbb{R}^n \to \mathbb{R}$ and a subset $S \subset \mathbb{R}^n$. The general optimization problem has the form:

$$\min \{f(x) | x \in S\}. \tag{7.1}$$

Recall that maximization

$$\max \{f(x) | x \in S\}$$

is equivalent to minimization

$$\min \{-f(x) | x \in S\}.$$

Definition 7.1 (Local Minimum). $\bar{x} \in S$ is called <u>local minimum</u> of (7.1), if there is an open neighbourhood $U \subset \mathbb{R}^n$ of $\bar{x}$ with

$$f(\bar{x}) \leq f(x) \text{ for all } x \in S \cap U.$$

$\bar{x} \in S$ is called <u>strong local minimum</u> of (7.1), if there is an open neighbourhood $U \subset \mathbb{R}^n$ of $\bar{x}$ with

$$f(\bar{x}) < f(x) \text{ for all } x \in S \cap U, x \neq \bar{x}.$$

Definition 7.2 (Global Minimum). $\bar{x} \in S$ is a <u>global minimum</u> of (7.1), if

$$f(\bar{x}) \leq f(x) \text{ for all } x \in S.$$

$\bar{x} \in S$ is strict global minimum of (7.1), if

$$f(\bar{x}) < f(x) \text{ for all } x \in S, x \neq \bar{x}.$$

Usually S is described by functional equations and inequalities. A general form reads as

$$S := \{x \in \mathbb{R}^n | g(x) \in K\}, \tag{7.2}$$

where $g : \mathbb{R}^n \to \mathbb{R}^m$ and $K \subset \mathbb{R}^m$ is convex. The problem (7.1) then reads as

$$\min \{f(x) | g(x) \in K\}. \tag{7.3}$$

Let k be the dimension of the affine hull of $K \subset \mathbb{R}^m$, $(0 \leq k \leq m)$. W.l.o.g., we can replace K with

$$K \times \{0_{m-k}\}, K \subset \mathbb{R}^k, \overset{o}{K} \neq \emptyset.$$

Accordingly, we can replace $g = (g_1, \dots, g_m)^T$ with

$$g := (g_1, \dots, g_k)^T \text{ und } h := (g_{k+1}, \dots, g_m)^T.$$

The feasible set S is represented via a system of inclusions and equalities $g(x) \in K, \overset{o}{K} \neq \emptyset$ and $h(x) = 0$:

$$S := \{x \in \mathbb{R}^n | g(x) \in K, h(x) = 0, \overset{o}{K} \neq \emptyset\}.$$

The problem (7.3) is then equivalent to

$$\min \{f(x) | g(x) \in K, h(x) = 0, \overset{o}{K} \neq \emptyset\}. \tag{7.4}$$

7.2 Formulation for the Standard Cone

If we choose in (7.4) for K the standard cone

$$K := \mathbb{R}^k_- = \{x \in \mathbb{R}^k | x_i \leq 0, i = 1, \dots, k\},$$

we get the standard problem of nonlinear optimization:

$$\min \{f(x) | g(x) \leq 0, h(x) = 0\}. \tag{7.5}$$

The inequalities $g(x) \leq 0$ need to be component-wise valid. Equivalently

$$\begin{aligned} &\min f(x) \\ &\text{s.t.:} \\ &g_i(x) \leq 0, \; i = 1, \dots, k \\ &g_j(x) = 0, \; j = k+1, \dots, m. \end{aligned} \tag{7.6}$$

For $k = 0$, i.e. $g(x) = 0$ the problem is only meaningful if $m \leq n$. Sign-constraints

$$x_i \geq 0, i = 1, \dots, r \; (r \leq n),$$

or variable bounds

$$a_i \leq x_i \leq b_i, i = 1, \dots, r, \; (r \leq n)$$

are modeled as inequalities.

We study differentiable optimization problems, thus, f and g are assumed to be continuously differentiable in an open neighbourhood of $\bar{x}$. The first derivtives $f'(x)$ or $g'(x)$ at $\bar{x}$ are given as

$$f'(\bar{x}) = \left(\frac{\partial f}{\partial x_1}(\bar{x}), \dots, \frac{\partial f}{\partial x_n}(\bar{x}) \right)$$

and the $m \times n$-matrix

$$g'(\bar{x}) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1}(\bar{x}) & \cdots & \frac{\partial g_1}{\partial x_n}(\bar{x}) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1}(\bar{x}) & \cdots & \frac{\partial g_m}{\partial x_n}(\bar{x}) \end{bmatrix}.$$

We will derive in the following necessary and sufficient optimality conditions for the standard optimization problem.

Chapter 8

Tangent Cone and Regularity

In order to set up a theory of necessary and sufficient optimality conditions we introduce the tangent cone $T(S, \bar{x})$ of a set $S \subset \mathbb{R}^n$ at $\bar{x} \in S$.

Example 8.1. Let $S = \mathbb{R}^n$ and consider

$$\min\{f(x) | x \in S\}.$$

Let $\bar{x}$ be a local minimum of f over S, see Fig. 8.1. By definition of a local minimum of f if we move away from $\bar{x}$ along a feasible direction the objective may not decrease. Suppose we speak of linear feasible directions $d \in \mathbb{R}^n$ at $\bar{x} \in S$, that is, there is $\bar{\alpha} > 0$ with $x = \bar{x} + \alpha d \in S$ for all $\alpha \leq \bar{\alpha}$.

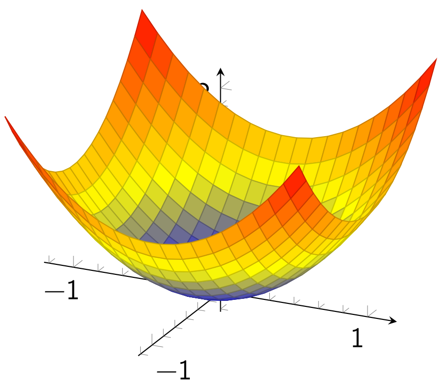
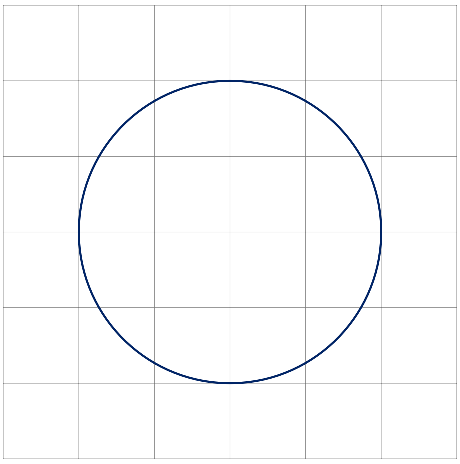


Figure 8.1: Example of a local min. at $\bar{x} = 0$.

For $\bar{x} \in \overset{o}{S}$, every $d \in \mathbb{R}^n$ is a feasible direction. For $\bar{x} \in \bar{S}$ the concept of linearly feasible directions is not enough.

Example 8.2. Let $S = \{x \in \mathbb{R}^2 | x_1^2 + x_2^2 = 1\}$ and consider

$$\min\{2x_1 + x_2^2 | x \in S\}.$$



Here $\overset{\circ}{S} = \emptyset$ and for <u>no</u> $\bar{x} \in S$, there is $\bar{\alpha} > 0$ with $x = \bar{x} + \alpha d \in S$ for all $\alpha \leq \bar{\alpha}$. Hence, we need a more general concept: infinitesimally feasible directions at $\bar{x}$.

8.1 Motivation of the Theory

$$\begin{aligned} \text{minimize } & f(x) \\ \text{s.t. } & h_i(x) = 0, \quad i \in \{1, \dots, k\}, \\ & g_j(x) \leq 0, \quad j \in \{k+1, \dots, m\}. \end{aligned} \tag{8.1}$$

with $x \in \mathbb{R}^n$and smooth functions f, h_i, g_j for all i, j. Notation: $h = (h_1, \dots, h_k), g = (g_{k+1}, \dots, g_m)$

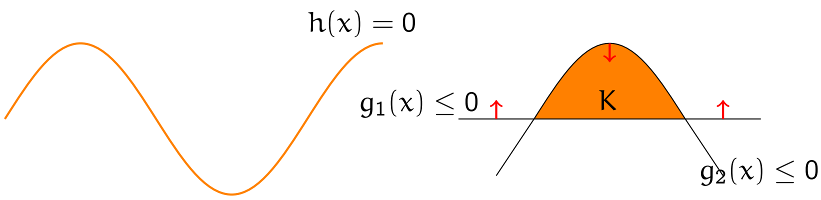


Figure 8.2: Illustration of S.

Let us now explain the idea of infinitesimally feasible directions. Consider the case $h(x) = 0$ as in Fig. 8.3. Under some assumptions (regularity as introduced later) the tangent plane $T_h(x) = \{v : \nabla h(x)^\intercal \cdot v = 0\}$ conaints the set of infinitesimally feasible directions at x (proof later).

Consider $f(x_1, x_2) = -x_1 - x_2$ with $-\nabla f(x_1, x_2) = (1, 1)$.

The intuition is as follows: Suppose there is a force tracking x with a rope along the hypersurface $h(x) = 0$ in the direction of the gradient $-\nabla f(x)$. The rope tracks the moving point continuously along $h(x) = 0$ in the direction of the force $-\nabla f(x)$. Note that there are two forces acting: the rope tracks x along the descent direction $-\nabla f(\bar{x})$

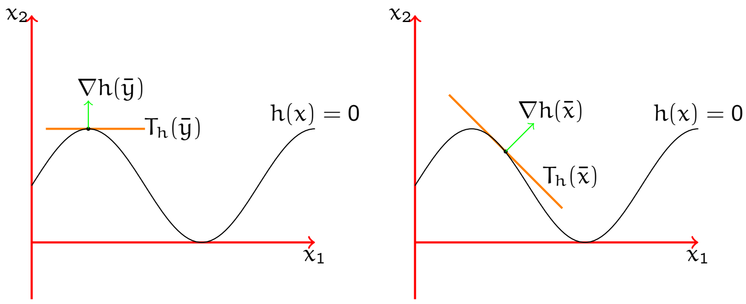


Figure 8.3: The set S is defined via $h(x) = 0$. The tangent plane is illustrated for $\bar{y}$ and $\bar{x}$.

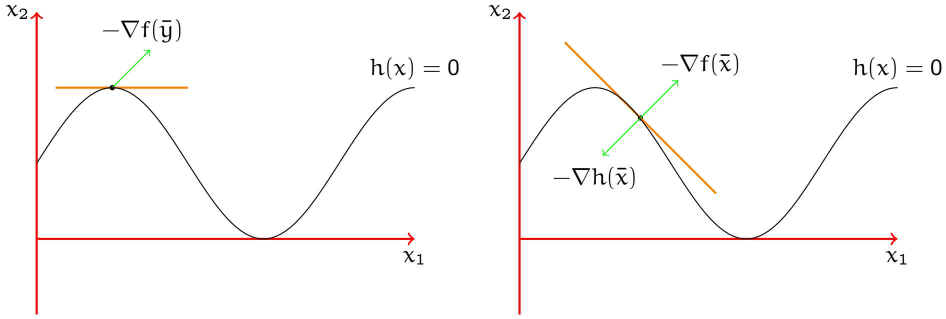


Figure 8.4: Gradients of h and f in a local minimum.

of f and the force $-\nabla h(\bar{x})$ keeps $\bar{x}$ on the hypersurface. The movement will stop if the forces $-\nabla f(x)$ and $\nabla h(x)$ act in opposite direction and an equilibrium of the two forces is reached.

Formally, for a local minimum at $\bar{x}$ for any movement from $\bar{x}$ along v for some v in the tangent plane, $\nabla f(\bar{x})^T v$ may only be nonnegative, that is, $\nabla f(\bar{x})^T v \geq 0$. As with $v \in T_h(\bar{x})$ we get $-v \in T_h(\bar{x})$ we follow $\nabla f(\bar{x})^T v = 0$. Hence, $\nabla f(\bar{x})^T$ and $\nabla h(\bar{x})^T$ are linearly dependent (without proof) and in $\bar{x}$ we have the condition

$$f'(\bar{x}) + \lambda h'(\bar{x}) = 0.$$

8.2 Tangent Cone and Variational Inequality

Definition 8.3 (Tangent Cone). The tangent cone $T(S, \bar{x})$ of $S \subset \mathbb{R}^n$ in $\bar{x}$ is defined as:

$$T(S, \bar{x}) := \left\{ v \in \mathbb{R}^n | \, \exists \, x_i \in S, t_i > 0, \text{ with } \lim_{i \to \infty} t_i = 0, v = \lim_{i \to \infty} \frac{x_i - \bar{x}}{t_i} \right\}. \tag{8.2}$$

The definition $v \in T(S, \bar{x})$ is equivalent to

$$x_i = \bar{x} + t_i v + r_i, r_i \in \mathbb{R}^n, \lim_{i \to \infty} \frac{r_i}{t_i} = 0. \tag{8.3}$$

With the Landau calculus, we have $r_i = o(t_i)$. In particular:

$$\lim_{i\to\infty} x_i = \bar{x}, \quad \lim_{i\to\infty} \frac{\|x_i - \bar{x}\|}{t_i} = \|v\|\,.$$

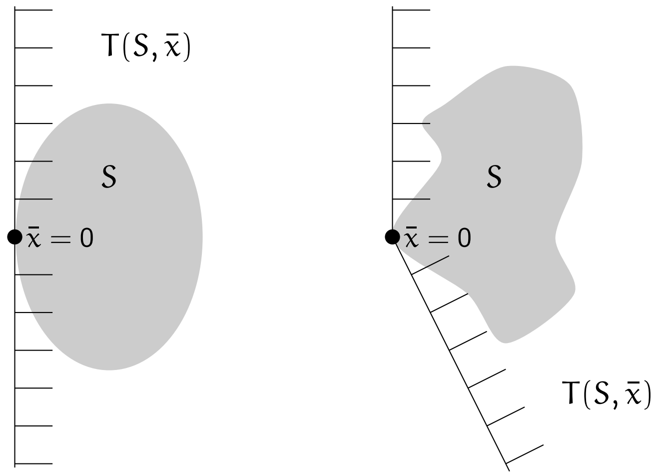


Figure 8.5: Tangent cone pointed at 0.

Notation.

Let us recap some terminology for sets in $\mathbb{R}^n$. For $K \subset \mathbb{R}^n$ we zse:

$\mathrm{aff}(K)$: the smallest affine subspace in $\mathbb{R}^n$ containing K,

$\mathrm{span}(K)$: the smallest linear subspace in $\mathbb{R}^n$ containing K,

$\mathrm{int}(K) = \overset{o}{K}$: topological interior of K wrt. $\mathbb{R}^n$,

K^i: relative topological interior of K wrt. $\mathrm{aff}(K)$,

$c\ell(K) =: \bar{K}$: topological closure of K,

$\partial K = \bar{K} \setminus \overset{o}{K}$: boundary of K.

Conic hull: Let $K \subset \mathbb{R}^n$ and $x \in K$. The cone

$$K(x) := \{\alpha(y - x) | y \in K, \alpha > 0\} = \bigcup_{\alpha > 0} \alpha(K - x)$$

is called <u>conical hull</u> of K wrt. x. Per definition $K(x)$ corresponds to the conic hull of $K - x$.

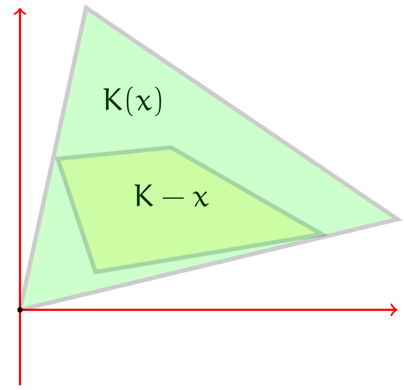


Figure 8.6: Illustration of the conic hull $K(x)$.

Lemma 8.4. For every $S \subset \mathbb{R}^n$ and $\bar{x} \in S$ the set $T(S, \bar{x})$ is a closed cone pointed at 0.

Proof. Obviously $T(S, \bar{x})$ is a cone pointed at 0. It remains to show that $T(S, \bar{x})$ is closed. Let $v \in c\ell(T(S, \bar{x}))$ and $(v_l)_{l \in \mathbb{N}}$ a sequence with

$$v_l \to v, \; v_l \in T(S, \bar{x}), \text{ for which w.l.o.g. } \|v_l - v\| \leq \frac{1}{l}.$$

We need to show that $v \in T(S, \bar{x})$. Per definition of the tangent cone for every $l \in \mathbb{N}$ there are sequences $(x_{l,k})_{k \in \mathbb{N}}$ and $(t_{l,k})_{k \in \mathbb{N}}$ with

$$x_{l,k} \in S, t_{l,k} > 0, \text{ with } \lim_{k \to \infty} t_{l,k} = 0, v_l = \lim_{k \to \infty} \frac{x_{l,k} - \bar{x}}{t_{l,k}}. \tag{8.4}$$

Hence, for all $l \in \mathbb{N}$ there is an index k_l with

$$\left\| \frac{x_{l,k_l} - \bar{x}}{t_{l,k_l}} - v_l \right\| \leq \frac{1}{l}, \; \|x_{l,k_l} - \bar{x}\| \leq \frac{1}{l} \text{ und } t_{l,k_l} \leq \frac{1}{l}.$$

For $l \to \infty$ we get $x_{l,k_l} \to \bar{x}$ and $t_l \to 0$. Moreover, wit the triangle inequality we get

$$\left\| \frac{x_{l,k_l} - \bar{x}}{t_{l,k_l}} - v \right\| \leq \left\| \frac{x_{l,k_l} - \bar{x}}{t_{l,k_l}} - v_l \right\| + \|v_l - v\| \leq \frac{2}{l}.$$

It follows that $\frac{x_{l,k_l} - \bar{x}}{t_{l,k_l}} \to v$ and hence $v \in T(S, \bar{x})$. □

For convex sets $K \subset \mathbb{R}^n$ we can easily compute the tangent cone.

Lemma 8.5. Let $K \subset \mathbb{R}^n$ be convex and let $\bar{x} \in K$. Then, $T(K, \bar{x})$ is the closure of the conical hull of K in $\bar{x}$, i.e.,

$$T(K, \bar{x}) = c\ell \left(\cup_{\alpha > 0} \alpha (K - \bar{x}) \right) =: \overline{K(\bar{x})}.$$

Proof. $T(K, \bar{x}) \supset \overline{K(\bar{x})}$: Let $x \in K$ and $\alpha > 0$. We need to show that $\alpha(x - \bar{x}) \in T(K, \bar{x})$.

As K convex and $x \in K$, we get

$$\begin{aligned} x_i &= \bar{x} + \frac{\alpha}{i}(x - \bar{x}) \in K, \text{ for } i \geq \alpha \\ \Rightarrow \frac{x_i - \bar{x}}{1/i} &= \alpha(x - \bar{x}) \\ \Rightarrow \alpha(x - \bar{x}) &= \lim_{i \to \infty} \frac{x_i - \bar{x}}{1/i} = \nu \in T(K, \bar{x}). \end{aligned}$$

As $T(K, \bar{x})$ is closed, we get $T(K, \bar{x}) \supset \overline{K(\bar{x})}$.

$T(K, \bar{x}) \subset \overline{K(\bar{x})}$: Let $\nu \in T(K, \bar{x})$ with

$$x_i = \bar{x} + t_i \nu + r_i, r_i \in \mathbb{R}^n, \lim_{i \to \infty} \frac{r_i}{t_i} = 0.$$

Then,

$$\underbrace{\frac{x_i - \bar{x}}{t_i}}_{\in K(\bar{x})} = \nu + \frac{r_i}{t_i} \Rightarrow \nu = \lim_{i \to \infty} \frac{x_i - \bar{x}}{t_i} \in \overline{K(\bar{x})}.$$

□

We derive a fundamental necessary optimality criterion.

Theorem 8.6 (Variational Inequality). Let $\bar{x}$ be a local minimum of

$$\min \{f(x) | x \in S\}$$

Then,

$$f'(\bar{x})\nu \geq 0 \text{ for all } \nu \in T(S, \bar{x}).$$

Proof. Let $\nu \in T(S, \bar{x})$. With $x_i = \bar{x} + \nu t_i + r_i$ we get

$$\begin{aligned} f(x_i) &= f(\bar{x}) + f'(\bar{x})(x_i - \bar{x}) + o(\|x_i - \bar{x}\|) \\ &= f(\bar{x}) + f'(\bar{x})\nu t_i + o(t_i). \end{aligned}$$

We obtain

$$\lim_{i \to \infty} \frac{f(x_i) - f(\bar{x})}{t_i} = \lim_{i \to \infty} \left(f'(\bar{x})\nu + \frac{o(t_i)}{t_i} \right) = f'(\bar{x})\nu.$$

Since $\bar{x}$ is a local minimum, we get for i large enough

$$f(x_i) \geq f(\bar{x}).$$

Thus,

$$0 \leq \lim_{i \to \infty} \frac{f(x_i) - f(\bar{x})}{t_i} = f'(\bar{x})\nu.$$

□

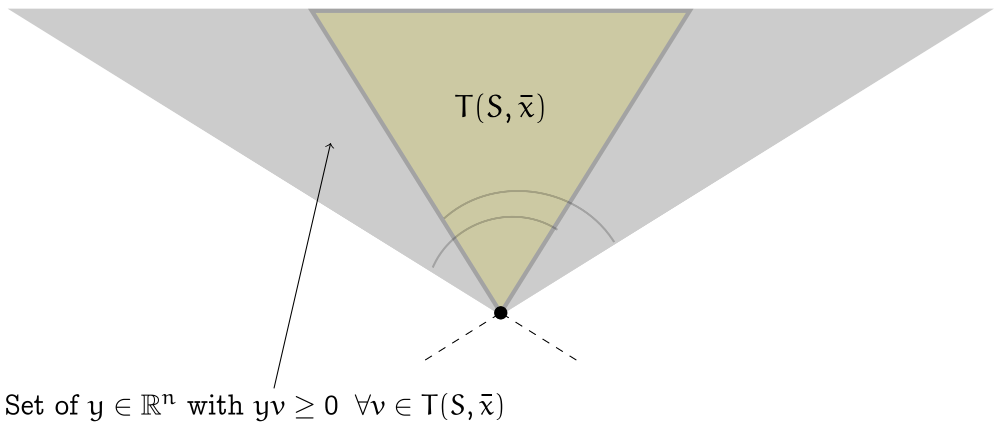


Figure 8.7: Set of vectors y for which $yv \geq 0$ for all $v \in T(S,\bar{x})$.

8.3 Linearized Cone

We consider the case that S is given as

$$S = \{x \in \mathbb{R}^n | g(x) \in K\},$$

where $g : \mathbb{R}^n \to \mathbb{R}^m$ is C^1 and $K \subset \mathbb{R}^m$ is convex. As the tangent cone is hard to compute, we introduce the linearized cone of S in $\bar{x}$.
As motivation for that cone, let us linearize the inclusion $g(x) \in K$ in $\bar{x}$, i.e., for $v \in \mathbb{R}^n$ we consider

$$g(\bar{x}) + g'(\bar{x})v \in K,$$

hence

$$g'(\bar{x})v \in K - g(\bar{x}) \subset \cup_{\alpha>0}\alpha(K - g(\bar{x})) =: K(g(\bar{x})).$$

This leads directly to the definition of the linearized cone of S in $\bar{x}$:

$$L(S,\bar{x}) = \{v \in \mathbb{R}^n | g'(\bar{x})v \in K(g(\bar{x}))\}. \tag{8.5}$$

Lemma 8.7. $L(S,\bar{x})$ is a convex cone pointed at 0.

Proof. We have $0 \in L(S,\bar{x})$, because $0 \in K(g(\bar{x}))$, which in turn follows from $g(\bar{x}) \in K$. The cone property $v \in L(S,\bar{x}) \Rightarrow \alpha v \in L(S,\bar{x})$ for all $\alpha \geq 0$ is also satisfied. We need to show convexity. Let $u, v \in L(S,\bar{x})$ and $\lambda \in (0,1)$. For $w = \lambda\, u + (1-\lambda)v$, we get

$$g'(\bar{x})w = \lambda \underbrace{g'(\bar{x})u}_{\in K(g(\bar{x}))} + (1-\lambda) \underbrace{g'(\bar{x})v}_{\in K(g(\bar{x}))} \in K(g(\bar{x}))$$

where the last inclusion follows by convexity of $K(g(\bar{x}))$ (see exercise). □

Exercise 8.8. Let $K \subset \mathbb{R}^m$ convex and $g : \mathbb{R}^n \to \mathbb{R}^m$ a C^1 mapping. Let $x \in \mathbb{R}^n$ with $g(x) \in K$. Show that the conical hull of K w.r.t. $g(x)$, that is, $K(g(x)) = \cup_{\alpha>0}\alpha(K - g(x))$, is convex.

Note that $L(S, \bar{x})$ depends not only on the set S but also on the chosen mapping g in order to represent S; see the following example.

Example 8.9. Let $S := \{0\} \subset \mathbb{R}, \bar{x} = 0$. With $K = \{0\} \subset \mathbb{R}, g(x) = x^2$ we get that S can be represented as $S = \{x \in \mathbb{R} | g(x) \in K\}$. With $g'(\bar{x}) = 0$ we have

$$L(S, \bar{x}) = \{v \in \mathbb{R} | g'(\bar{x})v = 0\} = \mathbb{R}.$$

If we choose $g(x) = x$, we still get $S = \{x \in \mathbb{R} | g(x) \in K\}$. But with $g'(\bar{x}) = 1$ we get $L(S, \bar{x}) = \{0\}$.

Another special case is $K = \{0\}$. Here

$$S = \{x \in \mathbb{R}^n | g(x) = 0\}$$

is an equation-defined mannifold and S in $\bar{x}$ is the linear subspace

$$L(S, \bar{x}) = \{v \in \mathbb{R}^n | g'(\bar{x})v = 0\}.$$

Another important special case is the standard cone $K = \mathbb{R}^k_- \times \{0_{m-k}\}$. Here, S is given as

$$S = \left\{ x \in \mathbb{R}^n \,\middle|\, \begin{array}{l} g_i(x) \leq 0, i = 1, \dots, k, \\ g_i(x) = 0, i = k+1, \dots, m \end{array} \right\}.$$

Let

$$I(\bar{x}) := \{i \in \{1, \dots, k\} | g_i(\bar{x}) = 0\}$$

be the set for which the inequalities $g_i(\bar{x}) \leq 0$ are active. We denote this set as the set of active indices.

We get that $K(g(\bar{x}))$ has the following form:

$$K(g(\bar{x})) = \left\{ y \in \mathbb{R}^m \,\middle|\, \begin{array}{l} y_i \leq 0, i \in I(\bar{x}), \\ y_i = 0, i = k+1, \dots, m \end{array} \right\}$$

and hence the linearized cone can be computed as

$$L(S, \bar{x}) = \left\{ v \in \mathbb{R}^n \,\middle|\, \begin{array}{l} g_i'(\bar{x})v \leq 0, i \in I(\bar{x}), \\ g_i'(\bar{x})v = 0, i = k+1, \dots, m \end{array} \right\}. \tag{8.6}$$

We obtain the following relationship between $T(S, \bar{x})$ and $L(S, \bar{x})$.

Lemma 8.10. If $K(g(\bar{x}))$ is closed, then

$$T(S, \bar{x}) \subset L(S, \bar{x}).$$

Proof. Let $v \in T(S, \bar{x})$ with

$$v = \lim_{i \to \infty} \frac{(x_i - \bar{x})}{t_i}, \ x_i \in S, \ t_i > 0.$$

As in the proof of Thm. 8.6 we get

$$g'(\bar{x})v = \lim_{i \to \infty} \frac{g(x_i) - g(\bar{x})}{t_i}.$$

With $x_i \in S \Rightarrow g(x_i) \in K$ and $t_i > 0$ using the definition of $K(g(\bar{x}))$ we get $\frac{g(x_i) - g(\bar{x})}{t_i} \in K(g(\bar{x}))$. Closedness of $K(g(\bar{x}))$ implies

$$g'(\bar{x})v \in K(g(\bar{x})), \text{ i.e. } v \in L(S, \bar{x}).$$

□

The reverse inclusion $L(S, \bar{x}) \subset T(S, \bar{x})$ is not valid in general as shown by the following example.

Example 8.11. Let $S := \{0\} \subset \mathbb{R}, \bar{x} = 0, T(S, \bar{x}) = \{0\}$.

1. With $K = \{0\} \subset \mathbb{R}, g(x) = x^2$ the set S is given as $S = \{x \in \mathbb{R} | g(x) \in K\}$. With $g'(\bar{x}) = 0$ we get
$$L(S, \bar{x}) = \{v \in \mathbb{R} | g'(\bar{x})v = 0\} = \mathbb{R}$$
and therefore $T(S, \bar{x}) \subsetneq L(S, \bar{x})$.
2. With $K = \{0\} \subset \mathbb{R}, g(x) = x$ we can represent S as $S = \{x \in \mathbb{R} | g(x) \in K\}$. With $g'(\bar{x}) = 1$ we get $L(S, \bar{x}) = \{0\}$ and therefore $T(S, \bar{x}) = L(S, \bar{x})$.

8.4 Regularity Conditions

For obtaining $L(S, \bar{x}) \subset T(S, \bar{x})$ one needs to impose additional conditions on g and K. Such conditions are known as constraint qualifications or regularity conditions. For a motivation, consider the case

$$S = \{x \in \mathbb{R}^n | g(x) = 0\}.$$

Definition 8.12. $\bar{x} \in S$ is called regular, if

$$\text{Im } g'(\bar{x}) = \mathbb{R}^m \tag{8.7}$$

where $\text{Im } g'(\bar{x})$ is the image of the linear mapping $x \mapsto g'(\bar{x})x$ for $x \in \mathbb{R}^n$.

$\bar{x}$ is regular iff the gradients $g_i'(\bar{x})^T, i = 1, \dots, m$ are linearly independent, or, equivalently

$$g'(\bar{x})g'(\bar{x})^T \text{ is non-singular.}$$

For later, we recap that exactly one of the following statements is true:

1. $\bar{x}$ is regular.
2. there is $\lambda \in \mathbb{R}^m, \lambda \neq 0$ with $\lambda g'(\bar{x}) = 0$.

Theorem 8.13. Let $S = \{x \in \mathbb{R}^n | g(x) = 0\}$ and let $\bar{x} \in S$ be regular. Then,

1. For $v \in \mathbb{R}^n$ with $g'(\bar{x})v = 0$ there is $\epsilon > 0$ and a curve $x : [-\epsilon, \epsilon] \to S$ with

$$x(0) = \bar{x}, \ \dot{x}(0) = \lim_{t \to 0} \frac{x(t) - \bar{x}}{t} = v.$$

2. We have $T(S, \bar{x}) = L(S, \bar{x}) = \{v \in \mathbb{R}^n | g'(\bar{x})v = 0\}$.

Proof. For (1): We define $F : \mathbb{R}^{m+1} \to \mathbb{R}^m$ via

$$F(y, t) := g(\bar{x} + tv + g'(\bar{x})^T y), \ y \in \mathbb{R}^m, t \in \mathbb{R}.$$

We have $F(0_m, 0) = g(\bar{x}) = 0$ and the partial derivatives of F wrt. y read as

$$\frac{\partial F}{\partial y}(0_m, 0) = g'(\bar{x})g'(\bar{x})^T.$$

This matrix is non-singular as $\bar{x}$ is regular. With the Implicit Function Theorem, there is $\epsilon > 0$ and a function $y : [-\epsilon, \epsilon] \to \mathbb{R}^m$ that is continuously differentiable in $t = 0$ with $y(0) = 0$ and

$$F(y(t), t) = 0 \text{ for } t \in [-\epsilon, \epsilon].$$

We obtain $\dot{y}(0) := \lim_{t \to 0} \frac{y(t) - y(0)}{t}$ and with the assumption $g'(\bar{x})v = 0$ we get

$$0 = \frac{dF(y(t), t)}{dt}\Big|_{t=0} = F_y(y(0), 0)\dot{y}(0) + F_t(y(0), 0) = g'(\bar{x})g'(\bar{x})^T \dot{y}(0),$$

where we used $F_t(y(0), 0) = g'(\bar{x})v = 0$. Hence, $\dot{y}(0) = 0$. By setting

$$x(t) = \bar{x} + vt + g'(\bar{x})^T y(t), \ t \in [-\epsilon, \epsilon],$$

we obtain a cruve $x(t)$ with the desired properties: $g(x(t)) = 0$, i.e. $x(t) \in S, x(0) = \bar{x}$ and $\dot{x}(0) = v$, because $\dot{y}(0) = 0$.
For (2): The claim follows from (1) and Lemma 8.10. □

We consider now the general case.

$$S = \{x \in \mathbb{R}^n | g(x) \in K\}, \ K \text{ convex.} \tag{8.8}$$

Definition 8.14. Let S as in (8.8). $\bar{x} \in S$ is called regular, if

$$\operatorname{Im} g'(\bar{x}) - K(g(\bar{x})) = \mathbb{R}^m. \tag{8.9}$$

This condition generalizes (8.7) and means geometrically that the linear subspace $\operatorname{Im} g'(\bar{x})$ and the cone $K(g(\bar{x}))$ are transversal to each other.
In the following, we derive easier but equivalent conditions.

Lemma 8.15. Let K be convex. The following conditions are equivalent:

1. $\operatorname{Im} g'(\bar{x}) - K(g(\bar{x})) = \mathbb{R}^m$,
2. $0 \in \operatorname{int}\left(\operatorname{Im} g'(\bar{x}) - K(g(\bar{x}))\right)$,
3. $0 \in \operatorname{int}\left(\operatorname{Im} g'(\bar{x}) + g(\bar{x}) - K\right)$,

Proof. The proof is done in the following ordfer: 1. $\Rightarrow$ 3. $\Rightarrow$ 2. $\Rightarrow$ 1.
1. $\Rightarrow$ 3. : The set $\operatorname{Im} g'(\bar{x}) + g(\bar{x}) - K$ is convex and contains 0 because $g(\bar{x}) \in K$. Suppose $0 \notin \operatorname{int}\left(\operatorname{Im} g'(\bar{x}) + g(\bar{x}) - K\right)$. With the separating hyperplane theorem (Theorem 2.8) there is $\lambda \in \mathbb{R}^m, \lambda \neq 0$ with

$$\lambda\left(g'(\bar{x})v + g(\bar{x}) - y\right) \geq 0 \text{ for all } v \in \mathbb{R}^n, y \in K.$$

With the definition

$$K(g(\bar{x})) = \cup_{\alpha>0}\alpha\left(K - g(\bar{x})\right)$$

we get

$$\lambda\left(g'(\bar{x})v - y\right) \geq 0 \text{ for all } v \in \mathbb{R}^n, y \in K(g(\bar{x})). \tag{8.10}$$

The assumption $\operatorname{Im} g'(\bar{x}) - K(g(\bar{x})) = \mathbb{R}^m$ yields $\lambda = 0$ in contradiction to the choice of λ.
3. $\Rightarrow$ 2. : Follows from $K - g(\bar{x}) \subset K(g(\bar{x}))$.
2. $\Rightarrow$ 1. : Let $B_\epsilon(0) := \{y \in \mathbb{R}^m | \, \|y\| \leq \epsilon\}$. Per assumption there is $\epsilon > 0$ with

$$B_\epsilon(0) \subset \operatorname{Im} g'(\bar{x}) - K(g(\bar{x})).$$

Since $\operatorname{Im} g'(\bar{x}) - K(g(\bar{x}))$ is a cone, we get

$$\mathbb{R}^m = \cup_{\alpha\geq 0}\alpha B_\epsilon(0) \subset \operatorname{Im} g'(\bar{x}) - K(g(\bar{x})),$$

implying 1. □

These conditions are due to S. Robinson (1976).
Using the proof of 1. $\Rightarrow$ 3. we get the following.

Corollary 8.16. $\bar{x} \in S$ is not regular if and only if there is $\lambda \in \mathbb{R}^m, \lambda \neq 0$ with

$$\lambda g'(\bar{x}) = 0, \ \lambda(-y) \geq 0 \text{ for all } y \in K(g(\bar{x})).$$

Exercise 8.17. Prove the statement of Corollary 8.16.

We specalize the conditions of Lemma 8.15 to the following case

$$S = \{x \in \mathbb{R}^n | g(x) \in K, h(x) = 0\}, \ \overset{o}{K} \neq \emptyset, K \text{ konvex}, \tag{8.11}$$

with $g = (g_1, \dots, g_k)^T, h = (g_{k+1}, \dots, g_m)^T$. Then, 2. from 8.15 is equivalent to

$$\operatorname{Im} h'(\bar{x}) = \mathbb{R}^{m-k} \text{ and there is } v \in \mathbb{R}^n \text{ with } h'(\bar{x})v = 0, \ g'(\bar{x})v \in \operatorname{int}(K(g(\bar{x}))). \tag{8.12}$$

and 3. is equivalent to

$$\operatorname{Im} h'(\bar{x}) = \mathbb{R}^{m-k} \text{ and there is } v \in \mathbb{R}^n \text{ with } h'(\bar{x})v = 0, \ g(\bar{x}) + g'(\bar{x})v \in \overset{o}{K}. \tag{8.13}$$

The conditions (8.12) and (8.13) are called local Slater-Conditions. For $K = \mathbb{R}^k_-$ we get from (8.12) the Mangasarian-Fromowitz-Conditions (cf. Mangasarian 1969).

Definition 8.18 (Mangasarian-Fromowitz). The gradients $g'_{k+1}(\bar{x})^T, \dots, g'_m(\bar{x})^T$ are linearly independent and there is $v \in \mathbb{R}^n$ with

$$\begin{aligned} &g'_i(\bar{x})v < 0, \ i \in I(\bar{x}) \text{ (set of active indices)} \\ &g'_i(\bar{x})v = 0, \ i = k+1, \dots, m. \end{aligned}$$

We now derive the central statement $L(S, \bar{x}) \subset T(S, \bar{x})$ under the aforementioned regularity conditions. We will do this for the case S being in the form (8.11).

Theorem 8.19. Let $\bar{x}$ be a regular point in

$$S = \{x \in \mathbb{R}^n | g(x) \in K, h(x) = 0\}, \ \overset{o}{K} \neq \emptyset, K \text{ convex}.$$

Then:

1. For $v \in \mathbb{R}^n$ with $h'(\bar{x})v = 0$, and $g(\bar{x}) + g'(\bar{x})v \in \overset{o}{K}$ there is $\epsilon > 0$ and a curve

   $$x : [0, \epsilon] \longrightarrow S$$

   with
   $$x(0) = \bar{x}, \ \dot{x}(0) = \lim_{t \downarrow 0} \frac{x(t) - \bar{x}}{t} = v.$$

2. We have $L(S, \bar{x}) \subset T(S, \bar{x})$.

Proof. For 1.: In case $k = m$ we set $x(t) = \bar{x} + tv$ for $t \in [0, \epsilon]$ and $\epsilon > 0$ small enough. If $k < m$ then using Thm. 8.13,(1.), there is $\delta > 0$ and a curve $x : [0, \delta] \to \mathbb{R}^n$ with

$$h(x(t)) = 0 \text{ for } t \in [-\delta, \delta],\ x(0) = \bar{x},\ \dot{x}(0) = v.$$

Using

$$g(\bar{x}) + g'(\bar{x})v \in \overset{o}{K},\ \lim_{t \to 0} \frac{g(x(t)) - g(\bar{x})}{t} = g'(\bar{x})v,$$

there is $\epsilon \leq \min\{\delta, 1\}$ with

$$g(\bar{x}) + \frac{g(x(t)) - g(\bar{x})}{t} \in K \text{ für } t \in [-\epsilon, \epsilon].$$

As $g(\bar{x}) \in K$ and with convexity of K, we get

$$g(x(t)) = (1 - t)g(\bar{x}) + t\left(g(\bar{x}) + \frac{g(x(t)) - g(\bar{x})}{t}\right) \in K \text{ for } 0 \leq t \leq \epsilon.$$

Thus, $x(t) \in S$ for all $0 \leq t \leq \epsilon$.
For 2.: We repeat the definition of $L(S, \bar{x})$:

$$L(S, \bar{x}) = \{v \in \mathbb{R}^n | g'(\bar{x})v \in K(g(\bar{x})),\ h'(\bar{x})v = 0\}.$$

Let $v \in L(S, \bar{x})$. Per definition of $K(g(\bar{x}))$ there is $r > 0$ with

$$g'(\bar{x})v \in r(K - g(\bar{x})),\ h'(\bar{x})v = 0.$$

This implies

$$g(\bar{x}) + g'(\bar{x})\frac{v}{r} \in K,\ \ h'(\bar{x})v = 0.$$

Using regularity of $\bar{x}$ and consequently (8.13), there is $v_0 \in \mathbb{R}^n$ with

$$g(\bar{x}) + g'(\bar{x})v_0 \in \overset{o}{K},\ \ h'(\bar{x})v_0 = 0.$$

We use the convex combination

$$v_\alpha := (1 - \alpha)v_0 + \alpha\frac{v}{r},\ \text{ for } 0 \leq \alpha < 1.$$

and get using the following statement

Exercise 8.20. Let K be convex and $x \in \overset{o}{K}, y \in \bar{K}$. Then,

$$(1 - \alpha)x + \alpha y \in \overset{o}{K},\ \text{für } 0 \leq \alpha < 1.$$

that

$$g(\bar{x}) + g'(\bar{x})v_\alpha \in \overset{o}{K},\ \ h'(\bar{x})v_\alpha = 0 \text{ for } 0 \leq \alpha < 1.$$

With part 1. we get

$$v_\alpha \in T(S,\bar{x}) \text{ for } 0 \leq \alpha < 1.$$

Using that $T(S,\bar{x})$ is closed, we get

$$v = \lim_{\alpha \to 1} r v_\alpha \in T(S,\bar{x}). \qquad \square$$

We give some examples.

Example 8.21. The set S is given as

$$g_1(x) = x_2 - x_1^3 \leq 0, \quad g_2(x) = x_2 \leq 0.$$

For $\bar{x} = 0$ we have $I(\bar{x}) = \{1, 2\}$.

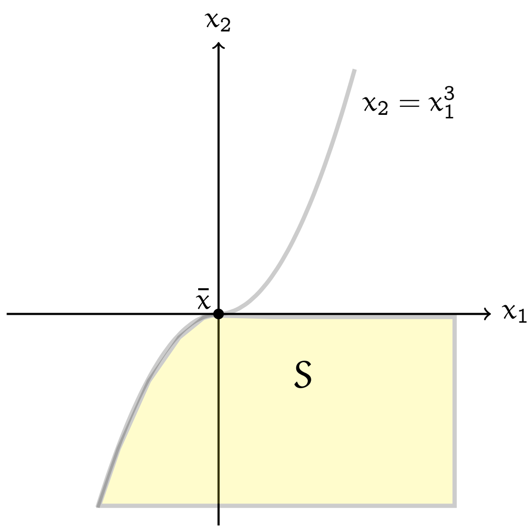


Figure 8.8: Set S.

The derivatives

$$g_1'(\bar{x}) = g_2'(\bar{x}) = (0, 1)$$

are not linearly independent. Yet, the point $\bar{x} = 0$ is regular in the sense of Definition 8.18, because for every $v \in \mathbb{R}^2$ with $v_2 < 0$ we get

$$g_i'(\bar{x})v = v_2 < 0, \ i = 1, 2.$$

We obtain

$$T(S,\bar{x}) = L(S,\bar{x}) = \{v \in \mathbb{R}^2 | v_2 \leq 0\}$$

according to Thm. 8.19, (2.).

Let us slightly change that example.

Example 8.22. S is given as

$$g_1(x) = x_2 - x_1^3 \leq 0, \quad g_2(x) = -x_2 \leq 0.$$

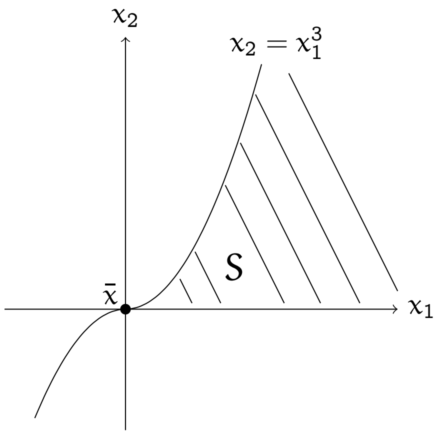


Figure 8.9: Set S.

For $\bar{x} = 0$ we get $I(\bar{x}) = \{1, 2\}$. The derivatives

$$g_1'(\bar{x}) = (0, 1) \text{ and } g_2'(\bar{x}) = (0, -1).$$

are linearly dependent. $\bar{x} = 0$ is not regular in the sense of Definition 8.18, since for no $v \in \mathbb{R}^2$, the conditions

$$g_1'(0)v = v_2 < 0, \text{ und } g_2'(0)v = -v_2 < 0$$

are satisfiable. We compute

$$T(S, \bar{x}) = \{v \in \mathbb{R}^2 | v_1 \geq 0, v_2 = 0\}$$

and with (8.6) we get

$$L(S, \bar{x}) = \{v \in \mathbb{R}^2 | v_2 = 0\}.$$

Here we have $T(S, \bar{x}) \subsetneq L(S, \bar{x})$.

Chapter 9

First Order Necessary Optimality Conditions

$$\min\{f(x)|\ g(x) \in K\} \tag{9.1}$$

We assume that $f, g \in C^1$.
Let $\bar{x}$ be a local minimum of (9.1). Then, there is a neighbourhood U of $\bar{x}$ such that the following subsets of $\mathbb{R} \times \mathbb{R}^m$ defined as

$$\{f(x) - f(\bar{x}), g(x)|x \in U\} \cap \{(r, y)|r < 0, y \in K\}$$

have an empty intersection. Thus, $(0, 0_m)$ lies on the boundary of

$$B = \{(f(x) - f(\bar{x}) + r, g(x) - y)|x \in U, r \geq 0, y \in K\} \subset \mathbb{R} \times \mathbb{R}^m. \tag{9.2}$$

Let us linearize the set B, i.e., we replace $f(x)$ and $g(x)$ with their first order Taylor expansion (without rest term) in $\bar{x}$, hence,

$$f(\bar{x}) + f'(\bar{x})v \text{ and } g(\bar{x}) + g'(\bar{x})v \text{ with } v = x - \bar{x},$$

and we obtain the convex set

$$\tilde{B} = \{(f'(\bar{x})v + r, g'(\bar{x})v + g(\bar{x}) - y|v \in \mathbb{R}^n, r \geq 0, y \in K\} \subset \mathbb{R} \times \mathbb{R}^m.$$

The conical hull of this set wrt. $(0, 0_m)$ is the following convex cone:

$$A = \{(f'(\bar{x})v + r, g'(\bar{x})v - y|v \in \mathbb{R}^n, r \geq 0, y \in K(g(\bar{x}))\} \subset \mathbb{R} \times \mathbb{R}^m. \tag{9.3}$$

Exercise 9.1. Show that indeed A is the conical hull of $\tilde{B}$ wrt. $(0, 0_m)$.

The set A can be interpreted as a convex approximation of the non-convex set B. The statement of the following theorem says that $(0, 0_m)$ lies on the boundary of A.

Theorem 9.2. Let $\bar{x}$ be a local minimum of (9.1). Then, the following statements are true.

1. Necessary Optimality Conditions of Fritz John:
   There is $(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m$, $(\lambda_0, \lambda) \neq (0, 0_m)$ with

$$\lambda_0 f'(\bar{x}) + \lambda g'(\bar{x}) = 0 \tag{9.4}$$

$$\lambda_0 \geq 0,\ \lambda(-y) \geq 0 \text{ for all } y \in K(g(\bar{x})). \tag{9.5}$$

2. Necessary Optimality Conditions of Karush-Kuhn-Tucker:
   If $\bar{x}$ is regular, we get $\lambda_0 > 0$ in (1) and w.l.o.g. $\lambda_0 = 1$ holds. Thus, there is $\lambda \in \mathbb{R}^m$ with

$$f'(\bar{x}) + \lambda g'(\bar{x}) = 0 \tag{9.6}$$

$$\lambda(-y) \geq 0 \text{ for all } y \in K(g(\bar{x})). \tag{9.7}$$

Proof. Both statements are proven together.
1. Case: $\bar{x}$ is not regular. With Corollary 8.16 there is $\lambda \in \mathbb{R}^m, \lambda \neq 0$ with

$$\lambda g'(\bar{x}) = 0,\ \lambda(-y) \geq 0 \text{ for all } y \in K(g(\bar{x})).$$

The statement of the first part of the theorem follows with $\lambda_0 = 0$.
2. Case: Let $\bar{x}$ be regular. The variational inequality of Thm. 8.6 reads as:

$$f'(\bar{x})v \geq 0 \text{ for all } v \in T(S, \bar{x}),\ \text{ where } S := \{x \mid g(x) \in K\}.$$

With the regularity of $\bar{x}$, we get with Thm. 8.19 that

$$T(S, \bar{x}) \supset L(S, \bar{x}) = \{v \in \mathbb{R}^n | g'(\bar{x})v \in K(g(\bar{x}))\},$$

and hence

$$f'(\bar{x})v \geq 0 \text{ for all } v \in \mathbb{R}^n \text{ with } g'(\bar{x})v \in K(g(\bar{x})). \tag{9.8}$$

We consider the convex cone in (9.3):

$$A = \{(f'(\bar{x})v + r, g'(\bar{x})v - y | v \in \mathbb{R}^n, r \geq 0, y \in K(g(\bar{x}))\} \subset \mathbb{R} \times \mathbb{R}^m.$$

Because of (9.8) we have that $(0, 0_m)$ lies on the boundary of A (set $v = 0, r = 0, y = 0$) and with the separating hyperplane theorem there is $(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m$, $(\lambda_0, \lambda) \neq 0$, with

$$\lambda_0(f'(\bar{x})v + r) + \lambda(g'(\bar{x})v - y) \geq 0 \text{ for all } v \in \mathbb{R}^n, r \geq 0, y \in K(g(\bar{x})). \tag{9.9}$$

For $r = 0$ and $y = 0$ we get

$$\lambda_0(f'(\bar{x})v) + \lambda g'(\bar{x})v \geq 0 \text{ for all } v \in \mathbb{R}^n$$

and therefore

$$\lambda_0(f'(\bar{x})) + \lambda g'(\bar{x}) = 0.$$

For $\nu = 0$ and $y = 0$ we get $\lambda_0 r \geq 0$ for all $r \geq 0$ and hence $\lambda_0 \geq 0$. Finally, (9.9) implies for $\nu = 0$ and $r = 0$

$$\lambda(-y) \geq 0 \text{ for all } y \in K(g(\bar{x})).$$

The case $\lambda_0 = 0$ contradicts the statement of Cor. 8.16 for regular points $\bar{x}$. Thus, $\lambda_0 > 0$ and w.l.o.g. $\lambda_0 = 1$.

□

9.1 Lagrange-Function and Multipliers

The last theorem is known as the Lagrange-Multiplier rule. The vector $\lambda_i, i = 0, \dots, m$ is called Lagrange-Multiplier. The Lagrange-Function is defined as

$$L(x, \lambda_0, \lambda) := \lambda_0 f(x) + \lambda g(x),\ (\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m$$

and for regular points

$$L(x, \lambda) := f(x) + \lambda g(x),\ \lambda \in \mathbb{R}^m.$$

The last theorem then reads as:

1. John: $L_x(\bar{x}, \lambda_0, \lambda) = 0, \lambda_0 \geq 0, \lambda(-y) \geq 0$ for all $y \in K(g(\bar{x}))$
2. KKT: $L_x(\bar{x}, \lambda) = 0, \lambda(-y) \geq 0$ for all $y \in K(g(\bar{x}))$,

where L_x denote the partial derivative of L wrt. x.

Exercise 9.3. Show that the following sets

$$\Lambda(\bar{x}) := \{\lambda \in \mathbb{R}^m | L_x(\bar{x}, \lambda) = 0, \lambda(-y) \geq 0 \text{ for all } y \in K(g(\bar{x}))\} \tag{9.10}$$

are convex and closed.
If $\bar{x}$ is regular, then the above set is the set of Lagrange-Multipliers at $\bar{x}$.

Theorem 9.4. Let $\bar{x}$ be a local minimum of (9.1). Then, the following statements are equivalent.

1. $\bar{x}$ is regular.
2. $\Lambda(\bar{x}) \neq \emptyset$ and $\Lambda(\bar{x})$ are bounded.

Proof. 1.⇒2.: $\Lambda(\bar{x}) \neq \emptyset$ follows from Theorem 9.2, (2.).
Assumption: $\Lambda(\bar{x})$ is unbounded. Then, there is a sequence $\lambda_i \in \Lambda(\bar{x}), i \in \mathbb{N}$, with $\|\lambda_i\| \to \infty$ for $i \to \infty$. Per definition of $\Lambda(\bar{x})$ we have

$$f'(\bar{x}) + \lambda_i g'(\bar{x}) = 0, \lambda_i(-y) \geq 0 \text{ for all } y \in K(g(\bar{x})).$$

With $\lambda_i \neq 0$ for i large enough, we get

$$\frac{1}{\|\lambda_i\|} f'(\bar{x}) + \frac{\lambda_i}{\|\lambda_i\|} g'(\bar{x}) = 0, \quad \frac{\lambda_i}{\|\lambda_i\|}(-y) \geq 0 \text{ for all } y \in K(g(\bar{x})). \tag{9.11}$$

As the boundary of $B_1(0)$ is compact in $\mathbb{R}^m$ we may assume that $\frac{\lambda_i}{\|\lambda_i\|} \to \lambda$ with $\|\lambda\| = 1$. Equation (9.11) yields for $i \to \infty$

$$\lambda g'(\bar{x}) = 0, \lambda(-y) \geq 0 \text{ for all } y \in K(g(\bar{x})).$$

With Corollary 8.16 we get a contradiction to the regularity at $\bar{x}$.
1.⇐2.: Let $\lambda_1 \in \Lambda(\bar{x})$.
Assumption: $\bar{x}$ is not regular.
With Corollary 8.16 there is $\lambda_2 \in \mathbb{R}^m, \lambda_2 \neq 0$ with

$$\lambda_2 g'(\bar{x}) = 0, \quad \lambda_2(-y) \geq 0 \text{ for all } y \in K(g(\bar{x})).$$

This implies $\lambda_1 + r\lambda_2 \in \Lambda(\bar{x})$ for all $r \geq 0$ in contradiction to the boundedness of $\Lambda(\bar{x})$. □

We get the following implication.

Corollary 9.5. If $\bar{x}$ is a regular local minimum of (9.1), then $\Lambda(\bar{x})$ is a nonempty, compact and convex subset of $\mathbb{R}^m$.

Now we discuss uniqueness of the Lagrange-Multipliers.

Definition 9.6. $\bar{x}$ is called normal, if

$$\operatorname{Im} g'(\bar{x}) - V = \mathbb{R}^m, \quad V := K(g(\bar{x})) \cap (-K(g(\bar{x}))). \tag{9.12}$$

Recall that $V = K(g(\bar{x})) \cap (-K(g(\bar{x})))$ is the largest linear subspace contained in $K(g(\bar{x}))$. With $K(g(\bar{x})) \supset V$ we get: $\bar{x}$ is normal $\Rightarrow$ $\bar{x}$ is regular.

Theorem 9.7. If $\bar{x}$ is a normal local minimum of (9.1), then $\Lambda(\bar{x})$ is a singleton.

Proof. We get $\Lambda(\bar{x}) \neq \emptyset$, as $\bar{x}$ is regular. For $\lambda_1, \lambda_2 \in \Lambda(\bar{x})$ we show $\lambda_1 = \lambda_2$. Per definition of $\Lambda(\bar{x})$ we have

$$f'(\bar{x}) + \lambda_i g'(\bar{x}) = 0, \quad \lambda_i(-y) \geq 0 \text{ for all } y \in K(g(\bar{x})) \quad (i = 1, 2).$$

The vector $\lambda := \lambda_1 - \lambda_2$ satisfies

$$\lambda g'(\bar{x}) = 0, \quad \lambda y = 0 \text{ for all } y \in V.$$

With condition (9.12) in Definition 9.6 (note $\operatorname{Im} g'(\bar{x}) - V = \mathbb{R}^m$) we get $\lambda = 0$ and therefore $\lambda_1 = \lambda_2$. □

9.2 Specialization for the Standard Cone

Exercise 9.8. Show that if the set K in formulation 9.1 is a convex cone, we get $K(g(\bar{x})) = K + \mathbb{R}\, g(\bar{x})$.

$$\lambda(-y) \geq 0 \text{ for all } y \in K(g(\bar{x}))$$

is in this case equivalent to

$$\lambda(-y) \geq 0 \text{ for all } y \in K,\ \lambda g(\bar{x}) = 0. \tag{9.13}$$

The equation $\lambda g(\bar{x}) = 0$ is called complementarity conditions.

Exercise 9.9. Show that for $K(g(\bar{x})) = K + \mathbb{R}\, g(\bar{x})$ we get that

$$\lambda(-y) \geq 0 \text{ for all } y \in K(g(\bar{x}))$$

is equivalent to

$$\lambda(-y) \geq 0 \text{ for all } y \in K,\ \lambda g(\bar{x}) = 0.$$

Now we consider the standard problem of nonlinear optimization.

$$\begin{aligned} \min \{f(x)|\ & g_i(x) \leq 0,\ i = 1, \dots, k, \\ & g_i(x) = 0,\ i = k+1, \dots, m\} \end{aligned} \tag{9.14}$$

This problem is obtained as special case of 9.1 by setting $K = \mathbb{R}^k_- \times \{0_{m-k}\}$. Recall:

$$\begin{aligned} I(\bar{x}) &:= \{i \in \{1, \dots, k\}|\ g_i(\bar{x}) = 0\} \\ J(\bar{x}) &:= I(\bar{x}) \cup \{k+1, \dots, m\} \end{aligned}$$

For the Lagrange-Multipliers $\lambda = (\lambda_1, \dots, \lambda_m)$ we get from (9.13)

$$\lambda_i \geq 0 \text{ for all } i \in I(\bar{x}),\ \lambda_i = 0 \text{ for all } i \notin J(\bar{x}).$$

We compute

$$K(g(\bar{x})) = \{y \in \mathbb{R}^m | y_i \leq 0,\ i \in I(\bar{x}),\ y_i = 0,\ i = k+1, \dots, m\}$$

and the linear subspace V in (9.12) is given by

$$V = \{y \in \mathbb{R}^m | y_i = 0 \text{ for all } i \in J(\bar{x})\}.$$

Corollary 9.10. $\bar{x}$ is <u>normal</u> for 9.14 if and only if the gradients

$$g_i'(\bar{x})\ i \in J(\bar{x}) \text{ are linearly independent.} \tag{9.15}$$

Theorem 9.11. Let $\bar{x}$ be a local minimum of (9.14). Then, there is $\lambda_0 \geq 0$ and $\lambda \in \mathbb{R}^m$ with $(\lambda_0, \lambda) \neq (0, 0_m)$, such that:

1. $L_x(\bar{x}, \lambda_0, \lambda) = \lambda_0 f'(\bar{x}) + \lambda g'(\bar{x}) = \lambda_0 f'(\bar{x}) + \sum_{i=1}^m \lambda_i g_i'(\bar{x}) = 0 \in \mathbb{R}^n$.
2. $\lambda_i = 0$ for $i \notin J(\bar{x})$, i.e., $i \in \{1, \dots, k\}$ with $g_i(\bar{x}) < 0$.
3. $\lambda_i \geq 0$ for $i \in I(\bar{x})$.

We have $\lambda_0 > 0$, if $\bar{x}$ regular. Then, w.l.o.g. $\lambda_0 = 1$ (divide Lagrange-Function by $\lambda_0 > 0$). If $\bar{x}$ is normal, then $\lambda \in \mathbb{R}^m$ is unique.

Every feasible $\bar{x}$ with multipliers (λ_0, λ) satisfying the conditions of Thm. 9.11are called <u>critical point</u>. It turns out that not every critical point is a local minimum of (9.14). In particular, for problems 9.11 with equations, the required conditions for the minimization

$$\min\{f(x) | g(x) = 0\}$$

and maximization variant

$$\max\{f(x) | g(x) = 0\}$$

coincide.

Example 9.12.

$$\min \{ f(x) = x_1 + x_2$$
$$g(x) = x_1^2 + x_2^2 - 2 = 0. \}$$

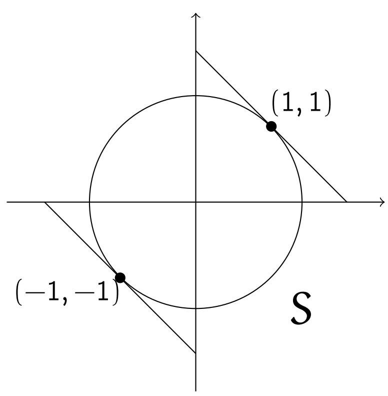


Figure 9.1: Set S and critical points.

Every point of S defined as

$$S := \{(x_1, x_2) \in \mathbb{R}^2 | x_1^2 + x_2^2 - 2 = 0\}$$

is regular. The necessary KKT-conditions are

$$f'(\bar{x}) + \lambda g'(\bar{x}) = (1 + 2\lambda\bar{x}_1, 1 + 2\lambda\bar{x}_2) = 0.$$

This implies $\lambda \neq 0$. Together with $g(\bar{x}) = 0$ we get

$$\bar{y} = (1, 1)^T, \ \lambda = -1/2$$
$$\bar{x} = (-1, -1)^T, \ \lambda = 1/2$$

Obviously $\bar{y}$ is a local maximum, while $\bar{x}$ is a local minimum. With the sufficient optimality conditions that come up in the next section, we can show this formally.

We slightly modify the example.

Example 9.13.

$$\min \{f(x) = x_1 + x_2$$
$$g(x) = x_1^2 + x_2^2 - 2 \leq 0.\}$$

Every $\bar{x} \neq 0$ of the set S defined as

$$S := \{(x_1, x_2) \in \mathbb{R}^2 | x_1^2 + x_2^2 - 2 \leq 0\}$$

is regular. Th necessary KKT-conditions read as

$$f'(\bar{x}) + \lambda g'(\bar{x}) = (1 + 2\lambda\bar{x}_1, 1 + 2\lambda\bar{x}_2) = 0, \ \lambda \geq 0 \text{ for } g(\bar{x}) = 0.$$

We get the unique solution

$$\bar{x} = (-1, -1)^T, \ \lambda = 1/2$$

Hence, the sign constraint $\lambda \geq 0$ sorts out the solution $\bar{y}$.

We give another example.

Example 9.14.

$$\min \{f(x) = -x_2$$
$$g_1(x) = x_1^2 + x_2 \leq 0$$
$$g_2(x) = -x_1^2 + x_2 \leq 0.\}$$

$\bar{x} = (0, 0)^T$ is the global minimum. We get $I(\bar{x}) = \{1, 2\}$ and

$$g_1'(0, 0) = (0, 1), g_2'(0, 0) = (0, 1)$$

are linearly dependent, thus, $\bar{x}$ is not normal. But $\bar{x}$ is regular, as it satisfies the Mangasarian-

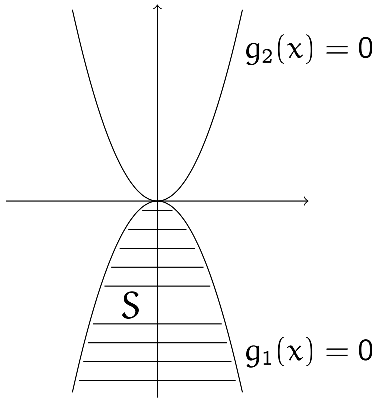


Figure 9.2: Set S and critical points.

Fromowitz-conditions for $v = (0, -1)^\top$.

$$g_1'(0,0)v = g_2'(0,0)v = (0,1) \cdot (0,-1)^\top = -1 < 0.$$

The necessary KKT-conditions read as

$$f'(\bar{x}) + \lambda g'(\bar{x}) = (0, -1 + \lambda_1 + \lambda_2) = 0,\ \lambda_1 \geq 0, \lambda_2 \geq 0.$$

The set of multipliers is given by

$$\Lambda(\bar{x}) = \{(\lambda_1, \lambda_2) \in \mathbb{R}^2_+ |\ \lambda_1 + \lambda_2 = 1\}.$$

This set is convex and compact in compliance with Corollary 9.5.

Chapter 10

Second-Order Necessary and Sufficient Optimality Conditions

$$\min \{f(x)| \; g(x) \in K\} \tag{10.1}$$

where we assume $f, g \in C^2$.
The Hesse-matrix of f in $\bar{x}$ is denoted by

$$f''(\bar{x}) = \left(\frac{\partial^2 f(\bar{x})}{\partial x_i \partial x_j} \right)_{i,j=1,\dots,n}.$$

The Hesse-matrix of the Lagrange function reads as

$$L_{xx}(\bar{x}, \lambda_0, \lambda) = \lambda_0 f''(\bar{x}) + \lambda g''(\bar{x}) = \lambda_0 f''(\bar{x}) + \sum_{i=1}^{m} \lambda_i g_i''(\bar{x}).$$

The linearized cone of S in $\bar{x}$ is per definition in (8.5) given as

$$L(S, \bar{x}) = \{v \in \mathbb{R}^n | g'(\bar{x})v \in K(g(\bar{x}))\}.$$

We now derive second order conditions using the Hesse-matrix of the Lagrange-function

$$\lambda_0 f''(\bar{x}) + \lambda g''(\bar{x})$$

wrt. the convex cone

$$\begin{aligned} C &:= \{v \in \mathbb{R}^n | f'(\bar{x})v \le 0, \; g'(\bar{x})v \in K(g(\bar{x}))\} \\ &= \{v \in \mathbb{R}^n | f'(\bar{x})v \le 0\} \cap L(S, \bar{x}). \end{aligned} \tag{10.2}$$

The set C represents the set of vectors v of the linearized cone $L(S, \bar{x})$ that constitute descent directions of f in $\bar{x}$.

Theorem 10.1. Let $\bar{x} \in S$ and $K(g(\bar{x}))$ closed.

1. Second Order Necessary Conditions:
   If $\bar{x}$ is a local minimum of (10.1), then, for every $v \in C$ there is $(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m$, $(\lambda_0, \lambda) \neq 0$ with
   (a) $\lambda_0 \geq 0$, $\lambda(-y) \geq 0$ for all $y \in K(g(\bar{x}))$.
   (b) $\lambda_0 f'(\bar{x}) + \lambda g'(\bar{x}) = 0$
   (c) $v^\top(\lambda_0 f''(\bar{x}) + \lambda g''(\bar{x}))v \geq 0$
2. Second Order Sufficient Conditions:
   Suppose for every $v \in C \setminus \{0\}$ there exists $(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m$, $(\lambda_0, \lambda) \neq 0$ with
   (a) $\lambda_0 \geq 0$, $\lambda(-y) \geq 0$ for all $y \in K(g(\bar{x}))$.
   (b) $\lambda_0 f'(\bar{x}) + \lambda g'(\bar{x}) = 0$
   (c) $v^\top(\lambda_0 f''(\bar{x}) + \lambda g''(\bar{x}))v > 0$.

   Then, there is $\epsilon > 0$ and a constant $c > 0$ with
   $$f(x) \geq f(\bar{x}) + c\,\|x - \bar{x}\|^2 \text{ for all } x \in S \text{ with } \|x - \bar{x}\| \leq \epsilon. \tag{10.3}$$
   In particular, $\bar{x}$ is a strong local minimum of (10.1).
3. If $\bar{x}$ is regular, then $\lambda_0 > 0$, i.e., w.l.o.g. $\lambda_0 = 1$ can be chosen in (1.) and (2.). If $\bar{x}$ is normal, then $\lambda_0 = 1$ and λ in (1.) and (2.) is unique and independent of $v \in C$.

Proof. For 1: The proof is similar to that for the first order conditions, see Lempio and Zowe (1981) for a complete proof.
For 2: We assume that (10.3) is false. Then, there is a sequence $\{x_i\} \subset S$ with
$$x_i \neq \bar{x}, \quad \lim_{i\to\infty} x_i = \bar{x}.$$
and
$$f(x_i) < f(\bar{x}) + \frac{1}{i}\|x_i - \bar{x}\|^2. \tag{10.4}$$
We set $v_i := x_i - \bar{x} \neq 0$. The boundary of the unit ball is compact and hence we can assume w.l.o.g. that
$$v := \lim_{i\to\infty} \frac{v_i}{\|v_i\|} \quad (\text{with } \|v\| = 1).$$
With (10.4) we get
$$f'(\bar{x})v = \lim_{i\to\infty} \frac{f(x_i) - f(\bar{x})}{\|x_i - \bar{x}\|} \leq 0,$$
and with the closedness of $K(g(\bar{x}))$ we get
$$g'(\bar{x})v = \lim_{i\to\infty} \frac{g(x_i) - g(\bar{x})}{\|x_i - \bar{x}\|} \in K(g(\bar{x})).$$

Per definition of the cone (10.2) we get $v \in C \setminus \{0\}$. For such v there is per assumption $(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m$, $(\lambda_0, \lambda) \neq 0$, satisfying $(a), (b), (c)$. The Taylorexpansion in $\bar{x}$ of second order of f and g yields together with (10.4):

$$f'(\bar{x})v_i + \frac{1}{2}v_i^T f''(\bar{x})v_i + o(\|v_i\|^2) = f(x_i) - f(\bar{x}) \leq \frac{1}{i}\|v_i\|^2 \tag{10.5}$$

$$g'(\bar{x})v_i + \frac{1}{2}v_i^T g''(\bar{x})v_i + o(\|v_i\|^2) = g(x_i) - g(\bar{x}) \in K(g(\bar{x})). \tag{10.6}$$

Multiplying (10.5) with λ_0 and (10.6) with λ, we get via addition of both equalities and considering $(a), (b)$ the inequality

$$\frac{1}{2}v_i^T(\lambda_0 f''(\bar{x}) + \lambda g''(\bar{x}))v_i + o(\|v_i\|^2) \leq \frac{\lambda_0}{i}\|v_i\|^2.$$

(We use in particular (a), i.e., $g(x_i) - g(\bar{x}) \in K(g(\bar{x}))$ and hence $\lambda(g(x_i) - g(\bar{x})) \leq 0$.) The division by $\|v_i\|^2$ yields in the limit $i \to \infty$ the inequality

$$\frac{1}{2}v^T(\lambda_0 f''(\bar{x}) + \lambda g''(\bar{x}))v \leq 0$$

in contradiction to (c).

For 3: If $\bar{x}$ is regular, we get $\lambda_0 > 0$ from Thm. 9.2(2.). If $\bar{x}$ is normal, we get the statement from Thm. 9.7. □

The cone C in (10.2) can also described without using f. Suppose that $\bar{x}$ is normal. If $\bar{x}$ is a local minimum, then there is a unique $\lambda \in \mathbb{R}^m$ with

$$f'(\bar{x}) + \lambda g'(\bar{x}) = 0, \quad \lambda(-y) \geq 0 \text{ for } y \in K(g(\bar{x})). \tag{10.7}$$

For $v \in C$ we get per definition

$$f'(\bar{x})v \leq 0, \quad \lambda(-y) \geq 0 \text{ for } y \in K(g(\bar{x})).$$

In conjunction with (10.7) we get

$$f'(\bar{x})v = -\lambda g'(\bar{x})v \leq 0, \lambda(-g'(\bar{x})v) \geq 0.$$

Hence,

$$f'(\bar{x})v = 0 \text{ and } \lambda g'(\bar{x})v = 0 \text{ for all } v \in C.$$

Thus, C has the form

$$C = \{v \in \mathbb{R}^n | \lambda g'(\bar{x})v = 0, \quad g'(\bar{x})v \in K(g(\bar{x}))\}. \tag{10.8}$$

10.1 Specialization to the Standard Cone

For $K = \{0\}$ we get

$$C = L(S, \bar{x}) = \{v \in \mathbb{R}^n | g'(\bar{x})v = 0\}.$$

For the standard problem we get using $I(\bar{x})$ and $J(\bar{x}) = I(\bar{x}) \cup \{k+1, \dots, m\}$ (see Section 9) for the multiplier $\lambda \in \mathbb{R}^m$:

$$\lambda_i \geq 0 \text{ for all } i \in I(\bar{x}),\ \lambda_i = 0 \text{ for all } i \notin J(\bar{x}).$$

Using

$$K(g(\bar{x})) = \{y \in \mathbb{R}^m | y_i \leq 0, i \in I(\bar{x}),\ y_i = 0, i = k+1, \dots, m\}$$

C has the form

$$\begin{aligned} C = \{v \in \mathbb{R}^n | g_i'(\bar{x})v \leq 0, i \in I(\bar{x}),\ \lambda_i = 0, \\ g_i'(\bar{x})v = 0, i \in I(\bar{x}),\ \lambda_i > 0, \\ g_i'(\bar{x})v = 0, i = k+1, \dots, m\}. \end{aligned} \tag{10.9}$$

Altogether we obtain the following conditions for problem (10.1).

Theorem 10.2. Let $\bar{x} \in S$ be normal, i.e., the gradients $g_i'(\bar{x})^T, i \in J(\bar{x})$ are linearly independent.

1. Second-order necessary conditions:
   If $\bar{x}$ is a local minimum (10.1), then there is a unique $\lambda_i \in \mathbb{R}, i \in J(\bar{x})$, with
   (a) $\lambda_i \geq 0$ for $i \in I(\bar{x})$,
   (b) $f'(\bar{x}) + \sum_{i \in J(\bar{x})} \lambda_i g_i'(\bar{x}) = 0$
   (c) $v^T(f''(\bar{x}) + \sum_{i \in J(\bar{x})} \lambda_i g_i''(\bar{x}))v \geq 0$ for all $v \in C$ mit C as in (10.9).
2. Second-order sufficient conditions:
   Suppose there are $\lambda_i \in \mathbb{R}, i \in J(\bar{x})$ with
   (a) $\lambda_i \geq 0$ for $i \in I(\bar{x})$,
   (b) $f'(\bar{x}) + \sum_{i \in J(\bar{x})} \lambda_i g_i'(\bar{x}) = 0$
   (c) $v^T(f''(\bar{x}) + \sum_{i \in J(\bar{x})} \lambda_i g_i''(\bar{x}))v > 0$ for all $v \in C \setminus \{0\}$ with C as in (10.9).

   Then, there is $\epsilon > 0$ and $c > 0$ with

$$f(x) \geq f(\bar{x}) + c\,\|x - \bar{x}\|^2 \text{ for all } x \in S \text{ with } \|x - \bar{x}\| \leq \epsilon. \tag{10.10}$$

   In particular, $\bar{x}$ is a strong local minimum of (10.1).

10.2 Examples

We revisit the example 9.12

$$\begin{aligned} \min \{f(x) &= x_1 + x_2 \\ g(x) &= x_1^2 + x_2^2 - 2 = 0.\} \end{aligned}$$

which has the two critical points

$$\begin{aligned} \bar{x}_1 &= (1,1)^T \text{ with } \lambda_1 = -1/2 \\ \bar{x}_2 &= (-1,-1)^T \text{ with } \lambda_2 = 1/2. \end{aligned}$$

The Hesse-matrix of the Lagrange-function in those points is

$$f''(\bar{x}) + \lambda g''(\bar{x}) = \lambda \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

For $\bar{x}_2$ and λ_2 we get that the matrix is pos. def. on $\mathbb{R}^2$, hence, the conditions of Thm. 10.2(2.) are satisfied with the subspace

$$C = \{v \in \mathbb{R}^2 | v_2 = -v_1\}.$$

Thus, $\bar{x}_2$ is a strong local minimum.

For $\bar{x}_1$ and λ_1 the matrix is negativ definite, and hence $\bar{x}_1$ is a strong local maximum.

We give another example.

$$\begin{aligned} \min \{f(x) &= x_1^2 + x_2 \\ g_1(x) &= x_1^2 + x_2^2 - 9 \le 0, \\ g_2(x) &= x_1 + x_2 - 1 \le 0.\} \end{aligned}$$

As candidate for a local minimnum consider $\bar{x} = (0,-3)^T$, for which $g_1(\bar{x}) = 0$ and $g_2(\bar{x}) < 0$. Hence, $I(\bar{x}) = \{1\}$. The point $\bar{x}$ is normal and the KKT-conditions in Thm. 9.11 has with $\lambda_2 = 0$ the solution $\lambda_1 = 1/6$. The cone C in (10.9) is the subspace

$$\begin{aligned} C &= \{v \in \mathbb{R}^2 | 2\bar{x}_1 v_1 + 2\bar{x}_2 v_2 = 0\} \\ &= \{v \in \mathbb{R}^2 | v_2 = 0\} \end{aligned}$$

The Hesse-matrix of the Lagrange-function

$$f''(\bar{x}) + \lambda_1 g_1''(\bar{x}) = \begin{pmatrix} 2(1+\lambda_1) & 0 \\ 0 & 2\lambda_1 \end{pmatrix}.$$

is positiv definite on $\mathbb{R}^2$ and in particular

$$v^\intercal(f''(\bar{x}) + \lambda_1 g_1'(\bar{x}))v = 2v_1^2(1+\lambda_1) > 0$$

for all $v \in C\setminus\{0\}$. With Thm. 10.2(2.), we get that $\bar{x} = (0,-3)^\intercal$ is a strong local minimum.

Chapter 11

Sensitivity Analysis

We consider the standard problem:

$$\min\{f(x)|\ g_i(x) \le 0,\ i = 1,\dots,k, \\ g_i(x) = 0,\ i = k+1,\dots,m\} \tag{11.1}$$

Suppose that the right-hand side (11.1) is perturbed:

$$\min\{f(x)|\ g_i(x) \le \epsilon_i,\ i = 1,\dots,k, \\ g_i(x) = \epsilon_i,\ i = k+1,\dots,m\} \tag{11.2}$$

We obtain a parameterized family of optimization problems depending on $\epsilon := (\epsilon_1,\dots,\epsilon_m) \in \mathbb{R}^m$ denoted by (11.1). For $\epsilon = 0$ the perturbed problem (11.2) becomes (11.1) which we denote as the unperturbed problem.
More generally for $\epsilon \in \mathbb{R}^p, p \ge 1$, we obtain:

$$\min\{f(x,\epsilon)|\ g_i(x,\epsilon) \le 0,\ i = 1,\dots,k, \\ g_i(x,\epsilon) = 0,\ i = k+1,\dots,m\} \tag{11.3}$$

where $f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ and $g : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^m$ are mappings with certain differentiability assumptions to be specified later and we allow even that f and g depend in a nonlinear way on ϵ.
For $K = \mathbb{R}^k_- \times \{0_{m-k}\}$. we obtain:

$$\min_{x\in\mathbb{R}^n}\{f(x,\epsilon)|g(x,\epsilon) \in K\}. \tag{11.4}$$

The feasible set (11.4) is given as

$$S(\epsilon) := \{x \in \mathbb{R}^n|g(x,\epsilon) \in K\},\ \epsilon \in \mathbb{R}^p. \tag{11.5}$$

The optimal value function of (11.4) is denoted as:

$$\begin{aligned} w : \mathbb{R}^p &\to \bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\} \text{ defined as} \\ w(\epsilon) &:= \inf_{x \in \mathbb{R}^n} \{f(x, \epsilon) | g(x, \epsilon) \in K\}, \epsilon \in \mathbb{R}^p. \end{aligned} \tag{11.6}$$

11.1 Local Sensitivity Analysis

Under which conditions can we embed a local minimum $x(0)$ of the unperturbed problem into a continuously differentiable family of perturbed local minima $x(\epsilon)$. W.l.o.g., we use the following notation:

$$\begin{aligned} I(\bar{x}) &= \{i \in \{1, \dots, k\} | g_i(\bar{x}, 0) = 0\} = \{1, \dots, k_0\}, \\ J(\bar{x}) &= \{1, \dots, k_0, k+1, \dots, m\}, \\ m_0 :&= |J(\bar{x})| = m + k_0 - k. \end{aligned} \tag{11.7}$$

Theorem 11.1. Let $\bar{x} \in S(0)$ be a local minimum of (11.4). Let $\bar{x}$ be normal, i.e., the gradients $g_i'(\bar{x})^T$ are linearly independent for $i \in J(\bar{x})$. Assume there are uniquely determined $\bar{\lambda}_i, i \in J(\bar{x})$ such that the second order sufficient optimality conditions of Thm. 10.1, (2) are satisfied with

1. $\bar{\lambda}_i > 0$ for $i = 1, \dots, k_0$,
2. $f_x(\bar{x}, 0) + \sum_{i \in J(\bar{x})} \bar{\lambda}_i g_{ix}(\bar{x}, 0) = 0$,
3. $v^T \left(f_{xx}(\bar{x}, 0) + \sum_{i \in J(\bar{x})} \bar{\lambda}_i g_{ixx}(\bar{x}, 0) \right) v > 0$ for all $v \neq 0$ with

$$v \in C = \{v \in \mathbb{R}^n | g_{ix}(\bar{x}, 0) v = 0, \ i \in J(\bar{x})\}.$$

Then there is a neighbourhood $E \subset \mathbb{R}^p$ of $\epsilon = 0$ and continuously differentiable functions $x : E \to \mathbb{R}^n, \lambda_i : E \to \mathbb{R}, i \in J(\bar{x})$, with:

1. $x(0) = \bar{x}, \ \lambda_i(0) = \bar{\lambda}_i, \ i \in J(\bar{x})$,
2. for all $\epsilon \in E$: $x(\epsilon), \lambda_i(\epsilon), i \in J(\bar{x})$ satisfy the conditions of Thm. 10.1(2.) for the perturbed problem (11.4). In particular, $x(\epsilon)$ is a local minimum of (11.4).

Proof. We define

$$G : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^{m_0}, \ m_0 = m + k_0 - k,$$

via

$$G(x, \epsilon) = (g_1(x, \epsilon), \dots, g_{k_0}(x, \epsilon), g_{k+1}(x, \epsilon), \dots, g_m(x, \epsilon))^T.$$

Per construction, $\bar{x}$ is a strong local minimum of

$$\min\{f(x, 0) | \ G(x, 0) = 0\}.$$

The wanted function $x(\epsilon)$ of local minima to (11.4) should be C^1, thus, they should satisfy

$$g_i(x(\epsilon), \epsilon) < 0 \text{ for } i \notin J(\bar{x}), \|\epsilon\| \text{ small enough.}$$

Hence, $x(\epsilon)$ should be a local minimum of the perturbed problem with equality constraints

$$\min\{f(x, \epsilon) | \; G(x, \epsilon) = 0\}. \tag{11.8}$$

The Lagrange-function therefore is constructed as

$$L(x, \nu, \epsilon) = f(x, \epsilon) + \nu G(x, \epsilon), \;\; \nu \in \mathbb{R}^{m_0}. \tag{11.9}$$

$x(\epsilon)$ and the multiplier $\nu(\epsilon) = (\lambda_i(\epsilon))_{i \in J(\bar{x})}$ needs to solve

$$F(x, \nu^T, \epsilon) := \begin{pmatrix} L_x(x, \nu, \epsilon)^T \\ G(x, \epsilon) \end{pmatrix} = 0, \tag{11.10}$$

where

$$F : \mathbb{R}^n \times \mathbb{R}^{m_0} \times \mathbb{R}^p \to \mathbb{R}^n \times \mathbb{R}^{m_0}.$$

Note that ν^T as the argument of F appears as a column vector. For $\epsilon = 0$ we get per assumption $F(\bar{x}, \bar{\nu}^T, 0) = 0$ with $\bar{\nu} := (\bar{\lambda}_i)_{i \in J(\bar{x})}$ and F is C^1 in a neighbourhood of $(\bar{x}, \bar{\nu}^T, 0)$. The Jacobi-matrix of F wrt. (x, ν^T) in $(\bar{x}, \bar{\nu}^T, 0)$ is given by the $(n+m_0) \times (n+m_0)$ matrix

$$A_0 := \frac{\partial F}{\partial(x, \nu^T)}(\bar{x}, \bar{\nu}^T, 0) = \begin{pmatrix} L_{xx}(\bar{x}, \bar{\nu}, 0) & G_x(\bar{x}, 0)^T \\ G_x(\bar{x}, 0) & 0 \end{pmatrix}. \tag{11.11}$$

In order to apply the implicit function theorem, we will show that A_0 is non-singular. Let $(v, w) \in \mathbb{R}^n \times \mathbb{R}^{m_0}$ with

$$A_0 \begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} L_{xx}(\bar{x}, \bar{\nu}, 0)\, v + G_x(\bar{x}, 0)^T\, w \\ G_x(\bar{x}, 0)\, v \end{pmatrix} = 0. \tag{11.12}$$

Multiplying the equation with $(v, 0)^T$ from the left and using $G_x(\bar{x}, 0)v = 0$, we get

$$v^T L_{xx}(\bar{x}, \bar{\nu}, 0) v = 0.$$

The assumption of the Thm. yield $v = 0$. The equation (11.12) reduces to

$$G_x(\bar{x}, 0)^T\, w = 0.$$

Using that $\bar{x}$ is normal, the matrix $G_x(\bar{x}, 0)^T$ has rank m_0 and therefore $w = 0$. Thus, A_0 is non-singular.

We can apply the implicit function theorem on the system of equations (11.10) and get the existence of a neighbourhood $E \subset \mathbb{R}^p$ of $\epsilon = 0$ and C^1 functions $x : E \to \mathbb{R}^n, \nu(\epsilon) = (\lambda_i(\epsilon))_{i \in J(\bar{x})} : E \to \mathbb{R}^{m_0}$, with:

$$\begin{aligned} F(x(\epsilon), \nu(\epsilon)^T, \epsilon) = 0 \text{ for all } \epsilon \in E, \\ x(0) = \bar{x}, \quad \nu(0) = \bar{\nu}. \end{aligned} \tag{11.13}$$

For completing the proof, we need to verify that $x(\epsilon)$ and $(\lambda_i(\epsilon))_{i \in J(\bar{x})}$ satisfy the 2. order sufficient optimality conditions. Because of

$$\begin{aligned} \lambda_i(0) = \bar{\lambda}_i > 0, \quad i = 1, \dots, k_0, \\ g_i(x(0), 0) = g_i(\bar{x}, 0) < 0, \ \text{for } i = k_0, \dots, k \end{aligned}$$

and the continuity of the functions, we can choose E small enough, such that for all $\epsilon \in E$ we get

$$\begin{aligned} \lambda_i(\epsilon) > 0, \quad i = 1, \dots, k_0, \\ g_i(x(\epsilon), \epsilon) < 0, \ \text{for } i = k_0, \dots, k \end{aligned}$$

With (11.10) and (11.13) we get that $x(\epsilon) \in S(\epsilon)$ for all $\epsilon \in E$ and moreover the KKT-conditions are satisfied:

$$L_x(x(\epsilon), \nu(\epsilon), \epsilon) = 0 \text{ for all } \epsilon \in E.$$

Again using continuity, we can choose E small enough such that for all $\epsilon \in E$:

$$v^T \Big(\lambda_0 f_{xx}(x(\epsilon), \epsilon) + \sum_{i \in J(x(\epsilon))} \bar{\lambda}_i g_{ixx}(x(\epsilon), \epsilon) \Big) v > 0$$

for all $v \neq 0$ with

$$v \in C = \{ v \in \mathbb{R}^n | g_{ix}(x(\epsilon), \epsilon) v = 0, \ i \in J(x(\epsilon)) \}.$$

The matrix $G_x(x(\epsilon), \epsilon)$ has rank m_0 implying that $x(\epsilon)$ is normal. Altogether, $x(\epsilon)$ is a strong local minimum of the perturbed problem (11.4). □

Corollary 11.2. For the functions $x : E \to \mathbb{R}^n, \nu(\epsilon) = (\lambda_i(\epsilon))_{i \in J(\bar{x})} : E \to \mathbb{R}^{m_0}$ appearing in Thm. 11.1, the following statements are true:

1. With the non-singular $(n + m_0) \times (n + m_0)$ matrix A_0 and the $(n + m_0) \times p$ matrix

$$B_0 = \begin{pmatrix} L_{x\epsilon}(\bar{x}, \bar{\nu}, 0) \\ G_\epsilon(\bar{x}, 0) \end{pmatrix}$$

we can compute $\dot{x}(0)$ and $\dot{\nu}(0)$ as

$$\begin{pmatrix} \dot{x}(0) \\ \dot{\nu}(0)^T \end{pmatrix} = -A_0^{-1} B_0.$$

2. Generalized Shadow-Price:

$$\frac{d}{d\epsilon} f(x(\epsilon), \epsilon)\Big|_{\epsilon=0} = L_\epsilon(\bar{x}, \bar{\nu}, 0) = f_\epsilon(\bar{x}, 0) + \bar{\nu} G_\epsilon(\bar{x}, 0).$$

Proof. For (1): The differentiation of (11.13) yieds with (11.10) and (11.11):

$$\frac{d}{d\epsilon} F(x(\epsilon), \nu(\epsilon)^T, \epsilon)\Big|_{\epsilon=0} = A_0 \begin{pmatrix} \dot{x}(0) \\ \dot{\nu}(0)^T \end{pmatrix} + B_0 = 0.$$

For (2): From $G(x(\epsilon), \epsilon) = 0$ we get

$$0 = \frac{d}{d\epsilon} G(x(\epsilon), \epsilon)\Big|_{\epsilon=0} = G_x(\bar{x}, 0)\ \dot{x}(0) + G_\epsilon(\bar{x}, 0).$$

Together with

$$f_x(\bar{x}, 0) = -\bar{\nu} G_x(\bar{x}, 0)$$

we get

$$\begin{aligned} \frac{d}{d\epsilon} f(x(\epsilon), \epsilon)\Big|_{\epsilon=0} &= f_x(\bar{x}, 0)\ \dot{x}(0) + f_\epsilon(\bar{x}, 0) \\ &= -\bar{\nu} G_x(\bar{x}, 0)\ \dot{x}(0) + f_\epsilon(\bar{x}, 0) \\ &= \bar{\nu} G_\epsilon(\bar{x}, 0) + f_\epsilon(\bar{x}, 0) \\ &= L_\epsilon(\bar{x}, \bar{\nu}, 0). \end{aligned}$$

□

11.2 Application to Real-Time Optimization

Part (1) of the Corollary allows to represent the solution $x(\epsilon)$ of the perturbed problem via a Taylor-expansion at the unperturbed solution :

$$\begin{pmatrix} x(\epsilon) \\ \nu(\epsilon)^T \end{pmatrix} \approx \begin{pmatrix} \bar{x} \\ \bar{\nu}^T \end{pmatrix} + \begin{pmatrix} \dot{x}(0) \\ \dot{\nu}(0)^T \end{pmatrix} \epsilon. \tag{11.14}$$

The formula 2. becomes for (11.2):

$$\frac{d}{d\epsilon_i} f(x(\epsilon), \epsilon)\Big|_{\epsilon=0} = \begin{cases} -\bar{\lambda}_i & \text{for } i \in J(\bar{x}) \\ 0, & \text{for } i \notin J(\bar{x}). \end{cases} \tag{11.15}$$

This formula shows how to interpret the Lagrange multipliers as shadow prices.

The formula (11.14) allows for an application in the area of real-time optimization. Suppose we compute offline a solution of the unperturbed problem. If the system data changes, a new solution can be computed in real-time (without resolving the equation system) via the approximation (11.14).

Example 11.3.

$$\begin{aligned} \min\ f(x, \epsilon) = -(0.5+\epsilon)\sqrt{x_1} - (0.5-\epsilon)x_2 \\ x_1 + x_2 \leq 1, \\ x_1 \geq 0.1, \\ x_2 \geq 0. \end{aligned}$$

for $\epsilon = 0$ the assumptions of Thm. 11.1 are satisfied with

$$\begin{aligned} \bar{x} = (0.25, 0.75),\ \bar{x} \text{ is normal} \\ J(\bar{x}) = \{1\},\ G_1(x_1, x_2) = x_1 + x_2 - 1 \\ \bar{\nu} = \bar{\lambda}_1 = 0.5 > 0. \end{aligned}$$

The Lagrange-function (11.9) is given by

$$L(\bar{x}, \bar{\nu}, \epsilon) = -(0.5+\epsilon)\sqrt{\bar{x}_1} - (0.5-\epsilon)\bar{x}_2 + \bar{\nu}(\bar{x}_1 + \bar{x}_2 - 1).$$

The sufficient conditions of 2. order are valid, because

$$\begin{aligned} L_{xx}(\bar{x}, \bar{\nu}, \epsilon) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\ C = \{v \in \mathbb{R}^2 | G_x(\bar{x})v = v_1 + v_2 = 0\} \\ v^T L_{xx}(\bar{x}, \bar{\nu}, \epsilon)v = v_1^2 > 0 \text{ for all } v \in C \setminus \{0\}. \end{aligned}$$

The formula gives

$$A_0 = \begin{pmatrix} L_{xx} & G_x^T \\ G_x & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

$$B_0 = \begin{pmatrix} L_{x\epsilon} \\ G_\epsilon \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} \dot{x}_1(0) \\ \dot{x}_2(0) \\ \dot{x}_3(0) \end{pmatrix} = -A_0^{-1} B_0 = \begin{pmatrix} 2 \\ -2 \\ -1 \end{pmatrix}.$$

Hence, the approximation of first order yields

$$\begin{pmatrix} x_1(\epsilon) \\ x_2(\epsilon) \\ x_3(\epsilon) \end{pmatrix} = \begin{pmatrix} 0.25 \\ 0.75 \\ 0.5 \end{pmatrix} + \begin{pmatrix} 2 \\ -2 \\ -1 \end{pmatrix} \epsilon.$$

For $\epsilon = 0.05$ we get

$$\begin{aligned} &approximation: \quad (0.35, 0.65, 0.45) \\ &exact\ \ value: \quad (0.373, 0.627, 0.45). \end{aligned}$$

Chapter 12

Duality

$$\min \{f(x)| \ g_i(x) \le 0, \ i = 1, \dots, k, \\ g_i(x) = 0, \ i = k+1, \dots, m\} \tag{12.1}$$

with equivalent representation $\min\{f(x)|x \in S\}$, where $S = \{x \in \mathbb{R}^n | g(x) \in K\}$ and $K = \mathbb{R}^k_- \times \{0_{m-k}\}$.

Consider the Lagrange-Function:

$$L(x, \lambda) := f(x) + \lambda g(x), \ \lambda \in \mathbb{R}^k_+ \times \mathbb{R}^{m-k}.$$

We define the Lagrangian-Dual:

$$\mu : \mathbb{R}^m \to \mathbb{R}$$
$$\mu(\lambda) = \inf_{x \in \mathbb{R}^n} L(x, \lambda) = \inf_{x \in \mathbb{R}^n} \{f(x) + \lambda g(x)\}.$$

We assume $\mu(\lambda, x) = -\infty$, if $L(x, \lambda)$ is not bounded from below on $\mathbb{R}^n$.

Theorem 12.1. If μ is finite on $R \subset \mathbb{R}^m$, them μ is concave on R.

Proof. Let $\lambda_1, \lambda_2 \in R$ and let $\alpha \in [0, 1]$. We obtain:

$$\begin{aligned} \mu(\alpha\lambda_1 + (1-\alpha)\lambda_2) &= \inf_{x \in \mathbb{R}^n} \{f(x) + (\alpha\lambda_1 + (1-\alpha)\lambda_2)g(x)\} \\ &\ge \inf_{y \in \mathbb{R}^n} \{\alpha f(y) + \alpha\lambda_1 g(y)\} + \inf_{z \in \mathbb{R}^n} \{(1-\alpha)f(z) + ((1-\alpha)\lambda_2)g(z)\} \\ &= \alpha\mu(\lambda_1) + (1-\alpha)\mu(\lambda_2). \end{aligned}$$

□

12.1 Dual Problem and Weak Duality

Let $p^* = \min\{f(x)|x \in S\}$ be the optimal value of (12.1). We show that for every Lagrange multiplier, i.e., $\lambda_i \ge 0, i = 1, \dots, k$, every value of $\mu(\lambda)$ yields a lower bound on p^*.

Theorem 12.2. For $\lambda \in \mathbb{R}^m$ with $\lambda_i \geq 0, i = 1, \dots, k$, we have:

$$\mu(\lambda) \leq p^*. \tag{12.2}$$

Proof. Let $\bar{x} \in S$, i.e. $g_i(\bar{x}) \leq 0, i = 1, \dots, k$ and $g_i(\bar{x}) = 0$ for $i = k+1, \dots, m$. Then,

$$\sum_{i=1}^{k} \lambda_i \, g_i(\bar{x}) + \sum_{i=k+1}^{m} \lambda_i \, g_i(\bar{x}) \leq 0.$$

We get

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

and obtain

$$\mu(\lambda) = \inf_{x \in \mathbb{R}^n} L(x, \lambda) \leq L(\bar{x}, \lambda) \leq f(\bar{x}).$$

□

For every Lagrange multiplier λ, the corresponding value $\mu(\lambda)$ yields a lower bound on p^*. The dual problem maximizes this lower bound:

$$\max\{\mu(\lambda) | \, \lambda \in \mathbb{R}^m, \lambda_i \geq 0, i = 1, \dots, k\} \tag{12.3}$$

This problem is termed dual problem while the original problem (12.1) is the primal problem.
We say that λ with $\lambda_i \geq 0, i = 1, \dots, k$ is dual feasible, if $\mu(\lambda) > -\infty$. We denote with λ^* the optimal Lagrange multiplier.

Remark 12.3. Problem (12.3) is a convex optimization problem, because the objective is concave and the feasible region is convex. This property is independent on whether or not the primal is convex.

Let d^* be an optimal solution to the dual problem. Thm. 12.2 implies weak duality:

$$d^* \leq p^*.$$

Note that weak duality also holds, if d^* and p^* are infinite. If $p^* = -\infty$, then $d^* = -\infty$ and the dual is infeasible. Otherwise, if $d^* = \infty$, we get $p^* = \infty$, hence the primal problem is infeasible. The difference $p^* - d^*$ is known as duality gap.

12.2 Strong Duality and Saddle Points

If

$$d^* = p^*,$$

we say that strong duality holds.

Theorem 12.4. Let x^* be feasible for (12.1) and λ^* feasible for (12.3). Suppose that strong duality holds, i.e.

$$f(x^*) = \mu(\lambda^*).$$

Then:

1. x^* and λ^* are globally optimal for (12.1) and (12.3).
2. $\lambda_i^* g_i(x^*) = 0$ for all $i = 1, \dots, m$.

Proof. For 1. let $\bar{x}$ be a global optimal solution of (12.1).

$$f(\bar{x}) \geq \mu(\lambda^*) = f(x^*),$$

where the first inequality follows from (12.2).
For 2. we observe that for all $i > k$, the feasibility of x^* yields $g_i(x^*) = 0$ and therefore $\lambda_i^* g_i(x^*) = 0$ is implied. For $i \leq k$ we consider the inequality:

$$f(x^*) = \mu(\lambda^*) \leq L(x^*, \lambda^*) = f(x^*) + \lambda^* g(x^*).$$

We get $\lambda^* g(x^*) \geq 0$. On the other hand $\lambda^* g(x^*) \leq 0$, and hence we get $\lambda^* g(x^*) = 0$. As for every $i \leq k$ we have $\lambda_i^* \geq 0$ and $g_i(x^*) \leq 0$, the claim follows for all $i \leq k$. □

We define a saddle point.

Definition 12.5. We are given a problem of the form (12.1). Let $(\bar{x}, \bar{\lambda})$ satisfy:

1. $\bar{\lambda}_i \geq 0, i = 1, \dots, k$
2. $L(\bar{x}, \lambda) \leq L(\bar{x}, \bar{\lambda})$ for all $\lambda \in \mathbb{R}^m$ with $\lambda_i \geq 0, i = 1, \dots, k$.
3. $L(\bar{x}, \bar{\lambda}) \leq L(x, \bar{\lambda})$ for all $x \in \mathbb{R}^n$.

Then, $(\bar{x}, \bar{\lambda})$ is called saddle point for (12.1). The conditions are called saddle point conditions.

Theorem 12.6 (saddle point theorem). Let $(\bar{x}, \bar{\lambda})$ be a saddle point for (12.1). Then strong duality holds for $\bar{x}$ and $\bar{\lambda}$.

Proof. Condition 2. implies

$$f(\bar{x}) + \nu g(\bar{x}) \leq f(\bar{x}) + \bar{\lambda} g(\bar{x}) \text{ for all } \nu \in \mathbb{R}^m \text{ with } \nu_i \geq 0, i = 1, \dots, k.$$

After basic manipulations, we get

$$(\nu - \bar{\lambda}) g(\bar{x}) \leq 0 \text{ for all } \nu \in \mathbb{R}^m \text{ with } \nu_i \geq 0, i = 1, \dots, k. \tag{12.4}$$

Inserting $\nu_i = \bar{\lambda}_i$ for $i = 1, \dots, k$, we get

$$\sum_{i=k+1}^{m} (\nu_i - \bar{\lambda}_i) g_i(\bar{x}) \le 0 \text{ for all } \nu_i \in \mathbb{R}, i = k+1, \dots, m. \tag{12.5}$$

This implies $g_i(\bar{x}) = 0$ for all $i = k+1, \dots, m$. Hence, inequality (12.4) reduces to

$$\sum_{i=1}^{k} (\nu_i - \bar{\lambda}_i) g_i(\bar{x}) \le 0 \text{ for all } \nu_i \ge 0, i = 1, \dots, k. \tag{12.6}$$

This condition implies $g_i(\bar{x}) \le 0$ for all $i = 1, \dots, k$. Thus, $\bar{x} \in S$. Moreover, (12.6) yields

$$\bar{\lambda}_i g_i(\bar{x}) = 0, \text{ for all } i = 1, \dots, k.$$

Exercise 12.7. Show that (12.6) implies the following:

1. $g_i(\bar{x}) \le 0$ for all $i = 1, \dots, k$.
2. $\bar{\lambda}_i g_i(\bar{x}) = 0$, for all $i = 1, \dots, k$.

Altogether, we obtain

$$f(\bar{x}) = f(\bar{x}) + \sum_{i=1}^{k} \bar{\lambda}_i g_i(\bar{x}) = f(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i g_i(\bar{x}) = L(\bar{x}, \bar{\lambda}).$$

The condition 3. yields

$$L(\bar{x}, \bar{\lambda}) \le L(x, \bar{\lambda}) \text{ for all } x \in \mathbb{R}^n.$$

Hence,

$$f(\bar{x}) = L(\bar{x}, \bar{\lambda}) \le L(x, \bar{\lambda}) \text{ for all } x \in \mathbb{R}^n,$$

which in turn implies

$$f(\bar{x}) \le \inf_{x \in \mathbb{R}^n} L(x, \bar{\lambda}) = \mu(\bar{\lambda}).$$

With weak duality, we get $f(\bar{x}) = \mu(\bar{\lambda})$. □

Exercise 12.8. Show that for any saddle point $(\bar{x}, \bar{\lambda})$ of (12.1) the KKT-conditions of Thm. 9.2 are satisfied.

12.3 Strong Duality for Convex Problems

We consider a convex optimization problem:

$$\begin{aligned} \min \ & f(x) \\ g_i(x) &\le 0, \ i = 1, \dots, k \\ Ax &= b, \end{aligned} \tag{12.7}$$

$$x \in \bar{S},$$

where $\bar{S} \subset \mathbb{R}^n$ convex, A being a $(m-k) \times n$ matrix of full rank and f and g_i convex for $i = 1, \ldots, k$. We assume that the following regularity conditions are satisfied (Slater): there is $\bar{x} \in \operatorname{int}(\bar{S})$ with

$$g_i(\bar{x}) < 0, \;\; i = 1, \ldots, k, \;\; A\,\bar{x} = b.$$

Theorem 12.9 (Strong Duality). Consider (12.7) and assume that the Slater-regularity conditions are satisfied. Then, $d^* = p^*$.

Proof. Suppose that p^* is finite. As the primal problem is feasible, only the case $p^* = -\infty$ can occur which implies $d^* = -\infty$ by weak duality. We define

$$A_1 \subseteq \mathbb{R}^k \times \mathbb{R}^{m-k} \times \mathbb{R}$$
$$A_1 = \{(u, v, t) | \exists x \in \bar{S},\; g_i(x) \le u_i,\; i = 1, \ldots, k,\; A_j\, x - b_j = v_j,\; j = 1, \ldots, m-k,\; f(x) \le t\}$$

Note that A_1 is convex. We define the convex set:

$$A_2 = \{(0, 0, s) \in \mathbb{R}^k \times \mathbb{R}^{m-k} \times \mathbb{R} | s < p^*\}.$$

We get that A_1 and A_2 do not intersect: Assume by contradiction $(u, v, t) \in A_1 \cap A_2$. With $(u, v, t) \in A_2$ we get $u = 0, v = 0$, and $t < p^*$. Because $(u, v, t) \in A_1$, we get x with $g_i(x) \le 0, i = 1, \ldots, k$, $Ax - b = 0$, and $f(x) \le t < p^*$, in contradiction to the optimality of p^*.
With the separating hyperplane theorem we get $(w_1, w_2, w_3) \neq 0$ and $\alpha \in \mathbb{R}$ with

$$(u, v, t) \in A_1 \Rightarrow w_1^T u + w_2^T v + w_3\, t \ge \alpha. \tag{12.8}$$

and

$$(u, v, t) \in A_2 \Rightarrow w_1^T u + w_2^T v + w_3\, t \le \alpha. \tag{12.9}$$

From (12.8) follows $w_1 \ge 0$ and $w_3 \ge 0$, as otherwise $w_1^T u + w_3\, t$ is unbounded from below on A_1 in contradiction to (12.8). Condition (12.9) implies $w_3\, t \le \alpha$ for all $t \le p^*$ and hence $w_3\, p^* \le \alpha$. Together with (12.8) we get for all $x \in \bar{S}$ (choose $g(x) = u$ and $Ax - b = v$), such that

$$\sum_{i=1}^{k} (w_1)_i\, g_i(x) + w_2^T (Ax - b) + w_3\, f(x) \ge \alpha \ge w_3\, p^*. \tag{12.10}$$

We argue by a case distinction: If $w_3 > 0$, we divide (12.10) by w_3 and obtain

$$L(x, w_1/w_3, w_2/w_3) \ge p^*,$$

for all $x \in \bar{S}$. We get $\mu(\lambda, \nu) \ge p^*$, where $\lambda = w_1/w_3$ and $\nu = w_2/w_3$. (Here $\lambda \in \mathbb{R}^k$ and $\nu \in R^{m-k}$ are the corresponding multipliers for $g(x) \le 0$ and $Ax = b$). With weak duality

$\mu(\lambda, \nu) \leq p^*$ we get equality.

Now, we consider the case $w_3 = 0$ and lead this to a contradiction. From (12.10), we get for all $x \in \bar{S}$:

$$\sum_{i=1}^{m} (w_1)_i \, g_i(x) + w_2^T (Ax - b) \geq 0. \tag{12.11}$$

We apply this to the point $\bar{x}$ satisfying the Slater-conditions and get

$$\sum_{i=1}^{k} (w_1)_i \, g_i(\bar{x}) \geq 0.$$

As $g_i(\bar{x}) < 0$ and $w_1 \geq 0$ we get $w_1 = 0$.

From $(w_1, w_2, w_3) \neq 0$ and $w_1 = 0, w_3 = 0$, we get $w_2 \neq 0$. Hence, (12.11) implies $w_2^T(Ax - b) \geq 0$ for all $x \in \bar{S}$. The point $\bar{x}$ satisfies $w_2^T(A\bar{x} - b) = 0$. With $\bar{x} \in \mathrm{int}(\bar{S})$ we get that $\bar{x} \pm \epsilon \in \mathrm{int}(\bar{S})$ for $\epsilon \in \mathbb{R}^n$ with $\|\epsilon\|$ small enough. We get $A^T w_2 = 0$, as otherwise there are points in $\mathrm{int}(\bar{S})$ with $w_2^T(Ax - b) < 0$, contradiction. Conditions $A^T w_2 = 0$ and $w_2 \neq 0$ contradict the assumption of A having full rank. □

Chapter 13

Numerical Methods

13.1 Unconstrained Optimization Problems

We first consider

$$\min\{f(x) | x \in \mathbb{R}^n\} \tag{13.1}$$

without constraints, where $f \in C^2$ on $\mathbb{R}^n$.

Definition 13.1. $v \in \mathbb{R}^n$ is called descent direction of f in $\bar{x} \in \mathbb{R}^n$, if there is $\bar{t} > 0$ with

$$f(\bar{x} + tv) < f(\bar{x}) \text{ for all } t \in (0, \bar{t}).$$

We obtain the following Lemma.

Lemma 13.2. If $f'(\bar{x})v < 0$, then v is a descent direction of f in $\bar{x}$.

Proof. Definie $\Phi(t) := f(\bar{x} + tv)$. We get $\dot{\phi}(0) = f'(\bar{x})v < 0$ and the claim follows. □

Remark 13.3.

- The condition $f'(\bar{x})v < 0$ means that the angle ϕ between v and $-f'(\bar{x})$ in $\bar{x}$ is less than $\pi/2$ (or 90°). Consider:

$$0 > f'(\bar{x})v \Rightarrow 0 < -f'(\bar{x})v = \cos(\phi) \left\|-f'(\bar{x})\right\| \|v\| .$$

  We get $\cos(\phi) > 0$ and hence $\phi \in [0, \pi/2)$.

- The criterion $f'(\bar{x})v < 0$ is not necessary. If $\bar{x}$ is a strict local maximum, then all $v \in \mathbb{R}^n$ are directions of descent for f in $\bar{x}$, but $f'(\bar{x})v < 0$ need not be satisfied.

Exercise 13.4. Let $B \in \mathbb{R}^{n \times n}$ be symmetric and positive definit. Then, $v = -B\, f'(\bar{x})$ is a descent direction of f in $\bar{x}$, if $f'(\bar{x}) \neq 0$.

We describe a descent method to compute $\bar{x}$ with $f'(\bar{x}) = 0$.

1. Choose $x^0 \in \mathbb{R}^n$, $k = 0$ and fix $\epsilon_0 > 0$;
2. If $\left\|f'(x^k)\right\| \leq \epsilon_0$:STOP (Termination);
3. Compute a descent direction v^k with $f'(x^k)v^k < 0$;
4. Compute a step-length t_k with $f(x^k + t_k v^k) < f(x^k)$;
5. Set $x^{k+1} \leftarrow x^k + t_k v^k$; $k \leftarrow k+1$ and go to step 2.

Figure 13.1: Generic Method of Descent.

Definition 13.5 (Gradient-Method, Newton-Method, Quasi-Newton-Method).

1. For
$$v^k := -f'(x^k)^\intercal$$
we obtain the Gradient-Method.
2. For the Newton-Method, we choose:
$$v^k := -f''(x^k)^{-1} f'(x^k)^\intercal.$$
In every point x^k in which the Hesse-matrix $f''(x^k)$ is positive definite, the vector v^k is a descent direction (assuming $f'(x^k) \neq 0$).
3. The Quasi-Newton-Method chooses
$$v^k := -H_k^{-1} f'(x^k)^\intercal,$$
for a suitable positive definite matrix H_k.

Theorem 13.6. If $H \in \mathbb{R}^{n \times n}$ is symmetric, positive definite and $f'(x) \neq 0$, then the gradient direction
$$v := \frac{H^{-1} f'(x)^\intercal}{\|H^{-1} f'(x)^\intercal\|_H}$$
maximizes the descent of $f'(x)v$ over all $v \in \mathbb{R}^n$ with $\|v\|_H = 1$, where $\|x\|_H := \sqrt{x^\intercal H x}$.

Proof. We prove the theorem only for $H = I$. With Cauchy-Schwarz-inequality we get using $\|v\| = 1$:
$$|f'(x)v| \leq \left\|f'(x)\right\| \|v\| = \left\|f'(x)\right\|.$$
This bound is attained for $v := \pm \frac{f'(x)^\intercal}{\|f'(x)^\intercal\|}$. □

We go back to the angle-condition discussed before.

Definition 13.7 (Angle-Condition). A generic descent method of the form (13.1) satisfies the angle-condition, if:

$$\text{there is } c > 0, \text{ such that for all } k \in \mathbb{N} \text{ we have: } c_k := -\frac{f'(x^k)v^k}{\|f'(x^k)\| \, \|v^k\|} \geq c. \tag{13.2}$$

A weaker condition is the Zoutendijk-Condition, which only requires $\sum_{k=0}^{\infty} c_k = \infty$.

13.2 Choice of the Step-Length

We discuss the degree of freedom regarding the step-length choice in 13.1. Let x^0 be the initial point and assume that $N(f, f(x^0)) = \{x \in \mathbb{R}^n | f(x) \leq f(x^0)\}$ is compact. With the continuity of $f''(x)$ on $N(f, f(x^0))$, there is $C > 0$ with

$$\left\|f''(x)\right\| \leq C \text{ for all } x \in N(f, f(x^0)).$$

We apply Taylor expansion of f in x in direction of $tv, v \in \mathbb{R}^n$:

$$\begin{aligned} f(x + tv) &= f(x) + tf'(x)v + \frac{t^2}{2} v^T f''(z) v \\ &\leq f(x) + tf'(x)v + \frac{t^2}{2} C \|v\|^2 . \end{aligned} \tag{13.3}$$

Here $z = x + \xi t v$ is an intermediate point with $0 < \xi < 1$. The bound (13.3) is valid for all $t > 0$ with $x + [0, t]v \subset N(f, f(x^0))$.

The last term of (13.3) is a polynomial $p(t)$ of degree two in t and attains at

$$t^* = -\frac{f'(x)v}{C \|v\|^2} > 0$$

its strict global minimum. Let $\bar{t}$ be the unique maximal step-length with

$$x + tv \in N(f, f(x^0)) \text{ for all } t \in [0, \bar{t}].$$

We get

$$p(\bar{t}) \geq f(x + \bar{t}v) \geq f(x) = p(0).$$

Using $p'(0) = f'(x)v < 0$ we get that t^* lies in $(0, \bar{t})$ and thus $x + t^*v \in N(f, f(x^0))$. With (13.3) we get

$$f(x + t^*v) \leq p(t^*) = f(x) - \frac{1}{2C} \left(\frac{f'(x)v}{\|v\|} \right)^2 .$$

This bounds the minimum descent. Since C is not known a priori, we define the following.

Definition 13.8. A step-length strategy $t(x, v)$ is called efficient, if for every $x^0 \in \mathbb{R}^n$, there is $\xi > 0$ with

$$f(x + t(x, v)v) \le f(x) - \xi \left(\frac{f'(x)v}{\|v\|} \right)^2 \text{ for all } x \in N(f, f(x^0)),$$

and v is a descent direction of f in x with $f'(x)v < 0$.

Under the assumptions we get that $t := \arg\min\{f(x + tv) | t > 0\}$ is efficient.

We obtain the following theorem on the general descent method (13.1) with efficient step-length strategies.

Theorem 13.9. Let $f \in C^2$ and let (13.1) with $\epsilon_0 = 0$ satisfy the condition (13.2). Suppose we choose an efficient step-length strategy. Then, one of the following statements is true:

1. After finitely many iterations we have $f'(x^k) = 0$.
2. $\lim_{k\to\infty} f(x^k) = -\infty$
3. $\lim_{k\to\infty} f'(x^k) = 0$, i.e., every accumulation point of $x^k, k \in \mathbb{N}$ is a zero of $f'(x)$.

Proof. If 13.1 terminates after finitely many iterations, we get using $\epsilon_0 = 0$ the condition $f'(x^k) = 0$.

So suppose that this does not hold. Using the angle- and efficiency condition we get for iteration k:

$$f(x^{k+1}) - f(x^k) \le -\xi \left(\frac{f'(x^k)v^k}{\|v^k\|} \right)^2 = -\xi c_k^2 \left\| f'(x^k) \right\|^2.$$

After $N \in \mathbb{N}$ iterations we get

$$f(x^N) - f(x^0) = \sum_{k=0}^{N-1} f(x^{k+1}) - f(x^k) \le -\xi \sum_{k=0}^{N-1} c_k^2 \left\| f'(x^k) \right\|^2.$$

We divide by $-\xi < 0$ and get

$$-\frac{f(x^N) - f(x^0)}{\xi} \ge \sum_{k=0}^{N-1} c_k^2 \left\| f'(x^k) \right\|^2.$$

If f is bounded from below we get

$$\lim_{N\to\infty} f(x^N) > -\infty$$

and therefore

$$\lim_{N\to\infty} -\frac{f(x^N) - f(x^0)}{\xi} < \infty.$$

Hence we get

$$\sum_{k=0}^{\infty} c_k^2 \left\|f'(x^k)\right\|^2 < \infty.$$

Using the angle- and efficiency condition, we get $\left\|f'(x^k)\right\| \to 0$.

□

We provide two additional step-length methods, the Armijo-Rule and the Goldstein-Rule.

Definition 13.10 (Armijo-Rule). For $\sigma \in (0,1), \alpha \in (0,1)$ we choose $t := \alpha^\ell$ with

$$\ell := \min\{j \in \mathbb{N}_0 | f(x + \alpha^j v) \leq f(x) + \sigma\alpha^j (f'(x)v)\}.$$

As for the interpretation of the Armijo-Rule, we define for

$$\Phi(t) = f(x + tv), t \geq 0$$

the auxiliary function

$$\Psi(t) = \Phi(t) - (f(x) + \sigma t(f'(x)v)).$$

Per construction, we get $\Psi(0) = 0$ and

$$\begin{aligned}\Psi'(0) &= \Phi'(0) - \sigma f'(x)v \\ &= f'(x)v - \sigma f'(x)v = (1-\sigma)f'(x)v < 0.\end{aligned}$$

Assuming $N(f, f(x^0))$ to be compact, we know that $\Phi(t)$ grows for t large enough. Hence, there is a unique $\ell \in \mathbb{N}_0$ and thus a maximal $t = \alpha^\ell > 0$ satisfying the Armijo-Condition. This t is usually computed via enumeration $\ell = 1, 2, \ldots$.

Remark 13.11.

- The Armijo-Method may not be efficient in general.
- The scaled Armijo-step-length works with a scaling factor $s > 0$ and is defined as

$$\ell := \min\{j \in \mathbb{N}_0 | f(x + s\alpha_j v) \leq f(x) + \sigma s\alpha_j (f'(x)v)\}.$$

  For large enough s, the Armijo-Variant is efficient.

Definition 13.12 (Goldstein-Rule). In order to avoid small step-length, we bound the feasible space from below. The step-length $t > 0$ satisfies the Goldstein-Condition, if for fixed $\sigma \in (0, 0.5)$, we have

$$\Phi_u(t) \leq \Phi(t) \leq \Phi_o(t), \tag{13.4}$$

where $\Phi_u(t)$ and $\Phi_o(t)$ are defined as follows:

$$\Phi_o(t) := \Phi(0) + \sigma t\Phi'(0), \Phi_u(t) := \Phi(0) + (1-\sigma)t\Phi'(0). \tag{13.5}$$

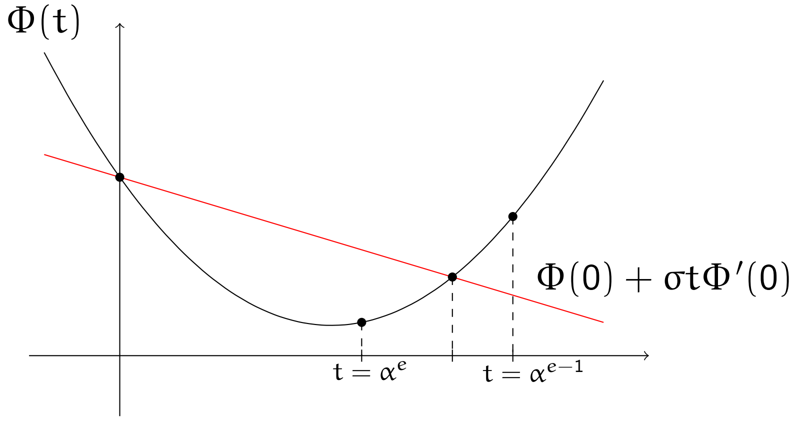


Figure 13.2: Armijo-step-length strategy.

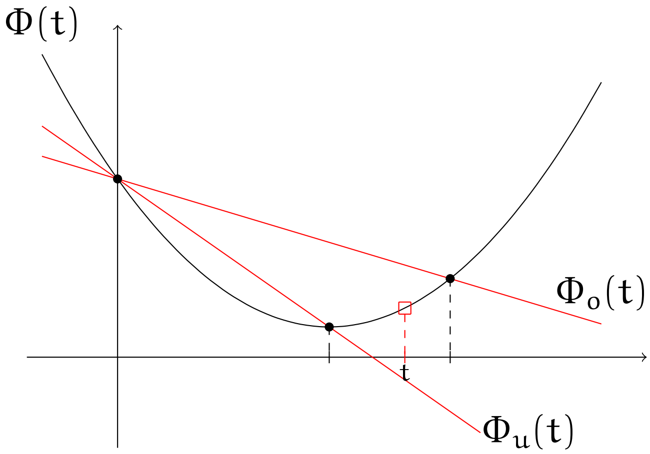


Figure 13.3: Goldstein-step-length strategy.

Any step-length method satisfying the Goldstein-Rule is efficient. We provide a concrete implementation.

Theorem 13.13. Let $f \in C^2$ and $x \in \mathbb{R}^n$ with $N(f, (f(x))$ compact. Let $v \in \mathbb{R}^n$ be a descent direction with $f'(x)v < 0$ and

$$C := \max\left\{ \left\| f''(y) \right\|^2 \mid y \in N(f, (f(x)) \right\}.$$

Then, every step-length t which satisfies the Goldstein-Rule also satisfies the following efficiency condition:

$$f(x + tv) \leq f(x) - \frac{\sigma}{C} \left(\frac{f'(x)v}{\|v\|} \right)^2 .$$

Proof. Let t^* be the infimum of the positive local minimizers of Φ. The value t^* is then a stationary point and Φ is strictly decreasing in $[0, t^*]$.

1. Set $t_u := 0$, $t_o > 0$ (arbitrary);
2. If $\Phi(t_o) < \Phi_u(t_o)$, set $t_u := t_o, t_o := 2t_o$.
   Repeat until $\Phi(t_o) \geq \Phi_u(t_o)$;
3. If $\Phi(t_o) \leq \Phi_o(t_o)$, set $t := t_o$, Stop.;
4. Repeat: set $t := (t_u + t_o)/2$
   If $\Phi(t) < \Phi_u(t)$, set $t_u := t$;
   If $\Phi(t) > \Phi_o(t)$, set $t_o := t$;
   Until: $\Phi_u(t) \leq \Phi(t) \leq \Phi_o(t)$, Stop.

Figure 13.4: Step-length choice after Goldstein.

Case 1: $t \leq t^*$
With $\Phi_u(t) := \Phi(0) + (1-\sigma)t\Phi'(0) \leq \Phi(t)$ we get using Taylor expansion

$$(1-\sigma)t\Phi'(0) \leq \Phi(t) - \Phi(0) = t\Phi'(0) + \frac{t^2}{2}\Phi''(\tilde{t}) \leq t\Phi'(0) + C\,\|v\|^2$$

and therefore

$$-\sigma t\Phi'(0) \leq \frac{t^2}{2}C\,\|v\|^2\,,$$

or equivalently

$$t \geq -\frac{2\sigma}{C}\frac{\Phi'(0)}{\|v\|^2} := \hat{t} > 0.$$

Using monotonicity of Φ on $0 < \hat{t} \leq t \leq t^*$, we get

$$\begin{aligned}\Phi(t) &\leq \Phi(\hat{t}) \leq \Phi(0) + \hat{t}\Phi'(0) + \frac{\hat{t}^2}{2}C\,\|v\|^2 \\ &= \Phi(0) - \frac{2\sigma(1-\sigma)}{C}\left(\frac{\Phi'(0)}{\|v\|^2}\right)^2 \\ &\leq \Phi(0) - \frac{\sigma}{C}\left(\frac{\Phi'(0)}{\|v\|^2}\right)^2\end{aligned}$$

The last inequality uses $\sigma \in (0, 0.5)$.

Case 2: $t > t^*$

We get

$$0 = \Phi'(t^*) = \Phi'(0) + t^*\Phi''(\tilde{t}) \leq \Phi'(0) + C\,\|v\|^2$$

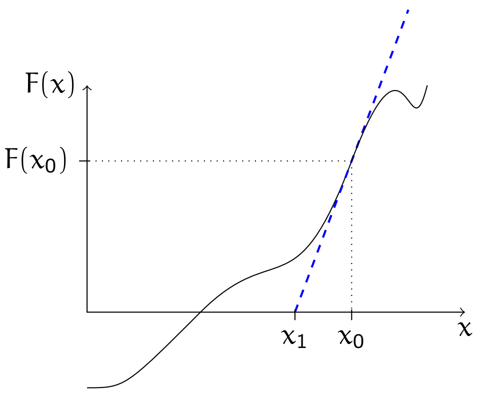


Figure 13.5: Illustration of the Newton-Method.

and thus $t^* \geq -\frac{\Phi'(0)}{C\|v\|^2}$. We obtain

$$\begin{aligned}\Phi(t) &\leq \Phi_o(t) = \Phi(0) + \sigma t \Phi'(0)\\ &\leq \Phi(0) + \sigma t^* \Phi'(0) \leq \Phi(0) - \frac{\sigma}{C}\left(\frac{\Phi'(0)}{\|v\|^2}\right)^2.\end{aligned}$$

□

13.3 Lagrange-Newton Method

We recap the Newton-Method. Given is a C^1 function

$$F : \mathbb{R}^n \to \mathbb{R}^n$$

and we search for

$$F(x) = 0.$$

The linear approximation in x^k is defined as

$$F_k(x) := F(x^k) + F'(x^k)(x - x^k).$$

Hence,

$$x^{k+1} = x^k - F'(x^k)^{-1} F(x^k),$$

if the inverse F'^{-1} exists. In a concrete implementation, we don't compute the inverse but a direction vector $v = x - x^k$ solving the Newton-Equation:

$$F'(x^k)v = -F(x^k).$$

For a solution v^k we set

$$x^{k+1} := x^k + v^k.$$

Theorem 13.14. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be C^1, $\bar{x}$ a root of F and the Jacobi-Matrix $F'(\bar{x})$ be regular. Then, there is an open ball U around $\bar{x}$ such that:

1. The Newton-Method is well-defined and produces a convergent sequence $x^k, k \in \mathbb{N}$ with limit point $\bar{x}$.
2. The convergence is superlinear.
3. If F' is locally Lipschitz, then the convergence is quadratic.

For a proof, see Kanzow und Geiger (Satz, 5.26).
We consider now a constrained optimization problem:

$$\min \{f(x)| \;\; g(x) = 0\}, \;\text{ where } g : \mathbb{R}^n \to \mathbb{R}^m. \tag{13.6}$$

Using the Lagrange-Function

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

the KKT-conditions read as

$$F(x, \lambda) := \begin{pmatrix} L_x(x, \lambda) \\ L_\lambda(x, \lambda) \end{pmatrix} = \begin{pmatrix} f'(x) + \lambda g'(x) \\ g(x) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{13.7}$$

The equation (13.7) is a nonlinear system of $n + m$ equations in x, λ. The corresponding Newton-Update reads as

$$\begin{pmatrix} x^{k+1} \\ \lambda^{k+1} \end{pmatrix} = \begin{pmatrix} x^k \\ \lambda^k \end{pmatrix} - F'(x^k, \lambda^k)^{-1} F(x^k, \lambda^k). \tag{13.8}$$

Formally, we get:
We obtain the following result regarding the Lagrange-Newton method.

Theorem 13.15. Let $(\bar{x}, \bar{\lambda})$ be a normal KKT-point of 13.6, satisfying the second order sufficient optimality conditions of Thm. 10.1, Condition (2). Then, the Lagrange-Newton method converges quadratically against a KKT-point of 13.6, where $\epsilon_0 = 0$ is assumed.

Proof. We only need to show that the Jacobi-Matrix $F'(\bar{x}, \bar{\lambda})$ is regular. This was already shown in the proof of Thm. 11.1. □

1. Choose $x^0 \in \mathbb{R}^n, \lambda^0 \in \mathbb{R}^m$, $k = 0$ and fix $\epsilon_0 > 0$;
2. If $\left\|F(x^k, \lambda^k)\right\| \leq \epsilon_0$:STOP (Termination);
3. Compute $(\Delta x^k, \Delta\lambda^k)$ as solution of

$$F'(x^k, \lambda^k) \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = -F(x^k, \lambda^k). \tag{13.9}$$

4. Set $x^{k+1} \leftarrow x^k + \Delta x^k; \lambda^{k+1} \leftarrow \lambda^k + \Delta\lambda^k;\ k \leftarrow k+1$ and go to step 2.

Figure 13.6: Lagrange-Newton Method.