

**Skript zur Vorlesung  
Einführung in die Optimierung**

Christoph Buchheim und Christian Meyer

Stand: 28. Juli 2017

## Inhaltsverzeichnis

Kapitel 1. Einführung	5
1 Grundbegriffe	5
2 Klassifikation und Beispiele	6
3 Notation und Wiederholung	8
<b>Teil 1. Konvexe Optimierung</b>	<b>11</b>
Kapitel 2. Konvexe Mengen und Funktionen	13
4 Konvexe Mengen	13
5 Konvexe Funktionen	20
6 Projektion auf konvexe Mengen	23
7 Trennungssätze	24
8 Farkas-Lemma	27
Kapitel 3. Konvexe Optimierungsprobleme	31
9 Variationsungleichung	31
10 KKT-Bedingungen und Lagrange-Dualität	32
11 Subgradientenverfahren	44
<b>Teil 2. Lineare Optimierung</b>	<b>47</b>
Kapitel 4. Polytope und Polyeder	49
12 Darstellungen von Polyedern	49
13 Seitenflächen, Facetten und Ecken	52
14 Satz von Weyl-Minkowski	57
Kapitel 5. Grundlagen der linearen Optimierung	63
15 Existenzaussagen	63
16 Dualitätssätze	65
Kapitel 6. Simplex-Algorithmus	71
17 Simplex-Algorithmus in primaler Form	71
18 Praktische Umsetzung des Simplex-Algorithmus	78
19 Reoptimierung	80

<b>Teil 3. Diskrete Optimierung</b>	83
Kapitel 7. Ganzzahlige Programmierung	85
20 Ganzzahlige Polyeder	85
21 Branch-and-Bound	88
<b>Teil 4. Nichtlineare Optimierung</b>	91
Kapitel 8. Freie Optimierung	93
22 Existenz optimaler Lösungen	93
23 Optimalitätsbedingungen	95
24 Algorithmen der freien Optimierung	97

## KAPITEL 1

# Einführung

### Inhalt

---

1	<b>Grundbegriffe</b>	5
2	<b>Klassifikation und Beispiele</b>	6
3	<b>Notation und Wiederholung</b>	8

---

### § 1 Grundbegriffe

Die mathematische Optimierung befasst sich mit Aufgaben der Form

$$\left. \begin{array}{l} \text{minimiere } f(x) \\ \text{s.t. } x \in \mathcal{F}, \end{array} \right\} \quad (\text{P})$$

wobei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  die *Zielfunktion* bezeichnet, die minimiert werden soll, und  $\mathcal{F} \subset \mathbb{R}^n$  die *zulässige Menge*, die Restriktionen an die Optimierungsvariable  $x \in \mathbb{R}^n$  modelliert. Die Abkürzung „s.t.“ steht für „subject to“, engl. für „unter der Nebenbedingung“.

Die wesentlichen Fragestellungen bei der Diskussion von (P) sind:

- Existiert eine Lösung von (P)?
- Wodurch werden Lösungen von (P) charakterisiert (Optimalitätsbedingungen)?
- Wie kann man mit Hilfe numerischer Algorithmen Lösungen (approximativ) ausrechnen?

Zur Beantwortung dieser Fragen müssen wir uns zunächst überlegen, was wir unter einer „Lösung“ überhaupt verstehen.

#### Definition 1.1

Ein Punkt  $x^* \in \mathcal{F}$  heißt *globaler Optimierer* oder *globaler Minimierer* von (P), falls

$$f(x^*) \leq f(x) \quad \forall x \in \mathcal{F} \quad (1.1)$$

gilt. Die Menge der globalen Minimierer von (P) wird mit

$$\arg \min_{x \in \mathcal{F}} f(x)$$

bezeichnet. Wir nennen  $x^*$  *lokalen Minimierer*, falls ein  $\varepsilon > 0$  existiert, so dass

$$f(x^*) \leq f(x) \quad \forall x \in \mathcal{F} \cap B_\varepsilon(x^*), \quad (1.2)$$

wobei  $B_\varepsilon(x^*)$  die Kugel um  $x^*$  mit Radius  $\varepsilon$  bezeichnet. Ein globaler bzw. lokaler Minimierer heißt *strikt*, falls (1.1) bzw. (1.2) für  $x \neq x^*$  mit echter Ungleichung erfüllt sind.  $\diamond$

**Bemerkung 1.2**

Die Betrachtung von Minimierungsaufgaben stellt keine Einschränkung dar, da Maximierungsaufgaben wegen

$$\arg \max_{x \in \mathcal{F}} f(x) = \arg \min_{x \in \mathcal{F}} (-f(x))$$

und

$$\max_{x \in \mathcal{F}} f(x) = - \min_{x \in \mathcal{F}} (-f(x))$$

auf Minimierungsaufgaben zurückgeführt werden können.  $\diamond$

**§ 2 Klassifikation und Beispiele**

Die analytischen Methoden aber auch die numerischen Algorithmen zur Lösung von (P) hängen stark von der jeweiligen Struktur des Optimierungsproblems ab. Je nachdem welche Struktur die Zielfunktion und die zulässige Menge besitzt, unterscheiden wir:

- *Nichtlineares Programm (NLP)*:  
Hierbei ist  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  eine (glatte) nichtlineare Zielfunktion und die zulässige Menge ist durch

$$\mathcal{F} = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\}$$

gegeben, wobei  $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$  und  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $m, p \in \mathbb{N}$ , (glatte) nichtlineare Funktionen sind.

- *Diskretes Optimierungsproblem*:  
Bei einem diskreten oder ganzzahligen Optimierungsproblem sind einige der Optimierungsvariablen auf ganzzahlige Punktmengen beschränkt, d.h. derartige Probleme haben die Form:

$$\left. \begin{array}{l} \text{minimiere } f(x) \\ \text{s.t. } x \in \mathcal{F}_1, \quad x_{\mathcal{I}} \in \mathcal{F}_2 \end{array} \right\} \quad (\text{DP})$$

mit einer Indexmenge  $\mathcal{I} \subset \{1, \dots, n\}$  und zulässigen Mengen  $\mathcal{F}_1 \subset \mathbb{R}^n$  und  $\mathcal{F}_2 \subset \mathbb{Z}^{\mathcal{I}}$ . Hierbei bezeichnet  $x_{\mathcal{I}}$  die Einschränkung des Vektors  $x$  auf  $\mathcal{I}$ .

- *Gemischt ganzzahliges nichtlineares Programm (MINLP)*<sup>1</sup>:  
Zusätzlich zu den Nebenbedingungen eines NLPs sollen hierbei einige Variablen ganzzahlig sein, d.h. die zulässige Menge ist durch

$$\mathcal{F} = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0, x_{\mathcal{I}} \in \mathbb{Z}^{\mathcal{I}}\}$$

gegeben.

- *Konvexes Optimierungsproblem*:  
Hierbei sind zulässige Menge und Zielfunktion *konvex*.
- *Quadratisches Programm (QP)*:  
QPs sind Spezialfälle von NLPs, bei denen die Zielfunktion quadratisch und die Funktionen in den Nebenbedingungen affin-linear sind. Solche Probleme haben also die Form

$$\left. \begin{array}{l} \text{minimiere } \frac{1}{2} x^{\top} A x + b^{\top} x + c \\ \text{s.t. } H x = d, \quad G x \leq e \end{array} \right\} \quad (\text{QP})$$

<sup>1</sup>engl. mixed integer nonlinear program

mit Matrizen  $A \in \mathbb{R}^{n \times n}$ ,  $H \in \mathbb{R}^{p \times n}$ ,  $G \in \mathbb{R}^{m \times n}$ , Vektoren  $b \in \mathbb{R}^n$ ,  $d \in \mathbb{R}^p$ ,  $e \in \mathbb{R}^m$  und  $c \in \mathbb{R}$ .

- *Lineares Programm (LP)*:  
Bei solchen Aufgabenstellungen sind nicht nur die Nebenbedingungen affin-linear, sondern auch die Zielfunktion.
- *Gemischt-ganzzahliges lineares Programm (MILP)*<sup>2</sup>:  
Hierbei sollen einige der Optimierungsvariablen eines LPs zudem ganzzahlig sein.

### Beispiel 2.1 (Transport-Problem)

Ein Unternehmen hat  $m$  Fabriken  $F_1, \dots, F_m$ , die jeweils pro Woche  $a_i$  Einheiten,  $i = 1, \dots, m$ , eines bestimmten Produkts herstellen können. Verkauft wird dieses Produkt in  $r$  Filialen  $V_1, \dots, V_r$ . Jede Verkaufsstelle  $V_j$  hat eine wöchentlichen Nachfrage von  $b_j$  Einheiten. Die Kosten, um eine Einheit des Produkts von Fabrik  $F_i$  zur Filiale  $V_j$  zu transportieren, betragen  $c_{ij}$ . Wieviele Einheiten des Produkts müssen von Fabrik  $F_i$  zu Filiale  $V_j$  transportiert werden, so dass die Nachfrage gedeckt, die Produktionskapazitäten nicht überschritten und die Transportkosten minimal sind? Wir wollen die Zahl der Einheiten, die von  $F_i$  nach  $V_j$  transportiert werden, mit  $x_{ij}$  bezeichnen. Sie sind die Optimierungsvariablen. Das zur oben genannten Fragestellung gehörige Optimierungsproblem lautet dann

$$\begin{aligned} \text{minimiere} \quad & \sum_{i=1}^m \sum_{j=1}^r c_{ij} x_{ij} && \text{(Minimierung der Transportkosten)} \\ \text{s.t.} \quad & \sum_{j=1}^r x_{ij} \leq a_i, \quad i = 1, \dots, m && \text{(Kapazität von Fabrik } F_i \text{ einhalten)} \\ & \sum_{i=1}^m x_{ij} \geq b_j, \quad j = 1, \dots, r && \text{(Bedarf von Filiale } V_j \text{ decken)} \end{aligned}$$

Bei dieser Aufgabe handelt es sich um ein lineares Programm. ◇

### Beispiel 2.2 (Rucksack-Problem)

Ein Einbrecher kann nur eine limitierte Menge an Gegenständen aus einem Haus entwenden, da sein Rucksack nur begrenzte Kapazität hat. Es gibt  $n$  „interessante“ Gegenstände  $G_1, \dots, G_n$  im Haus mit einer Wertigkeit  $c_i$  und einem Volumen  $a_i$ ,  $i = 1, \dots, n$ . Das Volumen des Rucksacks ist  $b > 0$ . Will der Einbrecher den Wert seiner Beute maximieren, dann muss er folgendes Problem lösen:

$$\begin{aligned} \text{maximiere} \quad & c^\top x && \text{(Maximierung des Werts der Beute)} \\ \text{s.t.} \quad & a^\top x \leq b && \text{(Kapazität des Rucksacks einhalten)} \\ & x_i \in \{0, 1\}, \quad i = 1, \dots, n. \end{aligned}$$

Ist  $x_i = 1$ , dann packt der Einbrecher den Gegenstand  $G_i$  in seinen Rucksack, ist hingegen  $x_i = 0$ , dann verbleibt  $G_i$  im Haus. Man beachte, dass  $c^\top x$  und  $a^\top x$  dann genau dem Wert bzw. dem Volumen der Beute entsprechen. Bei diesem Problem handelt es sich um ein ganzzahliges lineares Programm. ◇

### Beispiel 2.3 (Strömungstechnische Anwendung)

Mit der Strömung eines Gases aus einem kugelförmigen Kessel soll eine Turbine

<sup>2</sup>engl. mixed integer linear program

angetrieben werden. Im Kessel liegt der Gasdruck  $p_K$  vor, die Turbine wird bei Umgebungsdruck  $p_0$  betrieben. Die Geschwindigkeit  $u$  des (idealen) Gases am Ausgang des Kessels berechnet sich nach der *Ausflußformel von Saint-Venant-Wantzel* durch

$$u = \sqrt{2 \frac{\gamma}{\gamma - 1} \frac{p_K}{\rho_K} \left( 1 - \left( \frac{p_0}{p_K} \right)^{(\gamma-1)/\gamma} \right)}. \quad (2.1)$$

Hierbei sind  $\rho_K$  die Dichte des Gases im Kessel und  $\gamma > 0$  der Isentropenexponent des Gases, ein konstanter Materialparameter. Die Dichte im Kessel ist durch

$$\rho_K = \frac{m}{V} = \frac{3}{4\pi} \frac{m}{r^3} \quad (2.2)$$

gegeben, wobei  $m$  die Masse an Gas im Kessel und  $V$  das Volumen und  $r$  den Radius des Kessels bezeichnen. Mit Hilfe der *idealen Gasgleichung* kann man näherungsweise den Kesseldruck bestimmen:

$$p_K = \rho_K R T_K = \frac{3}{4\pi} R T_K \frac{m}{r^3} \quad (2.3)$$

mit der universellen Gaskonstanten  $R$  und der Kesseltemperatur  $T_K$ , die wir ebenfalls als konstant annehmen. Setzt man (2.2) und (2.3) in (2.1) ein, erhält man eine nichtlineare Abbildung  $u : (m, r) \mapsto u(m, r)$ . Um die Turbine effizient betreiben zu können, darf eine Mindestgeschwindigkeit  $u_{\min}$  am Kesselausgang nicht unterschritten werden. Diese soll mit minimalem Aufwand erreicht werden, d.h. mit möglichst wenig Gas und möglichst kleinem Kessel. Das zugehörige Optimierungsproblem lautet dann:

$$\begin{aligned} \underset{m, r \in \mathbb{R}}{\text{minimiere}} \quad & \alpha 4\pi r^2 + \beta m \quad (\text{Minimierung von Kesseloberfläche und Gasmasse}) \\ \text{s.t.} \quad & u(m, r) \geq u_{\min} \quad (\text{Mindestgeschwindigkeit am Kesselausgang}) \\ & m \geq 0, r \geq 0 \quad (\text{Positivität von Masse und Radius}), \end{aligned}$$

wobei  $\alpha, \beta > 0$  Gewichtungsfaktoren sind. Wir erhalten also ein nichtlineares Programm.  $\diamond$

### § 3 Notation und Wiederholung

Ist  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  eine in  $x \in \mathbb{R}^n$  differenzierbare Funktion, dann bezeichnen wir mit

$$\nabla f(x) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}$$

den Gradienten von  $f$  an der Stelle  $x$ . Die Ableitung von  $f$  ist  $f'(x) := \nabla f(x)^\top$ . Ist  $f$  zweimal differenzierbar, dann bezeichnen wir mit

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{pmatrix} \in \mathbb{R}^{n \times n}$$



die Hessematrix von  $f$  in  $x$ . Ist  $f$  zweimal stetig differenzierbar, dann ist  $\nabla^2 f(x)$  bekanntlich symmetrisch. Ist  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  differenzierbar, dann bezeichnet

$$g'(x) = \begin{pmatrix} \nabla g_1(x)^\top \\ \vdots \\ \nabla g_m(x)^\top \end{pmatrix} \in \mathbb{R}^{m \times n}$$

die Jacobi-Matrix von  $g$ .

Mit  $\|\cdot\|$  bezeichnen wir die euklidische Norm eines Vektors, d.h. für  $a \in \mathbb{R}^n$

$$\|a\| = \sqrt{\sum_{i=1}^n a_i^2}.$$

Die Menge

$$B_\rho(a) := \{x \in \mathbb{R}^n : \|x - a\| < \rho\}$$

ist die offene Kugel um  $a$  mit Radius  $\rho > 0$ . Das Innere, den Abschluss und den Rand einer Menge  $M \subset \mathbb{R}^n$  bezeichnen wir mit  $\text{int}(M)$ ,  $\overline{M}$  und  $\partial M = \overline{M} \setminus \text{int}(M)$ . Wir verstehen Mengeneinklusiven der Form  $A \subset B$  immer als  $A \subseteq B$ .

Sind  $a, b \in \mathbb{R}^n$  beliebige Vektoren, so ist die Ungleichung  $a \leq b$  komponentenweise zu verstehen, d.h.  $a_i \leq b_i$  für alle  $i = 1, \dots, n$ . Entsprechend sind  $a \geq b$ ,  $a < b$  und  $a > b$  definiert. Mit  $\mathbb{R}_+^n$  bezeichnen wir die Menge

$$\mathbb{R}_+^n := \{x \in \mathbb{R}^n : x \geq 0\}.$$

Seien  $A \in \mathbb{R}^{m \times n}$  und  $b \in \mathbb{R}^m$  beliebig. Für  $i \in \{1, \dots, m\}$  und  $j \in \{1, \dots, n\}$  bezeichnen

$$A_{(i)} \in \mathbb{R}^n \quad \text{und} \quad A_{[j]} \in \mathbb{R}^m$$

den  $i$ -ten Zeilen- bzw.  $j$ -ten Spaltenvektor von  $A$ . Hierbei ist zu beachten, dass  $A_{(i)} \in \mathbb{R}^{n \times 1}$  (nicht  $\mathbb{R}^{1 \times n}$ !). Für beliebige Indexmengen  $\mathcal{I} \subset \{1, \dots, m\}$ ,  $\mathcal{J} \subset \{1, \dots, n\}$  bezeichnen wir mit

$$b_{\mathcal{J}} = (b_i)_{i \in \mathcal{J}} \in \mathbb{R}^{|\mathcal{J}|}, \quad A_{(\mathcal{I})} = (A_{(i)}^\top)_{i \in \mathcal{I}} \in \mathbb{R}^{|\mathcal{I}| \times n}, \quad A_{[\mathcal{J}]} = (A_{[j]})_{j \in \mathcal{J}} \in \mathbb{R}^{m \times |\mathcal{J}|}$$

die entsprechenden Teilvektoren bzw. Teilmatrizen. Man beachte, dass

$$(A_{(\mathcal{I})})^\top = (A^\top)_{[\mathcal{I}]} =: A_{[\mathcal{I}]}^\top \quad \text{und} \quad (A_{[\mathcal{J}]})^\top = (A^\top)_{(\mathcal{J})} =: A_{(\mathcal{J})}^\top. \quad (3.1)$$

Die Einheitsmatrix des  $\mathbb{R}^n$  bezeichnen wir mit  $I \in \mathbb{R}^{n \times n}$ . Der Vektor  $\mathbf{1} \in \mathbb{R}^n$  ist definiert durch  $\mathbf{1}_i = 1$  für alle  $i = 1, \dots, n$ .

Wir sagen im Folgenden, zwei Optimierungsprobleme sind äquivalent, falls ihre Lösungsmengen übereinstimmen, d.h. die Optimierungsprobleme

$$\left. \begin{array}{l} \text{minimiere } f_1(x) \\ \text{s.t. } x \in \mathcal{F}_1 \end{array} \right\} \quad \text{und} \quad \left\{ \begin{array}{l} \text{minimiere } f_2(x) \\ \text{s.t. } x \in \mathcal{F}_2 \end{array} \right.$$

sind äquivalent, falls

$$\arg \min_{x \in \mathcal{F}_1} f_1(x) = \arg \min_{x \in \mathcal{F}_2} f_2(x).$$

Gelegentlich nennen wir zwei Optimierungsprobleme auch dann äquivalent, falls man die Lösungsmenge des einen durch eine bijektive Transformation in die des anderen umwandeln kann.



Teil 1

# Konvexe Optimierung



## Konvexe Mengen und Funktionen

### Inhalt

---

4	<b>Konvexe Mengen</b>	13
5	<b>Konvexe Funktionen</b>	20
6	<b>Projektion auf konvexe Mengen</b>	23
7	<b>Trennungssätze</b>	24
8	<b>Farkas-Lemma</b>	27

---

In den folgenden Abschnitten betrachten wir konvexe Optimierungsprobleme. Derartige Aufgaben haben sehr angenehme Eigenschaften und sind i.A. wesentlich einfacher zu diskutieren als allgemeine nichtlineare oder auch diskrete Probleme.

### § 4 Konvexe Mengen

#### Definition 4.1

Eine Menge  $C \subset \mathbb{R}^n$  heißt *konvex*, falls

$$x, y \in C \quad \Rightarrow \quad \lambda x + (1 - \lambda)y \in C \quad \forall \lambda \in [0, 1],$$

das bedeutet die Verbindungsstrecke zweier Punkte aus  $C$  liegt ebenfalls in  $C$ . ◇

#### Beispiel 4.2

Beispiele für konvexe Mengen sind:

- Der ganze  $\mathbb{R}^n$  und jeder lineare Unterraum.
- Die Lösungsmenge  $\{x \in \mathbb{R}^n : Ax = b\}$  eines linearen Gleichungssystems mit Matrix  $A \in \mathbb{R}^{m \times n}$  und rechter Seite  $b \in \mathbb{R}^m$ .
- Die offene und die abgeschlossene Kugel  $B_r(a)$  und  $\overline{B_r(a)}$  (mit  $a \in \mathbb{R}^n$  und  $r > 0$  beliebig).
- Eine einpunktige Menge  $\{a\}$  (mit  $a \in \mathbb{R}^n$  beliebig). ◇

Weitere Beispiele für konvexe Mengen sind *Hyperebene* und *Halbraum*:

#### Definition 4.3

Seien  $a \in \mathbb{R}^n$ ,  $a \neq 0$ , und  $\alpha \in \mathbb{R}$  gegeben. Dann heißt die Menge

$$H(a, \alpha) = \{x \in \mathbb{R}^n : a^\top x = \alpha\}$$

von  $a$  und  $\alpha$  induzierte *Hyperebene*. Eine Hyperebene  $H(a, \alpha)$  erzeugt durch

$$H(a, \alpha)_\geq = \{x \in \mathbb{R}^n : a^\top x \geq \alpha\} \quad \text{und} \quad H(a, \alpha)_\leq = \{x \in \mathbb{R}^n : a^\top x \leq \alpha\}$$

zwei (abgeschlossene) *Halbräume*. ◇

**Definition 4.4**

Für beliebige  $M_1, M_2 \subset \mathbb{R}^n$  und  $\alpha \in \mathbb{R}$  definieren wir

$$\alpha M_1 := \{\alpha x : x \in M_1\} \quad \text{und} \quad M_1 + M_2 := \{x_1 + x_2 : x_1 \in M_1, x_2 \in M_2\} \quad \diamond$$

**Satz 4.5**

Folgende Operationen erhalten Konvexität:

- (1) Sei  $\{C_j\}_{j \in J}$  eine Familie konvexer Mengen. Dann ist der Durchschnitt  $\bigcap_{j \in J} C_j$  ebenfalls konvex.
- (2) Sind  $C_i \subset \mathbb{R}^{n_i}$ ,  $i = 1, \dots, m$ , konvex, dann auch das kartesische Produkt  $C_1 \times C_2 \times \dots \times C_m$ .
- (3) Ist  $C \subset \mathbb{R}^n$  konvex, dann auch  $\alpha C$  für jedes  $\alpha \in \mathbb{R}$ .
- (4) Sind  $C_1$  und  $C_2$  konvex, dann auch  $C_1 + C_2$ .
- (5) Abschluss und Inneres einer konvexen Menge sind ebenfalls konvex.

*Beweis:* (1)–(4): s. Übung.

zu (5):

Sei  $C \subset \mathbb{R}^n$  eine beliebige konvexe Menge. Um die Konvexität des Abschlusses zu sehen, seien  $x, y \in \overline{C}$  und  $\lambda \in [0, 1]$  beliebig. Dann gibt es Folgen  $\{x^k\}_{k \in \mathbb{N}} \subset C$  und  $\{y^k\}_{k \in \mathbb{N}} \subset C$  mit  $x^k \rightarrow x$  und  $y^k \rightarrow y$  für  $k \rightarrow \infty$ . Wegen Konvexität von  $C$  ist auch  $\lambda x^k + (1 - \lambda)y^k \in C$  für alle  $k$ . Daraus folgt, dass

$$\lambda x + (1 - \lambda)y = \lim_{k \rightarrow \infty} (\lambda x^k + (1 - \lambda)y^k) \in \overline{C},$$

was die Konvexität des Abschlusses zeigt.

Seien nun  $x, y \in \text{int}(C)$  und  $\lambda \in [0, 1)$  beliebig. Dann gibt es ein  $\delta > 0$ , so dass  $B_\delta(y) \subset C$ . Wir setzen  $z = \lambda x + (1 - \lambda)y$  und zeigen, dass  $B_{(1-\lambda)\delta}(z) \subset C$ , woraus  $z \in \text{int}(C)$  und somit die Konvexität von  $\text{int}(C)$  folgt. Dazu sei  $v \in B_{(1-\lambda)\delta}(z)$  beliebig. Wir definieren  $w := \frac{1}{1-\lambda}(v - \lambda x)$ . Dann folgt, dass

$$\|w - y\| = \frac{1}{1 - \lambda} \|v - \underbrace{(\lambda x + (1 - \lambda)y)}_{=z}\| \leq \delta \quad \Rightarrow \quad w \in B_\delta(y) \subset C.$$

Die Konvexität von  $C$  impliziert dann, dass  $v = \lambda x + (1 - \lambda)w \in C$ , und da  $v \in B_{(1-\lambda)\delta}(y)$  beliebig war, folgt die Behauptung.  $\square$

**Bemerkung 4.6**

Beim Beweis der Konvexität des Inneren einer konvexen Menge haben wir gar nicht ausgenutzt, dass  $x \in \text{int}(C)$ . Der Beweis zeigt also auch, dass für jedes  $x \in C$  und jedes  $y \in \text{int}(C)$  die Verbindungsstrecke zwischen  $x$  und  $z$  im Inneren von  $C$  liegt, d.h.

$$\{\lambda x + (1 - \lambda)y : 0 \leq \lambda < 1\} \subset \text{int}(C). \quad (4.1)$$

Mit einer leichten Modifikation des Beweises kann man zeigen, dass (4.1) sogar für alle  $x \in \overline{C}$  und  $y \in \text{int}(C)$  gilt, s. [?, Lemma 2.9].  $\diamond$

Konvexität einer Menge lässt sich auch mit folgendem Begriff charakterisieren:

**Definition 4.7 (Konvexkombination)**

Seien  $k \in \mathbb{N}$ ,  $x^1, \dots, x^k \in \mathbb{R}^n$  und  $\lambda_1, \dots, \lambda_k \in \mathbb{R}$  mit  $\lambda_i \geq 0$  und  $\sum_{i=1}^k \lambda_i = 1$  gegeben. Dann heißt

$$\sum_{i=1}^k \lambda_i x_i$$

Konvexkombination der Vektoren  $x^1, \dots, x^k$ .  $\diamond$

**Lemma 4.8**

Eine Menge  $C \subset \mathbb{R}^n$  ist genau dann konvex, wenn sie für jedes  $k \in \mathbb{N}$  alle Konvexkombinationen von Punkten  $x_1, \dots, x_k$  aus  $C$  enthält.  $\diamond$

*Beweis:* Enthält die Menge  $C$  jede Konvexkombination von  $k$  ihrer Punkte für beliebiges  $k \in \mathbb{N}$ , dann gilt dies natürlich auch für  $k = 2$ , was nach Definition 4.1 die Konvexität von  $C$  impliziert.

Die umgekehrte Inklusion beweisen wir mittels Induktion nach  $k$ . Sei dazu  $C$  eine konvexe Menge. Für  $k = 2$  ist die Aussage nach der Definition konvexer Mengen wahr. Als Induktionsvoraussetzung nehmen wir nun an, die Aussage sei für  $k \geq 2$  wahr. Seien  $x^1, \dots, x^{k+1} \in C$  und  $\lambda_1, \dots, \lambda_{k+1} \in \mathbb{R}$  mit

$$\lambda_i \geq 0, \quad i = 1, \dots, n, \quad \text{und} \quad \sum_{i=1}^{k+1} \lambda_i = 1$$

beliebig. Dann müssen wir zeigen, dass die Konvexkombination  $x := \sum_{i=1}^{k+1} \lambda_i x^i$  ebenfalls in  $C$  liegt. Ist  $\lambda_{k+1} = 1$ , dann folgt  $x = x^{k+1} \in C$ . Ist hingegen  $\lambda_{k+1} \neq 1$ , dann definieren wir

$$z := \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x_i.$$

Wegen

$$\frac{\lambda_i}{1 - \lambda_{k+1}} \geq 0 \quad \forall i = 1, \dots, k, \quad \text{und} \quad \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} = \frac{1}{1 - \lambda_{k+1}} \underbrace{\sum_{i=1}^k \lambda_i}_{=1 - \lambda_{k+1}} = 1$$

ist  $z$  eine Konvexkombination aus  $k$  Elementen aus  $C$ , so dass  $z \in C$  nach Induktionsvoraussetzung. Nach Konstruktion gilt

$$x = \lambda_{k+1} x_{k+1} + (1 - \lambda_{k+1}) z$$

und daher  $x \in C$  wegen Konvexität von  $C$ .  $\square$

**Definition 4.9**

Die *konvexe Hülle*  $\text{conv}(A)$  einer Menge  $A \subset \mathbb{R}^n$  ist der Durchschnitt aller konvexen Mengen, die  $A$  enthalten, d.h.

$$\text{conv}(A) = \bigcap \{C \subset \mathbb{R}^n : C \text{ ist konvex, } A \subset C\}. \quad (4.2)$$

**Lemma 4.10**

Die konvexe Hülle einer Menge  $A \subset \mathbb{R}^n$  ist die kleinste konvexe Menge, die  $A$  enthält.

*Beweis:* Da der Schnitt konvexer Mengen konvex ist, existiert eine kleinste konvexe Menge  $\mathcal{C} \subset \mathbb{R}^n$ , die  $A$  enthält. Die Definition von  $\text{conv}(A)$  impliziert, dass  $\text{conv}(A) \subset \mathcal{C}$ . Andererseits ist  $\text{conv}(A)$  nach Satz 4.5(1) konvex und enthält  $A$  nach Definition, was  $\mathcal{C} = \text{conv}(A)$  liefert.  $\square$

**Lemma 4.11**

Die konvexe Hülle  $\text{conv}(A)$  einer Menge  $A \subset \mathbb{R}^n$  ist gleich der Menge aller Konvexkombinationen von Punkten aus  $A$ .

*Beweis:* Wir bezeichnen die Menge aller Konvexkombinationen mit  $M$ . Sei  $x \in M$  beliebig. Dann existieren  $k \in \mathbb{N}$ ,  $x^1, \dots, x^k \in A$  und  $\lambda_1, \dots, \lambda_k \in \mathbb{R}_+$ , so dass

$$\sum_{i=1}^k \lambda_i = 1 \quad \text{und} \quad x = \sum_{i=1}^k \lambda_i x^i. \quad (4.3)$$

Wegen  $x_i \in A \subset \text{conv}(A)$  für alle  $i = 1, \dots, k$  und Lemma 4.8 folgt  $x \in \text{conv}(A)$ , also  $M \subset \text{conv}(A)$ .

Um die umgekehrte Inklusion zu beweisen, zeigen wir, dass  $M$  konvex ist. Seien  $y, z \in M$  beliebig, so dass  $m \in \mathbb{N}$ ,  $y^1, \dots, y^m \in A$ ,  $\alpha_1, \dots, \alpha_m \in \mathbb{R}_+$  und  $\ell \in \mathbb{N}$ ,  $z^1, \dots, z^\ell \in A$ ,  $\beta_1, \dots, \beta_\ell \in \mathbb{R}_+$  existieren mit

$$\sum_{i=1}^m \alpha_i = 1, \quad y = \sum_{i=1}^m \alpha_i y^i \quad \text{und} \quad \sum_{i=1}^{\ell} \beta_i = 1, \quad z = \sum_{i=1}^{\ell} \beta_i z^i.$$

Sei  $\lambda \in [0, 1]$  beliebig. Dann folgt

$$w := \lambda y + (1 - \lambda)z = \lambda \sum_{i=1}^m \alpha_i y^i + (1 - \lambda) \sum_{j=1}^{\ell} \beta_j z^j$$

mit  $\lambda \sum_{i=1}^m \alpha_i + (1 - \lambda) \sum_{j=1}^{\ell} \beta_j = \lambda + (1 - \lambda) = 1,$

d.h.  $w$  ist eine Konvexkombination von Punkten aus  $A$  und somit  $w \in M$ , was die Konvexität von  $M$  zeigt. Außerdem ist  $A \subset M$  (da sich jeder Punkt aus  $A$  trivialerweise als einelementige Konvexkombination darstellen lässt). Wäre also  $\text{conv}(A) \not\subset M$ , dann wäre  $\text{conv}(A) \cap M$  eine kleinere konvexe Menge als  $\text{conv}(A)$ , die  $A$  enthält, im Widerspruch zu Lemma 4.10.  $\square$

#### Bemerkung 4.12

Aus den Lemmata 4.8 und 4.11 folgt sofort, dass für konvexe Mengen  $C \subset \mathbb{R}^n$  gilt, dass  $\text{conv}(C) = C$ .  $\diamond$

Man kann die Aussage von Lemma 4.11 noch wie folgt verschärfen:

#### Satz 4.13 (Carathéodory)

Die konvexe Hülle einer Menge  $A \subset \mathbb{R}^n$  ist gleich der Menge aller Konvexkombinationen aus  $(n + 1)$ -elementigen Teilmengen von  $A$ , d.h.

$$\text{conv}(A) = \left\{ x \in \mathbb{R}^n : \exists x^1, \dots, x^{n+1} \in A, \lambda_1, \dots, \lambda_{n+1} \in \mathbb{R}_+ \right.$$

$$\left. \text{mit } \sum_{i=1}^{n+1} \lambda_i = 1, x = \sum_{i=1}^{n+1} \lambda_i x^i \right\}. \quad (4.4)$$

*Beweis:* Wir bezeichnen die Menge auf der rechten Seite von (4.4) mit  $M$ . Zunächst gilt wegen Lemma 4.11, dass  $M \subset \text{conv}(A)$ .

Um die umgekehrte Inklusion zu zeigen, sei  $x \in \text{conv}(A)$  beliebig. Wiederum nach Lemma 4.11 existieren dann  $m \in \mathbb{N}$ ,  $x^1, \dots, x^m$  und  $\lambda_1, \dots, \lambda_m$ , so dass

$$\sum_{i=1}^m \lambda_i = 1 \quad \text{und} \quad x = \sum_{i=1}^m \lambda_i x^i,$$



wobei  $m$  minimal gewählt sei. Insbesondere ist also  $\lambda > 0$ .

Wir nehmen nun an, dass  $m > n + 1$  gilt, und führen dieses zum Widerspruch. Die Menge

$$\left\{ \begin{pmatrix} x^1 \\ 1 \end{pmatrix}, \dots, \begin{pmatrix} x^m \\ 1 \end{pmatrix} \right\} \subseteq \mathbb{R}^{n+1}$$

ist dann linear abhängig, es existiert also ein  $\gamma \in \mathbb{R}^m \setminus \{0\}$  mit  $\sum_{i=1}^m \gamma_i \begin{pmatrix} x^i \\ 1 \end{pmatrix} = 0$ . Die Betrachtung der letzten Zeile liefert  $\sum_{i=1}^m \gamma_i = 0$ , und wegen  $\gamma \neq 0$  existiert dann ein Index  $i$  mit  $\gamma_i > 0$ . Damit ist

$$\tau := \min \left\{ \frac{\lambda_i}{\gamma_i} \mid \gamma_i > 0 \right\} = \frac{\lambda_l}{\gamma_l}$$

wohldefiniert. Setze schließlich  $\bar{\lambda} := \lambda - \tau\gamma$ . Dann gilt nach Konstruktion  $\bar{\lambda} \geq 0$  und

$$\sum_{i=1}^m \bar{\lambda}_i = \underbrace{\sum_{i=1}^m \lambda_i}_{=1} - \tau \underbrace{\sum_{i=1}^m \gamma_i}_{=0} = 1$$

sowie

$$\sum_{i=1}^m \bar{\lambda}_i x^i = \underbrace{\sum_{i=1}^m \lambda_i x^i}_{=x} - \tau \underbrace{\sum_{i=1}^m \gamma_i x^i}_{=0} = x.$$

Wegen  $\bar{\lambda}_l = \lambda_l - \tau\gamma_l = 0$  ist das ein Widerspruch zur Minimalität von  $m$ .  $\square$

#### Lemma 4.14

Sei  $C \subset \mathbb{R}^n$  eine konvexe Menge mit nichtleerem Inneren. Dann gilt:

- (1)  $\text{int}(C) = \text{int}(\bar{C})$
- (2)  $\partial C = \partial \bar{C}$ ,

d.h. bei konvexen Mengen stimmen Inneres und Rand der Menge mit Innerem und Rand ihres Abschlusses überein.  $\diamond$

#### Bemerkung 4.15

Die scheinbar triviale Aussage von Lemma 4.14 stimmt keineswegs immer. Betrachtet man z.B. die Menge  $M = ([0, 1] \cap \mathbb{Q}) \cup [1, 2]$ , dann ist  $\text{int}(M) = (1, 2) \neq \text{int}(\bar{M}) = (0, 2)$  und  $\partial M = [0, 1] \cup \{2\} \neq \partial \bar{M} = \{0, 2\}$ .  $\diamond$

*Beweis von Lemma 4.14:* zu (1):

Sei  $x \in \text{int}(C)$  beliebig. Dann existiert ein  $\delta > 0$  mit  $B_\delta(x) \subset C \subset \bar{C}$ , was  $x \in \text{int}(\bar{C})$  und somit  $\text{int}(C) \subset \text{int}(\bar{C})$  beweist. Um die umgekehrte Inklusion zu zeigen, nehmen wir an, es gäbe  $x \in \text{int}(\bar{C}) \setminus \text{int}(C)$ . Dann gibt es ein  $\rho > 0$ , so dass  $B_\rho(x) \subset \bar{C}$ . Nach Voraussetzung existiert ein  $y \in \text{int}(C)$ . Definiere dann

$$z := x + \alpha(x - y) \quad \text{mit} \quad \alpha := \frac{\rho}{2\|x - y\|}, \quad (4.5)$$

so dass  $\|z - x\| \leq \rho/2$  und damit  $z \in B_\rho(x) \subset \bar{C}$ . Für  $x$  ergibt sich damit

$$x = \frac{1}{1 + \alpha} z + \left(1 - \frac{1}{1 + \alpha}\right) y, \quad (4.6)$$

so dass  $x$  als Konvexkombination aus  $z \in \bar{C}$  und  $y \in \text{int}(C)$  dargestellt werden kann, woraus nach Bemerkung 4.6  $x \in \text{int}(C)$  folgt.

zu (2):

Folgt wegen

$$\partial \bar{C} = \bar{C} \setminus \text{int}(\bar{C}) = \bar{C} \setminus \text{int}(\bar{C}) \stackrel{(1)}{=} \bar{C} \setminus \text{int}(C) = \partial C \quad (4.7)$$

sofort aus der ersten Aussage des Lemmas.  $\square$

Neben konvexen Mengen spielen Kegel in der konvexen Analysis eine wichtige Rolle:

**Definition 4.16**

Eine nichtleere Menge  $K \subset \mathbb{R}^n$  heißt *Kegel*, falls

$$x \in K \quad \Rightarrow \quad \alpha x \in K$$

für alle  $\alpha \geq 0$  gilt.  $\diamond$

Man beachte, dass mit dieser Definition immer  $0 \in K$  gilt.

**Beispiel 4.17**

Beispiele für Kegel sind

- $K_0 = \mathbb{R}_+^n$
- $K_1 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0\}$
- $K_2 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = x_2\}$
- $K_3 = \{0\}$

$\diamond$

**Lemma 4.18**

Ein Kegel  $K \subset \mathbb{R}^n$  ist genau dann konvex, falls  $K + K = K$ .

*Beweis:* Sei  $K$  ein konvexer Kegel und  $x, y \in K$  beliebig. Dann folgt

$$\frac{1}{2}(x + y) = \underbrace{\frac{1}{2}x}_{\in K} + \underbrace{\frac{1}{2}y}_{\in K} \in K \quad \Rightarrow \quad x + y \in K.$$

Dies zeigt  $K + K \subset K$ . Außerdem gilt für beliebige  $x \in K$ , dass

$$x = \frac{1}{2}x + \frac{1}{2}x \in K + K,$$

so dass  $K = K + K$ .

Sei nun  $K$  ein Kegel mit  $K = K + K$ . Seien weiter  $x, y \in K$  und  $\lambda \in [0, 1]$  beliebig. Dann folgt

$$\underbrace{\lambda x}_{\in K} + \underbrace{(1 - \lambda)y}_{\in K} \in K, \quad (4.8)$$

also Konvexität von  $K$ .  $\square$

**Definition 4.19 (Konische Kombination)**

Seien  $k \in \mathbb{N}$ ,  $x^1, \dots, x^k \in \mathbb{R}^n$  und  $\lambda_1, \dots, \lambda_k \in \mathbb{R}$  mit  $\lambda_i \geq 0$  gegeben. Dann heißt

$$\sum_{i=1}^k \lambda_i x_i$$

*konische Kombination* der Vektoren  $x^1, \dots, x^k$ .  $\diamond$

Man beachte, dass man bei der konischen Kombination im Gegensatz zur Konvexkombination nicht  $\sum_i \lambda_i = 1$  fordert. Analog zu Lemma 4.8 gilt:

**Lemma 4.20**

Eine nichtleere Menge  $K \subset \mathbb{R}^n$  ist genau dann ein konvexer Kegel, falls sie für jedes  $k \in \mathbb{N}$  alle konischen Kombinationen von Punkten  $x^1, \dots, x^k$  aus  $K$  enthält.

*Beweis:* Enthält  $K$  für jedes  $k \in \mathbb{N}$  alle  $k$ -elementigen konischen Kombinationen, dann natürlich auch für  $k = 1$ , was die Kegeleigenschaft zeigt. Da jede Konvexkombination natürlich auch eine konische Kombination ist, enthält  $K$  alle Konvexkombinationen von Punkten aus  $K$ , so dass Lemma 4.8 die Konvexität von  $K$  liefert.

Sei umgekehrt  $K$  ein konvexer Kegel und

$$x = \sum_{i=1}^k \alpha_i x^i \quad \text{mit} \quad \alpha_i \geq 0, x^i \in K, i = 1, \dots, k$$

eine beliebige konische Kombination. Ist  $\sum_i \alpha_i = 0$ , dann folgt  $x = 0$  und somit  $x \in K$ . Im Fall  $\sum_i \alpha_i \neq 0$  ist

$$\tilde{x} := \frac{1}{\sum_{i=1}^k \alpha_i} x = \sum_{i=1}^k \frac{\alpha_i}{\sum_{i=1}^k \alpha_i} x^i \quad (4.9)$$

eine Konvexkombination von Punkten aus  $K$ . Nach Lemma 4.8 liefert die Konvexität von  $K$  daher  $\tilde{x} \in K$  und die Kegeleigenschaft somit  $x = (\sum_i \alpha_i) \tilde{x} \in K$ .  $\square$

**Definition 4.21 (Konische Hülle)**

Sei  $A \subset \mathbb{R}^n$  beliebig. Die Menge

$$\text{cone}(A) = \bigcap \{K \subset \mathbb{R}^n : K \text{ konvexer Kegel, } A \subset K\}. \quad (4.10)$$

heißt *konische Hülle* von  $A$ .  $\diamond$

Analog zu Satz 4.5 zeigt man:

**Lemma 4.22**

Der Durchschnitt beliebig vieler konvexer Kegel ist ebenfalls ein konvexer Kegel.  $\diamond$

Daraus folgt, dass auch  $\text{cone}(A)$  ein konvexer Kegel ist. Wegen  $A \subset \text{cone}(A)$  nach Definition der konischen Hülle zeigt man damit völlig analog zu Lemma 4.10:

**Lemma 4.23**

Die konische Hülle einer Menge  $A \subset \mathbb{R}^n$  ist der kleinste konvexe Kegel, der  $A$  enthält.  $\diamond$

**Lemma 4.24**

Sei  $C \subset \mathbb{R}^n$  konvex. Dann gilt

$$\text{cone}(C) = \{\alpha x : \alpha \geq 0, x \in C\}. \quad (4.11)$$

*Beweis:* Wir bezeichnen die Menge auf der rechten Seite von (4.11) mit  $K$  und zeigen zunächst, dass  $K$  ein konvexer Kegel ist. Dass  $K$  ein Kegel ist, sieht man sofort. Zur Konvexität: Seien  $y^1, y^2 \in K$  und  $\lambda \in (0, 1)$  beliebig. Dann existieren  $x^1, x^2 \in C$  und  $\alpha_1, \alpha_2 \geq 0$ , so dass  $y^i = \alpha_i x^i$ ,  $i = 1, 2$ . Sind  $\alpha_1 = \alpha_2 = 0$ , dann ist  $\lambda y^1 + (1 - \lambda) y^2 = 0 \in K$ . Ist hingegen  $\alpha_1 + \alpha_2 > 0$ , dann definieren wir  $\beta := \lambda \alpha_1 + (1 - \lambda) \alpha_2$ . Es folgt, dass  $\beta > 0$  und  $\lambda \alpha_1 / \beta \leq 1$  und somit

$$z := \frac{1}{\beta} (\lambda y^1 + (1 - \lambda) y^2) = \frac{\lambda \alpha_1}{\beta} x^1 + \left(1 - \frac{\lambda \alpha_1}{\beta}\right) x^2 \in C$$

wegen Konvexität von  $C$ . Nach Definition von  $K$  folgt daraus aber gerade  $\lambda y^1 + (1 - \lambda)y^2 = \beta z \in K$ , was die Konvexität von  $K$  zeigt.

$K$  ist also ein konvexer Kegel. Die Definition von  $K$  liefert zudem sofort  $C \subset K$ , so dass  $\text{cone}(C) \subset K$  aus Lemma 4.23 folgt.

Um die umgekehrte Inklusion zu zeigen, sei  $y \in K$  beliebig. Dann existieren  $x \in C$  und  $\alpha \geq 0$ , so dass  $y = \alpha x$ . Wegen  $C \subset \text{cone}(C)$  folgt  $x \in \text{cone}(C)$ , und die Kegeleigenschaft impliziert  $y \in \text{cone}(C)$ . Daraus folgt  $K \subset \text{cone}(C)$ .  $\square$

### Lemma 4.25

Die konische Hülle einer Menge  $A \subset \mathbb{R}^n$  ist gleich der Menge aller konischen Kombinationen von Punkten aus  $A$ .

*Beweis:* Der Beweis ähnelt sehr stark dem von Lemma 4.11. Wir bezeichnen die Menge aller konischen Kombinationen mit  $Z$ . Sei  $x \in Z$  beliebig, d.h.

$$x = \sum_{i=1}^k \alpha_i x^i \quad \text{mit} \quad \alpha_i \geq 0, \quad x^i \in A, \quad i = 1, \dots, k$$

eine beliebige konische Kombination von Punkten aus  $A$ . Wegen  $x^i \in A \subset \text{cone}(A)$  ist  $x$  auch eine konische Kombination von Punkten aus  $\text{cone}(A)$ , so dass  $x \in \text{cone}(A)$  nach Lemma 4.20 und somit  $Z \subset \text{cone}(A)$ .

Für die Gleichheit der Mengen zeigen wir, dass  $Z$  ein konvexer Kegel ist. Ist  $x \in Z$  eine beliebige konische Kombination, dann ist auch  $\alpha x$  für beliebiges  $\alpha \geq 0$  eine konische Kombination, also  $\alpha x \in Z$ , was die Kegeleigenschaft beweist. Ganz ähnlich folgt auch die Konvexität von  $Z$ : sind  $x, y \in Z$  beliebige konische Kombinationen und  $\lambda \in [0, 1]$  beliebig, dann ist auch die Konvexkombination  $z := \lambda x + (1 - \lambda)y$  eine konische Kombination, so dass  $z \in Z$ , was die Konvexität zeigt. Außerdem ist trivialerweise jedes  $x \in A$  auch eine konische Kombination, so dass  $A \subset Z$ . Da  $\text{cone}(A)$  nach Lemma 4.23 aber der kleinste konvexe Kegel ist, der  $A$  enthält, folgt  $\text{cone}(A) = Z$ .  $\square$

## § 5 Konvexe Funktionen

### Definition 5.1

Es sei  $C \subset \mathbb{R}^n$  eine konvexe Menge.

- (1) Eine Funktion  $f : C \rightarrow \mathbb{R}$  heißt *konvex*, falls

$$f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y)$$

für alle  $\lambda \in [0, 1]$  und alle  $x, y \in C$ .

- (2) Eine Funktion  $f : C \rightarrow \mathbb{R}$  heißt *streng konvex*, falls

$$f((1 - \lambda)x + \lambda y) < (1 - \lambda)f(x) + \lambda f(y)$$

für alle  $\lambda \in (0, 1)$  und alle  $x, y \in C$  mit  $x \neq y$ .  $\diamond$

Anschaulich bedeutet die obige Definition, dass sich die Sekante zwischen  $x$  und  $y$  oberhalb des Graphs verläuft. Ein Beispiel für eine konvexe Funktion ist in Abbildung ?? gezeigt. Konvexität einer Funktion lässt sich auch über den so genannten *Epigraph* charakterisieren:

**Definition 5.2 (Epigraph)**

Sei  $C \subset \mathbb{R}^n$  eine gegebene Menge. Der *Epigraph* einer Funktion  $f : C \rightarrow \mathbb{R}$  ist definiert durch

$$\text{epi}(f) := \{(x, r) \in C \times \mathbb{R} : f(x) \leq r\}.$$

Man beachte, dass  $\text{epi}(f) \subset \mathbb{R}^{n+1}$ .  $\diamond$

Der Epigraph ist gerade der Bereich oberhalb des Graphs einer Funktion, s. Abb. ???. Aus bestimmten Eigenschaften des Epigraphs kann man Rückschlüsse auf die Funktion selber ziehen, beispielsweise gilt:

**Lemma 5.3**

Sei  $C \subset \mathbb{R}^n$  konvex. Eine Funktion  $f : C \rightarrow \mathbb{R}$  ist genau dann konvex, wenn ihr Epigraph  $\text{epi}(f)$  konvex ist.

*Beweis:* Sei  $f : C \rightarrow \mathbb{R}$  konvex. Dann gilt für beliebige  $(x_1, r_1), (x_2, r_2) \in \text{epi}(f)$ , dass

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \leq \lambda r_1 + (1 - \lambda)r_2 \quad \forall \lambda \in [0, 1],$$

woraus die Konvexität von  $\text{epi}(f)$  folgt. Sei andersrum der Epigraph einer beliebigen Funktion konvex, so dass  $(\lambda x_1 + (1 - \lambda)x_2, \lambda r_1 + (1 - \lambda)r_2) \in \text{epi}(f)$  für beliebige  $(x_1, r_1), (x_2, r_2) \in \text{epi}(f) \subset C \times \mathbb{R}$  und beliebige  $\lambda \in [0, 1]$ . Setzt man  $r_1 = f(x_1)$  und  $r_2 = f(x_2)$  ergibt sich daraus sofort die Konvexität von  $f$  auf  $C$ .  $\square$

Alternativ kann man konvexe Funktionen auch mit der Tangente statt der Sekante charakterisieren. Anhand des Beispiels in Abbildung ??? erkennt man, dass die Tangente immer unterhalb des Graphs der konvexen Funktion liegt. In der Tat gilt

**Satz 5.4**

Seien  $C \subset \mathbb{R}^n$  konvex,  $D \subset \mathbb{R}^n$  mit  $C \subset D$  offen und  $f : D \rightarrow \mathbb{R}$  differenzierbar. Dann gilt:

(1) Die Funktion  $f$  ist genau dann konvex auf  $C$ , wenn

$$f(y) - f(x) \geq \nabla f(x)^\top (y - x) \quad \forall x, y \in C. \quad (5.1)$$

(2) Die Funktion  $f$  ist genau dann streng konvex auf  $C$ , wenn

$$f(y) - f(x) > \nabla f(x)^\top (y - x) \quad \forall x, y \in C \text{ mit } x \neq y. \quad (5.2)$$

*Beweis:* zu (1):

Sei zunächst  $f : C \rightarrow \mathbb{R}$  konvex. Dann gilt nach Definition 5.1 für beliebige  $x, y \in C$  und alle  $0 < t \leq 1$

$$f(x + t(y - x)) - f(x) \leq tf(y) + (1 - t)f(x) - f(x) = t(f(y) - f(x)).$$

Da  $f$  diffbar ist, folgt daraus

$$f(y) - f(x) \geq \lim_{t \searrow 0} \frac{f(x + t(y - x)) - f(x)}{t} = \nabla f(x)^\top (y - x)$$

und damit der erste Teil von (1).

Um die Rückrichtung zu zeigen, gelte (5.1) und  $x, y \in C$  und  $0 \leq \lambda \leq 1$  seien beliebig. Wir definieren  $z := (1 - \lambda)x + \lambda y$ . Dann ist Konvexität von  $f$  äquivalent zu

$$(1 - \lambda)f(x) + \lambda f(y) - f(z) \geq 0.$$

Um dies nachzuweisen, berechnen wir unter Benutzung von (5.1)

$$\begin{aligned} (1 - \lambda)f(x) + \lambda f(y) - f(z) &= (1 - \lambda)(f(x) - f(z)) + \lambda(f(y) - f(z)) \\ &\geq (1 - \lambda)\nabla f(z)^\top(x - z) + \lambda\nabla f(z)^\top(y - z) \quad (5.3) \\ &= \nabla f(z)^\top((1 - \lambda)x + \lambda y - z) = 0 \end{aligned}$$

und damit Konvexität von  $f$ .

zu (2):

Sei  $f$  streng konvex und  $x, y \in C$ ,  $x \neq y$ , beliebig. Aus (1) folgt dann

$$f(1/2 x + 1/2 y) - f(x) \geq \frac{1}{2} \nabla f(x)^\top(y - x). \quad (5.4)$$

Die strenge Konvexität von  $f$  impliziert außerdem

$$\begin{aligned} f(1/2 x + 1/2 y) - f(x) &= f((1 - 1/2)x + 1/2 y) - f(x) \\ &< (1 - 1/2)f(x) + 1/2 f(y) - f(x) = \frac{1}{2}(f(y) - f(x)), \end{aligned}$$

was zusammen mit (5.4) die gewünschte Aussage ergibt.

Der Nachweis der Rückrichtung in (2) folgt genau wie bei Teil (1), indem man in (5.3) die strikte Ungleichung für  $x \neq y$  und  $0 < \lambda < 1$  verwendet.  $\square$

Schließlich untersuchen wir den Zusammenhang zwischen Krümmungsverhalten, d.h. der zweiten Ableitung, und Konvexität:

### Satz 5.5

Es sei  $C \subset \mathbb{R}^n$  eine offene konvexe Menge und  $f : C \rightarrow \mathbb{R}$  sei zweimal stetig differenzierbar. Dann gilt:

- (1) Die Funktion  $f$  ist genau dann konvex auf  $C$ , wenn die Hessematrix  $\nabla^2 f$  auf  $C$  positiv semidefinit ist.
- (2) Ist  $\nabla^2 f$  auf  $C$  positiv definit, dann ist  $f$  streng konvex auf  $C$ .

*Beweis:* zu (1):

Seien  $f$  konvex auf  $C$  und  $x \in C$  und  $s \in \mathbb{R}^n$  beliebig. Da  $C$  offen ist, gilt  $x + ts \in C$  für  $t > 0$  hinreichend klein. Aus Satz 5.4(1) folgt dann zusammen mit dem Mittelwertsatz, dass

$$t\nabla f(x)^\top s \leq f(x + ts) - f(x) = t\nabla f(x)^\top s + \frac{t^2}{2} s^\top \nabla^2 f(x + \tau_t s)s.$$

mit einem  $\tau_t \in [0, t]$ . Daraus folgt  $s^\top \nabla^2 f(x + \tau_t s)s \geq 0$  und, da  $\nabla^2 f$  nach Voraussetzung stetig ist, können wir zur Grenze  $t \searrow 0$  übergehen und erhalten

$$s^\top \nabla^2 f(x)s = \lim_{t \searrow 0} s^\top \nabla^2 f(x + \tau_t s)s \geq 0.$$

Da  $x \in C$  und  $s \in \mathbb{R}^n$  beliebig waren, folgt daraus die positive Semidefinitheit von  $\nabla^2 f$  auf  $C$ .

Um die Rückrichtung zu zeigen, sei  $\nabla^2 f$  nun positiv semidefinit auf  $C$  und  $x, y \in C$  beliebig. Wir definieren  $s := y - x$ . Wegen Konvexität von  $C$  ist dann  $x + ts = (1 - t)x + ty \in C$  für jedes  $t \in [0, 1]$  und damit  $s^\top \nabla^2 f(x + ts)s \geq 0$  für alle  $t \in [0, 1]$ . Der Mittelwertsatzes liefert die Existenz eines  $t \in [0, 1]$ , so dass

$$f(y) - f(x) = \nabla f(x)^\top s + \frac{1}{2} s^\top \nabla^2 f(x + ts)s \geq \nabla f(x)^\top s = \nabla f(x)^\top(y - x)$$

Nach Satz 5.4(1) folgt daraus die Konvexität von  $f$  auf  $C$ .

zu (2):

Ist  $\nabla^2 f$  positiv definit auf  $C$ , dann ergibt sich für  $x, y \in C$ ,  $x \neq y$  genau wie eben

$$f(y) - f(x) = \nabla f(x)^\top s + \frac{1}{2} s^\top \nabla^2 f(x + ts) s > \nabla f(x)^\top s = \nabla f(x)^\top (y - x)$$

mit einem  $t \in [0, 1]$ . Nach Satz 5.4(2) folgt die strenge Konvexität von  $f$ .  $\square$

### Bemerkung 5.6

- Für (2) und die Rückrichtung in (1) benötigt man die Offenheit der Menge  $C$  eigentlich nicht. Man könnte die beiden Aussagen also auch für zweimal stetig diffbare Funktionen  $f : D \rightarrow \mathbb{R}$  mit einer offenen Obermenge  $D \subset \mathbb{R}^n$  von  $C$  beweisen.
- Die Rückrichtung in (2) gilt i.A. *nicht*, wie man sich leicht am Beispiel der streng konvexen Funktion  $f(x) = x^4$  im Punkt  $x = 0$  überlegt.  $\diamond$

Der Begriff der strengen Konvexität lässt sich wie folgt verschärfen:

### Definition 5.7

Es sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  zweimal stetig differenzierbar auf einer konvexen Menge  $C \subset \mathbb{R}^n$ . Dann heißt  $f$  *gleichmäßig konvex*, wenn  $\nabla^2 f$  auf  $C$  gleichmäßig positiv definit ist, es also  $\mu > 0$  gibt mit

$$s^\top \nabla^2 f(x) s \geq \mu \|s\|^2 \quad \forall s \in \mathbb{R}^n, x \in C. \quad \diamond$$

Nach Satz 5.5(2) ist eine gleichmäßig konvexe Funktion automatisch streng konvex. Die Umkehrung gilt aber nicht, wie wir in Bemerkung 5.6 gesehen haben.

## § 6 Projektion auf konvexe Mengen

Wir betrachten nun eine erste konvexe Optimierungsaufgabe, die so genannte *Projektion* auf eine konvexe Menge. Dazu sei eine nichtleere, abgeschlossene und konvexe Menge  $C \subset \mathbb{R}^n$  und ein Punkt  $y \in \mathbb{R}^n$  gegeben. Dann ist die Projektion von  $y$  auf  $C$  als Lösung der folgenden Aufgabe definiert:

$$\left. \begin{array}{l} \text{minimiere } f(x) := \frac{1}{2} \|x - y\|^2 \\ \text{s.t. } x \in C, \end{array} \right\} \quad (6.1)$$

d.h. die Lösung  $x$  dieses Problems, ist genau der Punkt aus  $C$  mit minimalem Abstand zu  $y$ . Der folgende Satz zeigt, dass die Projektion überhaupt wohl definiert ist.

### Satz 6.1

Sei  $C$  nichtleer, abgeschlossen und konvex und  $y \in \mathbb{R}^n$  gegeben. Dann existiert eine eindeutige Lösung von (6.1), die wir mit  $\Pi_C(y)$  bezeichnen.

*Beweis:* Nach Voraussetzung existiert ein Punkt  $z \in C$ . Dieser habe von  $y$  den Abstand  $r = \|z - y\|$ . Außerhalb der Kugel  $\overline{B_r}(y)$  kann die Lösung von (6.1) natürlich nicht liegen, weshalb wir uns bei der Suche nach dem globalen Minimierer von (6.1) auf die Menge  $C_r = C \cap \overline{B_r}(y)$  beschränken können. Als Durchschnitt zweier abgeschlossener Mengen ist  $C_r$  selber abgeschlossen. Außerdem ist  $C_r$  nichtleer (wegen

$z \in C_r$ ) und natürlich beschränkt, folglich kompakt. Nach dem bekannten Satz von Weierstraß nimmt einer stetige Funktion auf kompakten Mengen ihr Maximum und Minimum an, so dass ein Minimierer von (6.1) existiert.

Zur Eindeutigkeit des Minimierers: wir nehmen an, die Aussage sei falsch und es gäbe zwei (globale) Minimierer  $x_1 \in C$  und  $x_2 \in C$ ,  $x_1 \neq x_2$ . Für die zweite Ableitung von  $f$  finden wir  $\nabla^2 f(x) = I$  für alle  $x \in \mathbb{R}^n$ , so dass  $f$  nach Satz 5.5(2) auf ganz  $\mathbb{R}^n$  streng konvex ist. Daraus folgt

$$f((1/2)x_1 + (1/2)x_2) < \frac{1}{2}f(x_1) + \frac{1}{2}f(x_2). \quad (6.2)$$

Wegen  $(1/2)x_1 + (1/2)x_2 \in C$  ist dies aber ein Widerspruch zur Optimalität von  $x_1$  und  $x_2$ .  $\square$

### Satz 6.2 (Projektionssatz)

Die Projektion wird eindeutig charakterisiert durch die *Variationsungleichung*

$$(\Pi_C(y) - y)^\top (x - \Pi_C(y)) \geq 0 \quad \forall x \in C, \quad (6.3)$$

d.h. jede Lösung von (6.1) löst auch (6.3) und umgekehrt.

*Beweis:* Es ist zu zeigen, dass ein  $x^* \in C$  genau dann optimal ist für (6.1), wenn für alle  $x \in C$  die Ungleichung  $(x^* - y)^\top (x - x^*) \geq 0$  gilt. Wir benutzen dazu jeweils  $\nabla f(x) = x - y$ .

„ $\Leftarrow$ “: Da  $f$  konvex ist, folgt mit Satz 5.4 (1)

$$0 \leq (x^* - y)^\top (x - x^*) = \nabla f(x^*)^\top (x - x^*) \leq f(x) - f(x^*),$$

also  $f(x^*) \leq f(x)$  für alle  $x \in C$ . Somit ist  $x^*$  optimal.

„ $\Rightarrow$ “: Sei  $x \in C$  beliebig. Für  $t \in [0, 1]$  gilt dann  $x^* + t(x - x^*) \in C$ , da  $C$  konvex ist. Aus der Optimalität von  $x^*$  folgt

$$f(x^* + t(x - x^*)) \geq f(x^*)$$

und somit

$$\begin{aligned} (x^* - y)^\top (x - x^*) &= \nabla f(x^*)^\top (x - x^*) \\ &= \lim_{t \searrow 0} \frac{f(x^* + t(x - x^*)) - f(x^*)}{t} \geq 0. \end{aligned}$$

$\square$

Anschaulich bedeutet die Aussage von Satz 6.2, dass der Winkel zwischen dem Lot von  $y$  auf die Menge  $C$  und Verbindungslinien zwischen Lotfußpunkt und beliebigen Punkten aus  $C$  immer stumpf sein muss, vgl. Abbildung ???. Satz 6.2 ist ein Spezialfall der Sätze 9.2 und 9.4 weiter unten.

## § 7 Trennungssätze

### Definition 7.1 (Trennende Hyperebene)

Eine Hyperebene  $H(a, \alpha) = \{x \in \mathbb{R}^n : a^\top x = \alpha\}$ ,  $a \neq 0$ , trennt zwei Mengen  $A, B \subset \mathbb{R}^n$ , falls  $A \subset H(a, \alpha)_\leq$  und  $B \subset H(a, \alpha)_\geq$ , wobei  $H(a, \alpha)_\leq$  und  $H(a, \alpha)_\geq$  wieder die durch die Hyperebene erzeugten Halbräume bezeichnen. Die Hyperebene trennt die beiden Mengen also genau dann, wenn

$$a^\top x \leq a^\top y \quad \forall x \in A, y \in B. \quad (7.1)$$



Man bezeichnet die Trennung als *strikt*, wenn  $A$  und  $B$  jeweils im Inneren der beiden Halbräume liegen, d.h. wenn

$$a^\top x < \alpha < a^\top y \quad \forall x \in A, y \in B \quad (7.2)$$

gilt.  $\diamond$

Beispiele für trennende Hyperebenen sind in Abbildung ?? zu sehen. Wir beginnen zunächst mit zwei Hilfsresultaten zur Trennung von konvexen Mengen und Punkten.

**Lemma 7.2**

Seien  $C \subset \mathbb{R}^n$ , nichtleer, abgeschlossen und konvex und  $x \notin C$ . Dann gibt es eine Hyperebene, die  $C$  und  $\{x\}$  strikt trennt, d.h.

$$\exists s \in \mathbb{R}^n, s \neq 0, \alpha \in \mathbb{R} \quad \text{mit} \quad s^\top x > \alpha > \sup_{y \in C} s^\top y.$$

*Beweis:* Nach den Sätzen 6.1 und 6.2 existiert eine eindeutige Projektion  $\Pi_C(x)$ , die (6.3) erfüllt. Wegen  $\Pi_C(x) \in C$  und  $x \notin C$ , ist  $s := x - \Pi_C(x) \neq 0$ . Außerdem erfüllt  $s$  wegen (6.3) für alle  $y \in C$ , dass

$$\begin{aligned} 0 &\geq s^\top (y - x + s) = s^\top y - s^\top x + \|s\|^2 \\ \Rightarrow \quad s^\top y &\leq s^\top x - \|s\|^2 < s^\top x - \frac{\|s\|^2}{2} < s^\top x, \end{aligned}$$

was der Aussage mit  $\alpha = s^\top x - \|s\|^2/2$  entspricht.  $\square$

**Korollar 7.3**

Eine abgeschlossene Menge ist konvex genau dann, wenn sie Durchschnitt abgeschlossener Halbräume ist.

Im Fall nicht abgeschlossener Mengen muss man auf die Striktheit der Trennung verzichten:

**Lemma 7.4**

Seien eine konvexe Menge  $C \subset \mathbb{R}^n$ ,  $C \neq \mathbb{R}^n$ , mit nichtleerem Inneren und ein Punkt  $x \notin \text{int}(C)$  gegeben. Dann existiert eine Hyperebene, die  $C$  und  $\{x\}$  trennt, d.h.

$$\exists s \in \mathbb{R}^n, s \neq 0 \quad \text{mit} \quad s^\top x \geq \sup_{y \in C} s^\top y.$$

*Beweis:* Für  $x \notin \overline{C}$  folgt die Aussage sofort aus Lemma 7.2. Sei deshalb  $x \in \overline{C} \setminus \text{int}(C) = \partial C$ . Da  $C$  nach Voraussetzung konvex und  $\text{int}(C) \neq \emptyset$  ist, gilt nach Lemma 4.14, dass  $\partial C = \partial \overline{C}$  und damit  $x \in \partial \overline{C}$ . Wegen  $C \neq \mathbb{R}^n$ , ist  $\overline{C} \neq \mathbb{R}^n$ . Daher gibt es eine Folge  $\{x^k\}$  mit

$$x^k \notin \overline{C} \quad \forall k \in \mathbb{N} \quad \text{und} \quad x^k \rightarrow x \quad \text{für} \quad k \rightarrow \infty.$$

Wegen  $x^k \notin \overline{C}$  gilt der Trennungssatz nach dem im ersten Teil bereits Bewiesenen für  $x^k$ , d.h. für jedes  $k \in \mathbb{N}$  existiert ein  $s^k$ ,  $s^k \neq 0$  mit

$$(s^k)^\top x^k > (s^k)^\top y \quad \forall y \in C.$$

Teilt man diese Ungleichung durch  $\|s^k\| \neq 0$ , so erhält man mit  $\sigma^k := s^k / \|s^k\|$

$$(\sigma^k)^\top x^k > (\sigma^k)^\top y \quad \forall y \in C \quad (7.3)$$

für alle  $k \in \mathbb{N}$ . Auf Grund von  $\|\sigma^k\| = 1$  für alle  $k \in \mathbb{N}$  gibt es eine konvergente Teilfolge  $\{\sigma^{k_j}\}$  und ein  $s \in \mathbb{R}^n$ , so dass

$$\sigma^{k_j} \rightarrow s \text{ für } j \rightarrow \infty, \quad \|s\| = 1.$$

Mit  $x^k \rightarrow x$  können wir daher in (7.3) mit der Teilfolge zur Grenze übergehen und erhalten

$$s^\top x \geq s^\top y \quad \forall y \in C.$$

Wegen  $\|s\| = 1$  ist  $s \neq 0$ , was die Aussage beweist.  $\square$

### Bemerkung 7.5

Die Aussage von Lemma 7.4 stimmt übrigens auch, wenn man auf die Voraussetzung  $\text{int}(C) \neq \emptyset$  verzichtet. Ist nämlich  $\text{int}(C) = \emptyset$ , dann kann man zeigen, dass  $C$  in einer Hyperebene liegen muss, womit sich die Trennungseigenschaft zeigen lässt, s. [?, Lemma 2.21].

### Satz 7.6 (Trennungssatz)

Seien  $C_1, C_2 \subset \mathbb{R}^n$  zwei nichtleere, konvexe Mengen mit  $C_1 \cap C_2 = \emptyset$ . Dann existiert ein  $s \in \mathbb{R}^n$ ,  $s \neq 0$ , mit

$$\sup_{x_1 \in C_1} s^\top x_1 \leq \inf_{x_2 \in C_2} s^\top x_2, \quad (7.4)$$

d.h. es existiert eine Hyperebene, die  $C_1$  und  $C_2$  trennt.

*Beweis:* Wir definieren die Menge

$$C := C_1 - C_2 = \{x_1 - x_2 : x_1 \in C_1, x_2 \in C_2\}.$$

Man sieht leicht, dass  $C$  wegen der Voraussetzungen an  $C_1$  und  $C_2$  nichtleer ist und konvex nach Satz 4.5. Außerdem ist  $0 \notin C$ , denn sonst gäbe es  $x_1 \in C_1$  und  $x_2 \in C_2$  mit  $x_1 = x_2$  im Widerspruch zu  $C_1 \cap C_2 = \emptyset$ . Damit ist insbesondere  $0 \notin \text{int}(C)$ , so dass nach Lemma 7.4 und Bemerkung 7.5 ein  $s \in \mathbb{R}^n$  existiert mit

$$0 = s^\top 0 \geq s^\top x \quad \forall x \in C. \quad (7.5)$$

Seien nun  $x_1 \in C_1$  und  $x_2 \in C_2$  beliebig. Dann ist  $x = x_1 - x_2 \in C$ , so dass aus (7.5) gerade (7.4) folgt.  $\square$

Um die strikte Trennung von zwei Mengen beweisen zu können, benötigen wir das folgende

### Lemma 7.7

Seien  $M_1 \subset \mathbb{R}^n$  abgeschlossen und  $M_2 \subset \mathbb{R}^n$  kompakt. Dann ist  $M_1 - M_2$  ebenfalls abgeschlossen.

*Beweis:* Wir definieren  $M := M_1 - M_2$ . Sei  $\{x^k\} \subset M$  eine konvergente Folge mit Grenzwert  $x \in \mathbb{R}^n$ . Wir müssen  $x \in M$  zeigen. Wegen  $x^k \in M$  existieren  $x_1^k \in M_1$  und  $x_2^k \in M_2$  mit  $x^k = x_1^k - x_2^k$  für alle  $k \in \mathbb{N}$ . Wegen der Kompaktheit von  $M_2$  existiert eine in  $M_2$  konvergente Teilfolge von  $\{x_2^k\}$ , die wir mit  $\{x_2^{k_\ell}\}$  bezeichnen, d.h.  $x_2^{k_\ell} \rightarrow x_2 \in M_2$  für  $\ell \rightarrow \infty$ . Daraus folgt

$$x_1^{k_\ell} = x^{k_\ell} + x_2^{k_\ell} \rightarrow x + x_2 =: x_1 \quad \text{für } \ell \rightarrow \infty.$$

Wegen Abgeschlossenheit von  $M_1$  ist  $x_1 \in M_1$ , was schließlich  $x = x_1 - x_2 \in M_1 - M_2 = M$  zeigt.  $\square$

**Satz 7.8 (Strikte Trennung)**

Seien  $C_1, C_2 \subset \mathbb{R}^n$  zwei nichtleere, konvexe, abgeschlossene und disjunkte Mengen. Darüber hinaus sei  $C_2$  kompakt. Dann existiert eine Hyperebene  $H(s, \alpha)$ , die  $C_1$  und  $C_2$  strikt trennt, d.h.

$$\exists s \in \mathbb{R}^n, s \neq 0, \alpha \in \mathbb{R} \quad \text{mit} \quad \sup_{x_1 \in C_1} s^\top x_1 < \alpha < \inf_{x_2 \in C_2} s^\top x_2. \quad (7.6)$$

*Beweis:* Der Beweis verläuft wie der von Satz 7.6. Wir betrachten die Menge  $C := C_1 - C_2$ . Nach Lemma 7.7 ist  $C$  abgeschlossen. Zudem ist  $C$  trivialerweise nichtleer und nach Satz 4.5 konvex. Wie im Beweis von Satz 7.6 folgt zudem  $0 \notin C$ , so dass sich  $\{0\}$  und  $C$  nach Lemma 7.2 durch eine Hyperebene  $H(s, \beta)$  strikt trennen lassen, d.h.

$$s^\top x < \beta < s^\top 0 = 0 \quad \forall x \in C. \quad (7.7)$$

Seien nun  $x_1 \in C_1$  und  $x_2 \in C_2$  beliebig. Dann ist  $x_1 - x_2 \in C$  und aus (7.7) folgt

$$s^\top x_1 < s^\top x_2 + \beta \quad \forall x_1 \in C_1, x_2 \in C_2.$$

Wegen  $\beta < 0$  nach (7.7) folgt daraus

$$\sup_{x_1 \in C_1} s^\top x_1 \leq \inf_{x_2 \in C_2} s^\top x_2 + \beta < \underbrace{\inf_{x_2 \in C_2} s^\top x_2 + \frac{\beta}{2}}_{=: \alpha} < \inf_{x_2 \in C_2} s^\top x_2$$

und damit die Aussage. Man beachte, dass  $\alpha > -\infty$ , da  $C_2$  kompakt ist.  $\square$

**§ 8 Farkas-Lemma**

Wir kommen nun zu einem wichtigen Hilfsmittel der endlich-dimensionalen Optimierung, dem so genannten *Farkas-Lemma*, auch *Alternativ-Satz von Farkas* genannt. Dazu beweisen wir zunächst folgendes

**Lemma 8.1**

Die konische Hülle von  $m \in \mathbb{N}$  Vektoren  $a^1, \dots, a^m \in \mathbb{R}^n$  ist ein abgeschlossener, konvexer Kegel.

*Beweis:* Dass  $\text{cone}\{a^1, \dots, a^m\}$  ein konvexer Kegel ist, folgt aus Lemma 4.23. Die Abgeschlossenheit von  $\text{cone}\{a^1, \dots, a^m\}$  ist vergleichsweise aufwendig zu zeigen. Wir beweisen sie mittels Induktion nach  $m$ . Zunächst gilt nach Lemma 4.20, dass

$$\text{cone}\{a^1, \dots, a^m\} = \left\{ \sum_{i=1}^m \lambda_i a^i : \lambda_i \geq 0, i = 1, \dots, m \right\}.$$

Für  $m = 1$  ist die konische Hülle also durch  $\text{cone}\{a^1\} = \{\lambda_1 a^1 : \lambda_1 \geq 0\}$  gegeben und daher eine Halbgerade. Diese ist abgeschlossen, wie man sehr einfach sieht: Dazu sei  $\lambda_k a_1 \in \text{cone}\{a^1\}$  eine konvergente Folge mit Grenzwert  $x \in \mathbb{R}^n$ . Es gilt dann  $x = \lambda a_1$  mit  $\lambda := \lim_{k \rightarrow \infty} \lambda^k$  und  $\lambda \geq 0$  wegen  $\lambda_k \geq 0$ .

Sei nun  $m \in \mathbb{N}$ ,  $m > 1$ , beliebig. Als Induktionsvoraussetzung nehmen wir an, die Aussage gelte für  $m-1$ , d.h. die konische Hülle von  $m-1$  Punkten sei abgeschlossen. Wir betrachten eine Folge  $\{x^k\} \subset \text{cone}\{a^1, \dots, a^m\}$ , die gegen  $x \in \mathbb{R}^n$  konvergiert, und müssen  $x \in \text{cone}\{a^1, \dots, a^m\}$  zeigen. Wegen  $x^k \in \text{cone}\{a^1, \dots, a^m\}$  existieren Folgen  $\{\lambda_{k,i}\} \subset \mathbb{R}_+$  mit

$$x^k = \sum_{i=1}^m \lambda_{k,i} a^i. \quad (8.1)$$

Die Menge

$$U = \left\{ \sum_{i=1}^m \alpha_i a^i : \alpha_i \in \mathbb{R} \right\}$$

ist als endlichdimensionaler Unterraum abgeschlossen. Demnach ist  $x \in U$ , d.h. es existieren  $\alpha_1, \dots, \alpha_m$  mit

$$x = \sum_{i=1}^m \alpha_i a^i. \quad (8.2)$$

Sind alle  $\alpha_i \geq 0$ , dann ist  $x \in \text{cone}\{a^1, \dots, a^m\}$ , und wir sind fertig. Andernfalls definieren wir für jedes  $k \in \mathbb{N}$

$$\beta_k := \min_{1 \leq i \leq m} \left\{ \frac{\lambda_{k,i}}{\lambda_{k,i} - \alpha_i} : \alpha_i < 0 \right\}. \quad (8.3)$$

Durch Übergang auf eine Teilfolge und Umnummerieren der Vektoren  $a^1, \dots, a^m$  können wir o.B.d.A. annehmen, dass das Minimum in (8.3) immer im Index  $m$  angenommen wird. Wegen  $\lambda_{k,i} \geq 0$  für alle  $k \in \mathbb{N}$  und  $i \in \{1, \dots, m\}$  und der Restriktion auf die Indizes mit  $\alpha_i < 0$  folgt  $\beta_k \in [0, 1]$ . Deshalb ist

$$z^k := \beta_k x + (1 - \beta_k) x^k$$

eine Konvexkombination aus  $x$  und  $x^k$ , und wegen  $\lim_{k \rightarrow \infty} x^k = x$  folgt daraus wiederum

$$z^k \rightarrow x \quad \text{für } k \rightarrow \infty. \quad (8.4)$$

Die Definition von  $z^k$  liefert zusammen mit (8.1) und (8.2)

$$z^k = \beta_k \sum_{i=1}^m \alpha_i a^i + (1 - \beta_k) \sum_{i=1}^m \lambda_{k,i} a^i = \sum_{i=1}^m \underbrace{(\beta_k \alpha_i + (1 - \beta_k) \lambda_{k,i})}_{=: \gamma_{k,i}} a^i.$$

Ist  $\alpha_i \geq 0$ , dann folgt wegen  $\beta_k \in [0, 1]$ , dass

$$\beta_k(\alpha_i - \lambda_{k,i}) \geq -\beta_k \lambda_{k,i} \geq -\lambda_{k,i}.$$

Ist hingegen  $\alpha_i < 0$ , dann ergibt sich

$$\beta_k(\alpha_i - \lambda_{k,i}) = \min_{\alpha_j < 0} \left\{ \frac{\lambda_{k,j}}{\lambda_{k,j} - \alpha_j} \right\} \underbrace{(\alpha_i - \lambda_{k,i})}_{< 0} \geq \frac{\lambda_{k,i}}{\lambda_{k,i} - \alpha_i} (\alpha_i - \lambda_{k,i}) = -\lambda_{k,i},$$

so dass in jedem Fall  $\gamma_{k,i} \geq 0$  für alle  $k \in \mathbb{N}$  und alle  $i = 1, \dots, m$  folgt. Außerdem folgt

$$\beta_k(\alpha_m - \lambda_{k,m}) = \min_{\alpha_j < 0} \left\{ \frac{\lambda_{k,j}}{\lambda_{k,j} - \alpha_j} \right\} (\alpha_m - \lambda_{k,m}) = \frac{\lambda_{k,m}}{\lambda_{k,m} - \alpha_m} (\alpha_m - \lambda_{k,m}) = -\lambda_{k,m},$$

also  $\gamma_{k,m} = 0$  und damit

$$z^k = \sum_{i=1}^{m-1} \gamma_{k,i} a^i,$$

und damit  $z^k \in \text{cone}\{a^1, \dots, a^{m-1}\}$  wegen  $\gamma_{k,i} \geq 0$ . Da  $\text{cone}\{a^1, \dots, a^{m-1}\}$  nach Induktionsvoraussetzung abgeschlossen ist, liefert (8.4), dass  $x \in \text{cone}\{a^1, \dots, a^{m-1}\} \subset \text{cone}\{a^1, \dots, a^m\}$ , was die Aussage beweist.  $\square$

Man beachte die Parallelen zum Beweis des Satzes von Carathéodory (Satz 4.13).

**Lemma 8.2**

Seien  $A \in \mathbb{R}^{m \times n}$  und  $B \in \mathbb{R}^{p \times n}$  gegebene Matrizen. Dann ist die Menge

$$K := \{x \in \mathbb{R}^n : x = A^\top u + B^\top v, v \in \mathbb{R}^p, u \geq 0\} \quad (8.5)$$

ein abgeschlossener, konvexer Kegel.

*Beweis:* Dass  $K$  ein konvexer Kegel ist, rechnet man sehr leicht nach. Die Abgeschlossenheit von  $K$  zeigen wir mit Lemma 8.1. Da sich jedes  $x \in K$  auch als

$$x = A^\top u + B^\top v = A^\top u + B^\top v^{(+)} - B^\top v^{(-)}$$

mit  $v^{(+)} := (\max\{0, v_i\})_{i=1}^p$  und  $v^{(-)} := (\max\{0, -v_i\})_{i=1}^p$  schreiben lässt, gilt

$$\begin{aligned} K &= \{x \in \mathbb{R}^n : x = (A^\top, B^\top, -B^\top)w, w \in \mathbb{R}^{m+2p}, w \geq 0\} \\ &= \text{cone}\{A_{[1]}^\top, \dots, A_{[m]}^\top, \dots, B_{[1]}^\top, \dots, B_{[p]}^\top, -B_{[1]}^\top, \dots, -B_{[p]}^\top\}, \end{aligned}$$

und die Abgeschlossenheit von  $K$  folgt in der Tat aus Lemma 8.1.  $\square$

**Satz 8.3 (Farkas-Lemma)**

Seien  $A \in \mathbb{R}^{m \times n}$ ,  $B \in \mathbb{R}^{p \times n}$  und  $b \in \mathbb{R}^n$  beliebig. Dann sind folgende Aussagen äquivalent:

- (1) Es gibt ein  $u \in \mathbb{R}^m$ ,  $u \geq 0$ , und ein  $v \in \mathbb{R}^p$ , so dass  $A^\top u + B^\top v = b$ .
- (2) Für alle  $d \in \mathbb{R}^n$  mit  $A d \leq 0$  und  $B d = 0$  gilt  $b^\top d \leq 0$ .

*Beweis:* Um die Implikation (1)  $\Rightarrow$  (2) zu zeigen, gelte  $A^\top u + B^\top v = b$  mit einem  $u \geq 0$ . Dann folgt für alle  $d \in \mathbb{R}^n$  mit  $A d \leq 0$  und  $B d = 0$ , dass

$$b^\top d = u^\top A d + v^\top B d \leq 0,$$

was gerade (2) entspricht.

Für die umgekehrte Richtung argumentieren wir per Kontraposition, d.h. wir zeigen  $\neg(1) \Rightarrow \neg(2)$ . Ist (1) falsch, dann ist  $b \notin K$  mit  $K$  wie in (8.5) definiert. Da  $K$  nach Lemma 8.2 eine abgeschlossene und konvexe Menge ist, können  $\{b\}$  und  $K$  nach Lemma 7.2 deshalb durch eine Hyperebene strikt getrennt werden. (Man beachte, dass  $K$  wegen  $0 \in K$  nichtleer ist.) Das bedeutet, es existiert ein  $s \in \mathbb{R}^n$ ,  $s \neq 0$ , und ein  $\alpha \in \mathbb{R}$  mit

$$s^\top b > \alpha > s^\top x \quad \forall x \in K.$$

Wegen  $0 \in K$  folgt daraus sofort

$$s^\top b > 0. \quad (8.6)$$

Da  $K$  außerdem ein Kegel ist, gilt für alle  $x \in K$ , dass

$$s^\top x < \frac{\alpha}{t} \quad \forall t \geq 0 \quad \Rightarrow \quad s^\top x \leq \lim_{t \rightarrow \infty} \frac{\alpha}{t} = 0 \quad \forall x \in K.$$

Seien nun  $u \in \mathbb{R}^m$ ,  $u \geq 0$  und  $v \in \mathbb{R}^p$  beliebig. Dann ist  $A^\top u + B^\top v \in K$  und daher

$$0 \geq s^\top (A^\top u + B^\top v) = u^\top A s + v^\top B s. \quad (8.7)$$

Wählen wir  $u = 0$  und  $v = B s$ , dann folgt  $\|B s\|^2 \leq 0$  also  $B s = 0$ . Wählen wir  $v = 0$  und  $u = e_i$ ,  $i = 1, \dots, m$ , wobei  $e_i$  der  $i$ -te Einheitsvektor ist, dann ergibt sich  $A s \leq 0$ . Daraus folgt insgesamt, dass  $s$

$$A s \leq 0, \quad B s = 0 \quad \text{und} \quad b^\top s > 0$$

erfüllt, so dass (2) nicht gilt, was die Aussage beweist.  $\square$

**Bemerkung 8.4**

Man kann das Farkas-Lemma auch äquivalent als *Alternativsatz* formulieren:

Seien  $A \in \mathbb{R}^{m \times n}$ ,  $B \in \mathbb{R}^{p \times n}$  und  $b \in \mathbb{R}^n$  beliebig. Dann ist entweder das System

$$A^\top u + B^\top v = b, \quad u \geq 0$$

lösbar, oder (exklusiv) das System

$$A d \leq 0, \quad B d = 0, \quad b^\top d > 0$$

hat eine Lösung. ◇

Aus Satz 8.3 ergibt sich mit  $p = 0$  der folgende Spezialfall:

**Korollar 8.5**

Für  $A \in \mathbb{R}^{m \times n}$  und  $b \in \mathbb{R}^n$  sind folgende Aussagen äquivalent:

(1) Für alle  $d \in \mathbb{R}^n$  mit  $A d \leq 0$  gilt  $b^\top d \leq 0$ .

(2) Es gibt ein  $u \in \mathbb{R}^m$ ,  $u \geq 0$ , so dass  $A^\top u = b$ . ◇

Es gibt auch eine „inhomogene“ Version dieses Korollars:

**Satz 8.6**

Seien  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^n$ ,  $c \in \mathbb{R}^m$  und  $\beta \in \mathbb{R}$  gegeben. Dann sind folgende Aussagen äquivalent:

(1) Für alle  $d \in \mathbb{R}^n$  mit  $A d \leq c$  gilt  $b^\top d \leq \beta$ .

(2) Es gibt ein  $u \in \mathbb{R}^m$ ,  $u \geq 0$ , so dass

(a)  $A^\top u = b$  und  $c^\top u \leq \beta$

(b) oder  $A^\top u = 0$  und  $c^\top u < 0$

(wobei das „oder“ inklusiv gemeint ist, wie der Fall  $\beta < 0$  und  $b = 0$  zeigt.)

*Beweis:* Aussage (1) ist äquivalent zur Unlösbarkeit des Systems

$$d \in \mathbb{R}^n, \quad A d \leq c, \quad b^\top d > \beta,$$

also auch zur Unlösbarkeit des homogenen Systems

$$t \in \mathbb{R}, \quad d \in \mathbb{R}^n, \quad t > 0, \quad A d - t c \leq 0, \quad b^\top d - t \beta > 0,$$

und damit zur Unlösbarkeit von

$$\varepsilon \in \mathbb{R}, \quad t \in \mathbb{R}, \quad d \in \mathbb{R}^n, \quad \varepsilon - t \leq 0, \quad A d - t c \leq 0, \quad \beta t + \varepsilon - b^\top d \leq 0, \quad \varepsilon > 0.$$

Wir wenden nun Korollar 8.5 an auf

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & -c & A \\ 1 & \beta & -b^\top \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

damit ist obige Unlösbarkeit äquivalent zu

$$\exists y, v, w \geq 0: \quad A^\top y = b v, \quad c^\top y + w = \beta v, \quad v + w = 1. \quad (8.8)$$

Zu zeigen bleibt also, dass (8.8) äquivalent zu (2) ist.

Gelte zunächst (8.8). Falls  $v = 0$ , folgt  $w = 1$ , setze dann  $u := y$ . Damit gilt (2) (b). Andernfalls setze  $u := \frac{1}{v} y$ , woraus (2) (a) folgt.

Gelte nun umgekehrt (2). Im Fall (a) gilt (8.8) mit  $v := \frac{1}{1 + \beta - c^\top u}$ ,  $w := 1 - v$ ,  $y := v u$ , im Fall (b) mit  $y := -\frac{1}{c^\top u} u$ ,  $v := 0$ ,  $w := 1$ . □

## Konvexe Optimierungsprobleme

### Inhalt

9	<b>Variationsungleichung</b>	31
10	<b>KKT-Bedingungen und Lagrange-Dualität</b>	32
11	<b>Subgradientenverfahren</b>	44

### § 9 Variationsungleichung

Wir betrachten nun Optimierungsaufgaben, bei denen die zulässige Menge konvex ist, d.h.

$$\left. \begin{array}{l} \text{minimiere } f(x) \\ \text{s.t. } x \in C, \end{array} \right\} \quad (9.1)$$

mit einer konvexen Menge  $C \subset \mathbb{R}^n$ . Ist zusätzlich auch die Zielfunktion konvex auf  $C$ , so spricht man von einem *konvexen Optimierungsproblem*. Derartige Aufgaben haben sehr vorteilhafte Eigenschaften, beispielsweise muss man nicht mehr zwischen globalen und lokalen Minimierern unterscheiden:

**Satz 9.1**

Seien  $C \subset \mathbb{R}^n$  und  $f : C \rightarrow \mathbb{R}$  konvex. Dann gilt:

- Jeder lokale Minimierer von  $f$  auf  $C$  ist auch globaler Minimierer. Die Lösungsmenge von (9.1), also die Menge aller Minimierer von  $f$  auf  $C$ , ist konvex.
- Ist  $f$  streng konvex auf  $C$ , existiert höchstens ein (globaler) Minimierer.

*Beweis:* s. Übung. □

Im Fall einer konvexen Nebenbedingung  $x \in C$  kann man sehr einfache Optimalitätsbedingungen in Form einer *Variationsungleichung* formulieren.

**Satz 9.2**

Seien  $C \subset \mathbb{R}^n$  konvex und  $x^*$  ein lokaler Minimierer von (9.1). Ist  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  differenzierbar in  $x^*$ , dann gilt die folgende *Variationsungleichung*:

$$\nabla f(x^*)^\top (x - x^*) \geq 0 \quad \forall x \in C. \quad (\text{VI})$$

*Beweis:* Die Umgebung der lokalen Optimalität von  $x^*$  sei mit  $B_r(x^*)$  bezeichnet. Sei  $x \in C$  beliebig. Da  $C$  konvex ist, liegt die Konvexkombination  $x^* + t(x - x^*)$ ,  $t \in [0, 1]$ , in  $C$  und für hinreichend kleine  $t \in [0, 1]$  in  $B_r(x^*)$ . Daraus folgt

$$\frac{f(x^* + t(x - x^*)) - f(x^*)}{t} \geq 0$$

für hinreichend kleine  $t > 0$ . Da  $f$  differenzierbar in  $x^*$  ist, können wir zur Grenze übergehen  $t \searrow 0$  übergehen und erhalten (VI).  $\square$

### Bemerkung 9.3

Der Beweis von Satz 9.2 zeigt, dass man die Differenzierbarkeit von  $f$  auf ganz  $\mathbb{R}^n$  eigentlich gar nicht braucht. Es reicht zu fordern, dass  $f$  in  $x^*$  richtungsdifferenzierbar in alle Richtungen  $x - x^*$  mit  $x \in C$  ist.  $\diamond$

Die Konvexität von  $f$  wurde in Satz 9.2 *nicht* gefordert. Wenn auch noch  $f$  konvex ist, ist die Variationsungleichung nicht nur notwendig, sondern auch *hinreichend* für Optimalität:

### Satz 9.4

Sei  $C \subset \mathbb{R}^n$  konvex und  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  differenzierbar und konvex auf  $C$ . Dann ist jede Lösung der Variationsungleichung (VI) ein globaler Minimierer von (9.1).

*Beweis:* Sei  $x^* \in C$  eine Lösung von (VI) und  $x \in C$  beliebig. Nach Satz 5.4(1) gilt

$$f(x) - f(x^*) \geq \nabla f(x^*)^\top (x - x^*) \geq 0,$$

wobei wir (VI) benutzt haben, was nach Voraussetzung erfüllt ist. Da  $x \in C$  beliebig war, folgt daraus die globale Optimalität von  $x^*$ .  $\square$

### Bemerkung 9.5

Wie bereits oben angedeutet ist Satz 6.2 ein Spezialfall der Sätze 9.2 und 9.4. Hierbei ist zu beachten, dass  $f(x) = 1/2 \|x - y\|^2$  natürlich konvex und differenzierbar für alle  $x \in \mathbb{R}^n$  ist.  $\diamond$

## § 10 KKT-Bedingungen und Lagrange-Dualität

Die Variationsungleichung (VI) ist für eine numerische Bestimmung des Optimums nur in bestimmten Fällen brauchbar. Viele Algorithmen basieren hingegen auf Optimalitätsbedingungen, die mit Hilfe von *Lagrange-Multiplikatoren*, auch *duale Variablen* genannt, formuliert werden. Wir werden im Folgenden die Existenz von Lagrange-Multiplikatoren für folgendes Problem beweisen:

$$\left. \begin{array}{l} \text{minimiere } f(x) \\ \text{s.t. } x \in C, \quad \Phi(x) \leq 0. \end{array} \right\} \quad (\text{CP})$$

Hierbei seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  und  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^s$  konvex und  $C \subset \mathbb{R}^n$  eine konvexe Menge.

### Definition 10.1

Ein Punkt  $\hat{x} \in C$  heißt *Slater-Punkt*, falls

$$\Phi(\hat{x}) < 0$$

gilt. Aufgabe (CP) erfüllt die *Slater-Bedingung*, falls ein Slater-Punkt existiert.

### Lemma 10.2

Die Slater-Bedingung sei mit einem Slater-Punkt  $\hat{x} \in C$  erfüllt. Gilt

$$f^* := \inf_{\substack{x \in C \\ \Phi(x) \leq 0}} f(x) > -\infty, \quad (10.1)$$

dann existiert ein  $\lambda \in \mathbb{R}^s$ ,  $\lambda \geq 0$ , mit

$$f^* = \inf_{x \in C} (f(x) + \Phi(x)^\top \lambda). \quad (10.2)$$



Wird das Infimum in (10.1) von einem  $x^* \in C$  angenommen (d.h.  $x^*$  löst (CP)), dann gilt zudem

$$f(x^*) + \Phi(x^*)^\top \lambda = \inf_{x \in C} (f(x) + \Phi(x)^\top \lambda) \quad (10.3a)$$

$$\Phi(x^*) \leq 0, \quad \Phi(x^*)^\top \lambda = 0, \quad \lambda \geq 0. \quad (10.3b)$$

*Beweis:* Der Beweis erfolgt in zwei Schritten. Zunächst konstruieren wir eine Hyperebene  $(\alpha, \lambda) \in \mathbb{R} \times \mathbb{R}^s$ , die die Mengen

$$A := \{(r, y) \in \mathbb{R} \times \mathbb{R}^s : \exists x \in C \text{ mit } r > f(x), y \geq \Phi(x)\}$$

$$B := \{(r, y) \in \mathbb{R} \times \mathbb{R}^s : r \leq f^*, y \leq 0\}.$$

trennt. Danach zeigen wir mit Hilfe der Slater-Bedingung, dass  $\alpha > 0$ , was (10.2) impliziert.

Zunächst ist wegen der Slater-Bedingung  $C \neq \emptyset$ , so dass auch  $A \neq \emptyset$ . Wegen  $(f^*, 0) \in B$  ist auch  $B$  nichtleer. Wegen (10.1) ist zudem  $A \cap B = \emptyset$ . Außerdem sind  $A$  und  $B$  konvex. Im Fall von  $B$  sieht man das sofort. Sind hingegen  $(r_i, y_i)$ ,  $i = 1, 2$ , zwei Elemente von  $A$ , so dass  $x_1, x_2 \in C$  mit  $r_i > f(x_i)$  und  $y_i \geq \Phi(x_i)$  gilt, dann folgt aus der Konvexität von  $f$ ,  $g$  und  $C$ , dass

$$tx_1 + (1-t)x_2 \in C,$$

$$tr_1 + (1-t)r_2 > tf(x_1) + (1-t)f(x_2) \geq f(tx_1 + (1-t)x_2),$$

$$ty_1 + (1-t)y_2 \geq t\Phi(x_1) + (1-t)\Phi(x_2) \geq \Phi(tx_1 + (1-t)x_2)$$

und damit  $(tr_1 + (1-t)r_2, ty_1 + (1-t)y_2) \in A$  für alle  $t \in [0, 1]$ , also Konvexität von  $A$ . Daher lassen sich  $A$  und  $B$  nach dem Trennungssatz 7.6 durch eine Hyperebene trennen, d.h. es existieren  $(\alpha, \lambda) \in \mathbb{R} \times \mathbb{R}^s$ ,  $(\alpha, \lambda) \neq (0, 0)$ , mit

$$\alpha r_1 + \lambda^\top y_1 \geq \alpha r_2 + \lambda^\top y_2 \quad \forall (r_1, y_1) \in A, (r_2, y_2) \in B. \quad (10.4)$$

Per Definition ist  $(f(\hat{x}) + 1, \Phi(\hat{x})) \in A$  und  $(f^*, \Phi(\hat{x})) \in B$ . Eingesetzt in (10.4) ergibt sich

$$\alpha \underbrace{(f(\hat{x}) + 1 - f^*)}_{>0, \text{ nach (10.1)}} \geq 0 \quad \Rightarrow \quad \alpha \geq 0. \quad (10.5)$$

Überdies gilt  $\lambda \geq 0$ . Um dies zu sehen, testen wir (10.4) mit  $(r_1, y_1) = (f(\hat{x}) + 1, \Phi(\hat{x})) \in A$  und  $(r_2, y_2) = (f^*, z + \Phi(\hat{x})) \in B$  mit beliebigem  $z \in \mathbb{R}^s$  mit  $z \leq 0$ . Dann folgt

$$\underbrace{\alpha}_{\geq 0} \underbrace{(f(\hat{x}) + 1 - f^*)}_{\geq 0} \geq \lambda^\top z \quad \forall z \in \mathbb{R}^s : z \leq 0 \quad \Rightarrow \quad \lambda^\top y \geq 0 \quad \forall y \in \mathbb{R}^s : y \geq 0$$

und damit  $\lambda \geq 0$  wie im Satz behauptet.

Wir zeigen nun  $\alpha > 0$ . Dazu nehmen wir an, dass  $\alpha = 0$ . (Man beachte, dass der Fall  $\alpha < 0$  wegen (10.5) sowieso ausgeschlossen ist.) Wegen  $(\alpha, \lambda) \neq (0, 0)$  ist dann zunächst  $\lambda \neq 0$ . Außerdem folgt im Fall  $\alpha = 0$  aus (10.4) mit  $(f^*, 0) \in B$  und  $(f(\hat{x}) + 1, \Phi(\hat{x})) \in A$ , dass

$$\lambda^\top \Phi(\hat{x}) \geq 0. \quad (10.6)$$

Andererseits existiert wegen  $\lambda \geq 0$  und  $\lambda \neq 0$  ein Index  $\ell \in \{1, \dots, s\}$  mit  $\lambda_\ell > 0$ . Daraus folgt zusammen mit der Slater-Bedingung, dass

$$\lambda^\top \Phi(\hat{x}) = \sum_{i=1}^s \underbrace{\lambda_i}_{\geq 0} \underbrace{\Phi_i(\hat{x})}_{< 0} \leq \lambda_\ell \Phi_\ell(\hat{x}) < 0$$

im Widerspruch zu (10.6). Also gilt  $\alpha > 0$ , und wir nehmen o.B.d.A. an, dass  $\alpha = 1$  (sonst argumentiere mit  $\lambda/\alpha$ ). Seien nun  $x \in C$  und  $\varepsilon > 0$  beliebig. Aus (10.4) folgt mit  $(r_1, y_1) = (f(x) + \varepsilon, \Phi(x)) \in A$  und  $(r_2, y_2) = (f^*, 0) \in B$  (beachte, dass  $\alpha = 1$ ):

$$f^* \leq f(x) + \varepsilon + \lambda^\top \Phi(x) \quad \forall x \in C, \varepsilon > 0.$$

Grenzübergang  $\varepsilon \searrow 0$  liefert dann

$$f^* \leq f(x) + \lambda^\top \Phi(x) \quad \forall x \in C. \quad (10.7)$$

Wegen  $\lambda \geq 0$  ist  $\lambda^\top \Phi(x) \leq 0$  für alle  $x$  mit  $\Phi(x) \leq 0$ , woraus mit (10.7)

$$f^* \leq \inf_{x \in C} (f(x) + \lambda^\top \Phi(x)) \leq \inf_{\substack{x \in C \\ \Phi(x) \leq 0}} f(x) = f^* \quad (10.8)$$

und damit die erste Aussage (10.2) folgt.

Falls ein  $x^* \in C$  existiert mit  $\Phi(x^*) \leq 0$  und  $f(x^*) = f^*$ , dann folgt aus (10.8), dass

$$f^* \leq f(x^*) + \lambda^\top \Phi(x^*) \leq f(x^*) = f^*$$

und somit  $f(x^*) + \lambda^\top \Phi(x^*) = \inf_{x \in C} (f(x) + \lambda^\top \Phi(x))$  und  $\lambda^\top \Phi(x^*) = 0$ .  $\square$

Wir betrachten nun eine spezielle Menge  $C$ , nämlich

$$C = \{x \in \mathbb{R}^n : \Psi(x) \leq 0, h(x) = 0\}, \quad (10.9)$$

wobei  $\Psi : \mathbb{R}^n \rightarrow \mathbb{R}^t$  und  $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$  affin-linear seien, d.h. es gibt Matrizen  $H \in \mathbb{R}^{p \times n}$ ,  $G \in \mathbb{R}^{t \times n}$  und Vektoren  $e \in \mathbb{R}^p$ ,  $d \in \mathbb{R}^t$ , so dass

$$h(x) = Hx - e \quad \text{und} \quad \Psi(x) = Gx - d.$$

Wegen der Linearität von  $h$  und  $\Psi$  ist diese Menge natürlich konvex. Des Weiteren setzen wir

$$m = s + t, \quad g : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad g(x) = (\Phi(x), \Psi(x)).$$

Man beachte, dass  $g$  wegen der Voraussetzung an  $\Phi$  und der Linearität von  $\Psi$  konvex ist. In diesem Fall lautet das konvexe Optimierungsproblem (CP) dann

$$\left. \begin{array}{l} \text{minimiere} \quad f(x) \\ \text{s.t.} \quad g(x) \leq 0, \quad h(x) = 0. \end{array} \right\} \quad (\text{CLP})$$

Darüber hinaus nehmen wir im Folgenden an, dass  $f$  und  $\Phi$  (und damit auch  $g$ ) differenzierbar sind.

### **Satz 10.3 (KKT-Bedingungen)**

Die Zielfunktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  sei differenzierbar und konvex. Darüber hinaus sei  $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$  affin-linear und  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  bestehe aus einem konvexen Teil  $\Phi$ , der differenzierbar sei, und einem affin-linearen Anteil  $\Psi$ . Ferner sei die Slater-Bedingung erfüllt, d.h. es existiere ein  $\hat{x} \in \mathbb{R}^n$  mit

$$h(\hat{x}) = 0, \quad \Psi(\hat{x}) \leq 0, \quad \Phi(\hat{x}) < 0. \quad (10.10)$$

Ist  $x^* \in \mathbb{R}^n$  ein (globaler = lokaler) Minimierer von (CLP), dann existieren *Lagrange-Multiplikatoren*  $\lambda^* \in \mathbb{R}^m$ ,  $\mu^* \in \mathbb{R}^p$ , so dass das folgende *Optimalitätssystem* erfüllt ist:

$$\text{Zulässigkeit:} \quad h(x^*) = 0, \quad g(x^*) \leq 0 \quad (10.11a)$$

$$\text{Gradientengleichung:} \quad \nabla f(x^*) + g'(x^*)^\top \lambda^* + h'(x^*)^\top \mu^* = 0 \quad (10.11b)$$

$$\text{Komplementäre Schlupfbedingung:} \quad \lambda^* \geq 0, \quad (\lambda^*)^\top g(x^*) = 0 \quad (10.11c)$$

*Beweis:* Die Hauptarbeit ist eigentlich mit dem Beweis des Farkas-Lemmas 8.3 und dem von Lemma 10.2 schon erbracht worden.

Zunächst ist zu beachten, dass (CLP) eine konvexe Optimierungsaufgabe darstellt, so dass wir nach Satz 9.1 nicht zwischen globalen und lokalen Optimierern unterscheiden müssen. Ist  $x^* \in \mathbb{R}^n$  ein (globaler = lokaler) Minimierer, dann ist es natürlich sowieso zulässig, was (10.11a) impliziert.

Per Konstruktion ist (CLP) ein Spezialfall von (CP). Wegen der Slater-Bedingung ist Lemma 10.2 anwendbar, so dass ein  $\tilde{\lambda} \in \mathbb{R}^s$  existiert mit

$$f(x^*) + \Phi(x^*)^\top \tilde{\lambda} = \inf_{x \in C} (f(x) + \Phi(x)^\top \tilde{\lambda}) \quad (10.12)$$

$$\Phi(x^*) \leq 0, \quad \Phi(x^*)^\top \tilde{\lambda} = 0, \quad \tilde{\lambda} \geq 0. \quad (10.13)$$

Nach (10.12) löst  $x^*$  also die Aufgabe

$$\left. \begin{array}{l} \text{minimiere } f(x) + \Phi(x)^\top \tilde{\lambda} \\ \text{s.t. } x \in C. \end{array} \right\} \quad (10.14)$$

Da deren zulässige Menge  $C$  konvex ist, erfüllt jede lokale Lösung, also insbesondere  $x^*$ , nach Satz 9.2 die folgende Variationsungleichung:

$$\underbrace{(\nabla f(x^*) + \Phi'(x^*)^\top \tilde{\lambda})^\top}_{=: -b} (x - x^*) \geq 0 \quad \forall x \in C = \{x \in \mathbb{R}^n : Gx \leq d, Hx = e\}. \quad (10.15)$$

Wir definieren nun die Indexmenge

$$\mathcal{A}_{\text{lin}}^* := \{i \in \{1, \dots, t\} : \Psi_i(x^*) = 0\} = \{i \in \{1, \dots, t\} : G_{(i)}^\top x^* = d_i\}.$$

Des Weiteren definieren wir

$$l := |\mathcal{A}_{\text{lin}}^*|, \quad G_a \in \mathbb{R}^{l \times n}, \quad G_a := G_{(\mathcal{A}_{\text{lin}}^*)}, \quad d_a \in \mathbb{R}^l, \quad d_a := d_{\mathcal{A}_{\text{lin}}^*}$$

d.h.  $G_a$  ist die Teilmatrix aus den Zeilenvektoren von  $G$ , deren Zeilenindex in  $\mathcal{A}_{\text{lin}}^*$  liegt. Sei nun  $s \in \mathbb{R}^n$  mit  $HS = 0$  und  $G_a s \leq 0$  beliebig. Dann erfüllt  $x_\tau := x^* + \tau s$  für jedes  $\tau \geq 0$  wegen der Zulässigkeit von  $x^*$

$$Hx_\tau = Hx^* + \tau Hs = e \quad \text{und} \quad G_a x_\tau = G_a x^* + \tau G_a s \leq d_a. \quad (10.16)$$

Für  $i \notin \mathcal{A}_{\text{lin}}^*$  existiert ein  $\tau_i > 0$  hinreichend klein, so dass

$$G_{(i)}^\top x_\tau = \underbrace{G_{(i)}^\top x^*}_{< d_i} + \tau G_{(i)}^\top s \leq d_i \quad \forall \tau \in [0, \tau_i].$$

Wählen wir  $0 < \tau \leq \min_{i \notin \mathcal{A}_{\text{lin}}^*} \tau_i$ , dann folgt daraus zusammen mit (10.16) schließlich  $x_\tau \in C$ , so dass  $x_\tau$  zulässig für die Variationsungleichung (10.15) ist. Diese impliziert dann (nach Division durch  $\tau > 0$ ), dass

$$b^\top s \leq 0 \quad \forall s \in \mathbb{R}^n : Hs = 0, \quad G_a s \leq 0.$$

Mit dem Farkas-Lemma aus Satz 8.3 folgt dann die Existenz von  $\hat{\lambda} \in \mathbb{R}^l$  und  $\mu^* \in \mathbb{R}^p$ , so dass

$$\hat{\lambda} \geq 0 \quad \text{und} \quad b = G_a^\top \hat{\lambda} + H^\top \mu^*. \quad (10.17)$$

Definieren wir den Vektor

$$\bar{\lambda} \in \mathbb{R}^t, \quad \bar{\lambda}_{\mathcal{A}_{\text{lin}}^*} = \hat{\lambda}, \quad \bar{\lambda}_i = 0 \quad \forall i \notin \mathcal{A}_{\text{lin}}^*,$$

dann folgt

$$\Psi'(x^*)^\top \bar{\lambda} = G^\top \bar{\lambda} = G_a^\top \hat{\lambda}$$

und damit aus (10.17):

$$\nabla f(x^*) + \Phi'(x^*)^\top \tilde{\lambda} + \Psi'(x^*)^\top \bar{\lambda} + h'(x^*)^\top \mu^* = 0.$$

Definieren wir  $\lambda^* = (\tilde{\lambda}, \bar{\lambda}) \in \mathbb{R}^m$  dann ergibt sich wegen

$$g'(x^*)^\top \lambda^* = (\Phi'(x^*)^\top, \Psi'(x^*)^\top) \begin{pmatrix} \tilde{\lambda} \\ \bar{\lambda} \end{pmatrix} = \Phi'(x^*)^\top \tilde{\lambda} + \Psi'(x^*)^\top \bar{\lambda}$$

gerade die Gradientengleichung (10.11b).

Es bleibt die komplementäre Schlupfbedingung zu zeigen. Diese folgt aber sofort aus der Definition von  $\lambda^*$ : Zunächst ist  $\tilde{\lambda} \geq 0$  nach Lemma 10.2. Außerdem ist  $\hat{\lambda} \geq 0$  nach dem Farkas-Lemma und daher nach Konstruktion auch  $\bar{\lambda} \geq 0$ , was die Vorzeichenbedingung an den Ungleichungsmultiplikator  $\lambda^*$  impliziert. Darüber hinaus gilt

$$g(x^*)^\top \lambda^* = \underbrace{\Phi(x^*)^\top \tilde{\lambda}}_{=0} + \underbrace{\Psi(x^*)^\top \bar{\lambda}}_{=0} = \underbrace{(G_a x^* - d_a)^\top \hat{\lambda}}_{=0} = 0,$$

wobei wir Lemma 10.2 und die Definition von  $\mathcal{A}_{\text{lin}}^*$  benutzt haben. Damit ist auch die komplementäre Schlupfbedingung bewiesen.  $\square$

#### Bemerkung 10.4

Die Gleichungen und Ungleichungen in (10.11) werden auch als *Karush-Kuhn-Tucker (KKT) Bedingungen* bezeichnet. KKT-Bedingungen stellen i.A. nur unter zusätzlichen Voraussetzungen notwendige Optimalitätsbedingungen dar. Im Fall von (CLP) sind dies die Voraussetzungen von Satz 10.3, also insbesondere die Slater-Bedingung. Derartige Voraussetzungen nennt man auch *Constraint Qualifications (CQ)*. Unter geeigneten CQ kann man auch für Probleme der Bauart (CLP) mit *allgemeinen* (nichtlinearen) Funktionen  $f$ ,  $g$  und  $h$  beweisen, dass (lokale) Optimierer die KKT-Bedingungen erfüllen. Die KKT-Theorie ist also nicht auf konvexe Probleme beschränkt. Dies wird in der Vorlesung ‘‘Nichtlineare Optimierung’’ näher untersucht.  $\diamond$

#### Bemerkung 10.5

Satz 10.3 zeigt insbesondere, dass man in dem Fall, dass alle Komponenten von  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  linear-affin ist, auf die Slater-Bedingung verzichten kann. In diesem Fall reicht also die Existenz eines zulässigen Punktes, d.h.

$$\exists \hat{x} \in \mathbb{R}^n \quad \text{mit} \quad h(\hat{x}) = 0, \quad g(\hat{x}) \leq 0,$$

anstelle von (10.10). *Im Fall linearer Nebenbedingungen ist also außer der Existenz eines zulässigen Punktes keine weitere CQ notwendig!*  $\diamond$

#### Definition 10.6

Ein Punkt  $x^* \in \mathbb{R}^n$  heißt *KKT-Punkt* oder auch *stationärer Punkt* von (CLP), falls Lagrange-Multiplikatoren  $\lambda^* \in \mathbb{R}^m$ ,  $\mu^* \in \mathbb{R}^p$  existieren, so dass das Optimalitätssystem (10.11) erfüllt ist. Ein Tripel  $(x^*, \lambda^*, \mu^*)$ , welches (10.11) erfüllt, heißt *KKT-Tripel*.  $\diamond$

Mit Hilfe der so genannten *Lagrange-Funktion* lässt sich das Optimalitätssystem kompakt aufschreiben.

**Definition 10.7 (Lagrange-Funktion)**

Die Funktion  $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$ , definiert durch

$$\mathcal{L}(x, \lambda, \mu) := f(x) + \lambda^\top g(x) + \mu^\top h(x), \quad (10.18)$$

heißt Lagrange-Funktion.  $\diamond$

Damit ist (10.11) äquivalent zu:

$$\left. \begin{array}{l} \text{Zulässigkeit:} \\ \text{Gradientengleichung:} \end{array} \right\} \begin{array}{l} \nabla_{\mu} \mathcal{L}(x^*, \lambda^*, \mu^*) = 0 \\ \nabla_{\lambda} \mathcal{L}(x^*, \lambda^*, \mu^*) \leq 0 \end{array} \quad (10.19a)$$

$$\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) = 0 \quad (10.19b)$$

$$\text{Komplementäre Schlupfbedingung:} \quad \lambda^* \geq 0, \quad (\lambda^*)^\top g(x^*) = 0. \quad (10.19c)$$

KKT-Tripel lassen sich im konvexen Fall nicht nur durch das Optimalitätssystem charakterisieren, sondern auch durch die so genannte Sattelpunkteigenschaft, wie wir im Folgenden sehen werden.

**Definition 10.8 (Sattelpunkt)**

Ein Tripel  $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$  heißt *Sattelpunkt* der Lagrange-Funktion, falls  $\lambda^* \geq 0$  und

$$\mathcal{L}(x^*, \lambda, \mu) \leq \mathcal{L}(x^*, \lambda^*, \mu^*) \leq \mathcal{L}(x, \lambda^*, \mu^*) \quad (10.20)$$

für alle  $(x, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$  mit  $\lambda \geq 0$  gilt.  $\diamond$

**Lemma 10.9**

Ein Tripel  $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$  ist genau dann ein Sattelpunkt der Lagrange-Funktion, falls es ein KKT-Tripel von (CLP) ist.

*Beweis:* Sei  $(x^*, \lambda^*, \mu^*)$  ein KKT-Tripel. Wir beweisen zunächst die rechte Ungleichung in (10.20). Das freie Optimierungsproblem

$$\underset{x \in \mathbb{R}^n}{\text{minimiere}} \quad \mathcal{L}(x, \lambda^*, \mu^*) \quad (10.21)$$

ist ein konvexes Problem, denn die Zielfunktion ist wegen Konvexität von  $f$  und  $g$ , Linearität von  $h$  und Nicht-Negativität von  $\lambda^*$  konvex, wie man leicht überprüft. Da  $(x^*, \lambda^*, \mu^*)$  ein KKT-Tripel ist, gilt die Gradientengleichung (10.19b), was gerade der notwendigen Optimalitätsbedingung von (10.21) entspricht, vgl. Satz 9.2. Wegen Konvexität sind die notwendigen Bedingungen auch hinreichend, s. Satz 9.4. Daher löst  $x^*$  die Aufgabe (10.21), was die rechte Ungleichung beweist.

Zur linken Ungleichung in (10.20): Aus der Zulässigkeit von  $x^*$  und der komplementären Schlupfbedingung folgt für alle  $\lambda \in \mathbb{R}^m$  mit  $\lambda \geq 0$ , dass

$$\mathcal{L}(x^*, \lambda, \mu) - \mathcal{L}(x^*, \lambda^*, \mu^*) = g(x^*)^\top \lambda - \underbrace{g(x^*)^\top \lambda^*}_{=0} + \underbrace{h(x^*)^\top (\mu - \mu^*)}_{=0} = g(x^*)^\top \lambda \leq 0$$

was schließlich die Sattelpunktungleichung liefert.

Sei  $(x^*, \lambda^*, \mu^*)$ ,  $\lambda^* \geq 0$ , nun ein Sattelpunkt der Lagrange-Funktion. Dann löst  $x^*$  das freie Problem (10.21) und erfüllt somit dessen notwendige Optimalitätsbedingung, die gerade der Gradientengleichung (10.19b) entspricht. Die Zulässigkeit von  $x^*$  folgt aus der linken Ungleichung in (10.20). Diese ist äquivalent zu

$$g(x^*)^\top (\lambda - \lambda^*) + h(x^*)^\top (\mu - \mu^*) \leq 0 \quad \forall \mu \in \mathbb{R}^p, \lambda \in \mathbb{R}^m : \lambda \geq 0. \quad (10.22)$$

Wählen wir  $\lambda = \lambda^* \geq 0$  und  $\mu = h(x^*) + \mu^*$ , dann folgt  $\|h(x^*)\|^2 \leq 0$ , also  $h(x^*) = 0$ . Um die Zulässigkeit für die Ungleichungsnebenbedingungen zu zeigen, sei  $i \in \{1, \dots, m\}$  beliebig. Wir wählen  $\lambda = e_i + \lambda^* \geq 0$  und  $\mu = \mu^*$  in (10.22), wobei  $e_i \in \mathbb{R}^m$  wieder der  $i$ -te Einheitsvektor ist. Dann folgt  $g(x^*)^\top e_i = g_i(x^*) \leq 0$ . Da  $i$  beliebig war, folgt insgesamt die Zulässigkeit von  $x^*$ . Es bleibt schließlich nur noch die komplementäre Schlupfbedingung zu zeigen. Dazu testen wir (10.22) mit  $\lambda = 0$  und  $\lambda = 2\lambda^* \geq 0$  (und jeweils  $\mu = \mu^*$ ). Dann ergibt sich sofort  $g(x^*)^\top \lambda^* = 0$ , und  $(x^*, \lambda^*, \mu^*)$  ist somit ein KKT-Tripel wie behauptet.  $\square$

### Bemerkung 10.10

Man beachte, dass wir für den zweiten Teil des Beweises die Konvexität von  $f$  und  $g$  und die Linearität von  $h$  gar nicht benötigt haben. Das bedeutet, Sattelpunkte sind auch für beliebige (differenzierbare)  $f$ ,  $g$  und  $h$  KKT-Tripel. Die Umkehrung aber gilt i.A. nur im konvexen Fall.  $\diamond$

Als nächstes betrachten wir eine Aufgabe, die formal genauso aussieht wie (CLP), d.h.

$$\left. \begin{array}{l} \text{minimiere } f(x) \\ \text{s.t. } g(x) \leq 0, \quad h(x) = 0, \end{array} \right\} \quad (\text{NLP})$$

allerdings sind die Funktionen  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  und  $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$  diesmal *beliebig*, d.h. insbesondere, dass  $f$  und  $g$  nicht konvex und  $h$  nicht affin-linear sein muss. Aufgabe (NLP) hat also die Struktur eines (allgemeinen) nichtlinearen Programms.

### Proposition 10.11

Ist  $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ ,  $\lambda^* \geq 0$ , Sattelpunkt der Lagrange-Funktion, dann löst  $x^*$  die Aufgabe (NLP).

*Beweis:* Die Zulässigkeit von  $x^*$  zeigt man genauso wie im zweiten Teil des Beweises von Lemma 10.9. (Man beachte, dass die Konvexität von  $f$  und  $g$  und die Linearität von  $h$  hierfür nicht gebraucht werden, vgl. Bemerkung 10.10). Zur Optimalität von  $x^*$ : Ebenfalls wie im zweiten Teil des Beweises von Lemma 10.9 zeigt man die komplementäre Schlupfbedingung. Zusammen mit  $h(x^*) = 0$  folgt aus dieser

$$f(x^*) = f(x^*) + g(x^*)^\top \lambda^* + h(x^*)^\top \mu^* \stackrel{(10.20)}{\leq} f(x) + g(x)^\top \lambda^* + h(x)^\top \mu^* \leq f(x)$$

für alle  $x \in \mathbb{R}^n$  mit  $g(x) \leq 0$  und  $h(x) = 0$ . Bei der letzten Ungleichung haben wir  $\lambda^* \geq 0$  ausgenutzt. Damit ist  $x^*$  global optimal für (NLP).  $\square$

Da KKT-Tripel nach Lemma 10.9 im konvexen Fall Sattelpunkte sind und diese nach Proposition 10.11 globale Minimierer darstellen, erhalten wir sofort das folgende

### Korollar 10.12

Jeder KKT-Punkt der konvexen Aufgabe (CLP) ist ein globaler Minimierer von (CLP).  $\diamond$

### Bemerkung 10.13

Die KKT-Bedingungen sind im konvexen Fall also *hinreichend* für Optimalität. Man beachte, dass sie i.A. nicht *notwendig* für Optimalität sind, es sei denn, eine CQ wie z.B. die Slater-Bedingung ist erfüllt, wie wir in Satz 10.3 gesehen haben.  $\diamond$

Die linke Ungleichung in (10.20), d.h.

$$\mathcal{L}(x^*, \lambda^*, \mu^*) = \sup_{\substack{\lambda \geq 0 \\ \mu \in \mathbb{R}^p}} \mathcal{L}(x^*, \lambda, \mu)$$

deutet daraufhin, dass auch die Lagrange-Multiplikatoren ein Optimierungsproblem lösen und zwar ein Maximierungsproblem. Dieses bezeichnet man als *duales Problem*.

### Definition 10.14

Das *primale Problem* zu (NLP) ist gegeben durch

$$\inf_{x \in \mathbb{R}^n} \sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \mathcal{L}(x, \lambda, \mu), \quad (\text{P})$$

wobei  $\mathcal{L}$  die Lagrange-Funktion aus (10.18) ist. Das *duale Problem* zu (NLP) ist definiert durch

$$\sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu). \quad (\text{D})$$

Die Extremwerte von (P) und (D) werden mit  $\inf(P), \sup(D) \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$  bezeichnet, d.h.

$$\inf(P) = \inf_{x \in \mathbb{R}^n} \sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \mathcal{L}(x, \lambda, \mu) \quad \text{und} \quad \sup(D) = \sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu). \quad \diamond$$

### Lemma 10.15

Das primale Problem (P) ist äquivalent zu (NLP) in dem Sinn, dass lokale (und damit auch globale) Lösungen von (P) auch (NLP) lösen und umgekehrt.

*Beweis:* Wegen

$$\sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \mathcal{L}(x, \lambda, \mu) = \sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} (f(x) + g(x)^\top \lambda + h(x)^\top \mu) = \infty,$$

falls  $g_i(x) \not\leq 0$  oder  $h(x) \neq 0$ , kommen unzulässige  $x \in \mathbb{R}^n$  bei der äußeren Minimierung in (P) nicht in Frage. Für zulässige Punkte, d.h.  $x \in \mathbb{R}^n$  mit  $h(x) = 0$  und  $g(x) \leq 0$ , folgt

$$\sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} (f(x) + \underbrace{g(x)^\top \lambda}_{\leq 0} + \underbrace{h(x)^\top \mu}_{=0}) = f(x),$$

so dass

$$\sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \mathcal{L}(x, \lambda, \mu) = f(x) \quad \forall x \in \mathbb{R}^n : g(x) \leq 0, h(x) = 0.$$

Daher stimmen die Zielfunktionen von (P) und (NLP) auf der zulässigen Menge überein, woraus die Aussage folgt.  $\square$

Es gilt also

$$(\text{P}) \Leftrightarrow \begin{cases} \text{minimiere} & f(x) \\ \text{s.t.} & g(x) \leq 0, \quad h(x) = 0. \end{cases}$$

Definiert man die so genannte *duale Zielfunktion*  $q : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R} \cup \{-\infty\}$  durch

$$q(\lambda, \mu) := \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu),$$

dann lässt sich das duale Problem auch schreiben als

$$(\text{D}) \Leftrightarrow \begin{cases} \text{maximiere} & q(\lambda, \mu) \\ \text{s.t.} & \lambda \geq 0. \end{cases}$$

Im Gegensatz zum primalen Problem ist das duale Problem eine Maximierungsaufgabe. Wozu das duale Problem nützlich ist, zeigt der sogenannte starke Dualitätssatz, den wir im Folgenden herleiten.

**Lemma 10.16**

Für jedes Optimierungsproblem der Form (NLP) gilt:

$$\sup(D) \leq \inf(P) \quad (10.23)$$

*Beweis:* Nach Lemma 10.15 gilt

$$\inf(P) = \inf_{x \in \mathbb{R}^n} \sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \mathcal{L}(x, \lambda, \mu) = \inf_{\substack{g(x) \leq 0 \\ h(x) = 0}} f(x).$$

Andererseits ist für alle  $\mu \in \mathbb{R}^p$ ,  $\lambda \in \mathbb{R}^m$  mit  $\lambda \geq 0$

$$\inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu) \leq \inf_{\substack{g(x) \leq 0 \\ h(x) = 0}} (f(x) + \underbrace{g(x)^\top \lambda}_{\leq 0} + \underbrace{h(x)^\top \mu}_{=0}) \leq \inf_{\substack{g(x) \leq 0 \\ h(x) = 0}} f(x),$$

was die Aussage liefert.  $\square$

**Bemerkung 10.17**

- Falls (P) und (D) beide zulässig sind, folgt also  $\inf(P) \in \mathbb{R}$  und  $\sup(D) \in \mathbb{R}$ .
- Das duale Problem gibt nach Lemma 10.16 eine untere Schranke für den optimalen Zielfunktionswert des primalen Problems an.  $\diamond$

Aus Lemma 10.16 folgt insbesondere, dass

$$f(x) \geq \inf(P) \geq \sup(D) \geq q(\lambda, \mu)$$

für alle  $(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p$  mit  $\lambda \geq 0$  und für alle  $x \in \mathbb{R}^n$ , die zulässig für das primale Problem (NLP) sind, d.h. für alle  $x$ , für die  $h(x) = 0$  und  $g(x) \leq 0$  gilt. Das führt auf den so genannten Schwachen-Dualitätssatz:

**Korollar 10.18 (Schwache Dualität)**

Ist  $x$  zulässig für das primale Problem und  $(\lambda, \mu)$  zulässig für das duale Problem, dann gilt  $f(x) \geq q(\lambda, \mu)$ .

Es gibt durchaus Optimierungsaufgaben, bei denen  $\inf(P) > \sup(D)$  gilt, wie das folgende Beispiel zeigt:

**Beispiel 10.19**

Betrachte das Problem

$$\left. \begin{array}{l} \text{minimiere } f(x) \\ \text{s.t. } x \geq 0 \end{array} \right\}$$

mit

$$f(x) = \begin{cases} x^2 - 2x, & x \geq 0 \\ x, & x < 0. \end{cases}$$

Man bestätigt leicht, dass das globale Minimum dieser Aufgabe bei  $x = 1$  liegt, so dass  $\inf(P) = f(1) = -1$  ist. Für die Lagrange-Funktion erhalten wir

$$\mathcal{L}(x, \lambda) = \begin{cases} x^2 - (2 + \lambda)x, & x \geq 0 \\ (1 - \lambda)x, & x < 0, \end{cases}$$



so dass

$$q(\lambda) = \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda) = \begin{cases} -(1/4)(2 + \lambda)^2, & \lambda \geq 1 \\ -\infty, & \lambda < 1. \end{cases}$$

Daraus folgt wiederum, dass

$$\sup(D) = \sup_{\lambda \geq 0} q(\lambda) = q(1) = -9/4 < -1 = \inf(P). \quad \diamond$$

Es kann also sein, dass eine so genannte *Dualitätslücke* auftritt, dass also  $\inf(P) > \sup(D)$ . Im Fall einer Dualitätslücke können keine Lagrange-Multiplikatoren existieren, wie das folgende Resultat zeigt:

**Proposition 10.20**

Ein Tripel  $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$  mit  $\lambda^* \geq 0$  ist genau dann Sattelpunkt der Lagrange-Funktion, falls  $x^*$  das primale Problem (P) löst,  $(\lambda^*, \mu^*)$  das duale Problem (D) und zudem

$$\sup(D) = \inf(P) \quad (10.24)$$

gilt.

*Beweis:* Sei  $(x^*, \lambda^*, \mu^*)$  ein Sattelpunkt. Dann folgt zusammen mit Lemma 10.16, dass

$$\begin{aligned} & \mathcal{L}(x^*, \lambda^*, \mu^*) \\ & \stackrel{(10.20)}{=} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda^*, \mu^*) \\ & \leq \sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu) = \sup(D) \\ & \stackrel{(10.23)}{\leq} \inf(P) = \inf_{x \in \mathbb{R}^n} \sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \mathcal{L}(x, \lambda, \mu) \\ & \leq \sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \mathcal{L}(x^*, \lambda, \mu) \stackrel{(10.20)}{=} \mathcal{L}(x^*, \lambda^*, \mu^*) \end{aligned}$$

und damit  $\inf(P) = \sup(D) = \mathcal{L}(x^*, \lambda^*, \mu^*)$ , was den ersten Teil der Aussage beweist.

Um die umgekehrte Implikation zu beweisen, seien  $x^*$  und  $(\lambda^*, \mu^*)$  Lösungen von (P) bzw. (D). Dann folgt

$$\begin{aligned} \mathcal{L}(x^*, \lambda^*, \mu^*) & \geq \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda^*, \mu^*) \\ & = \sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu) = \sup(D) \\ & \stackrel{(10.24)}{=} \inf(P) = \inf_{x \in \mathbb{R}^n} \sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \mathcal{L}(x, \lambda, \mu) \\ & = \sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \mathcal{L}(x^*, \lambda, \mu) \geq \mathcal{L}(x^*, \lambda^*, \mu^*), \end{aligned}$$

so dass

$$\begin{aligned} \mathcal{L}(x^*, \lambda^*, \mu^*) & = \sup_{\substack{\mu \in \mathbb{R}^p \\ \lambda \geq 0}} \mathcal{L}(x^*, \lambda, \mu) \\ \text{und } \mathcal{L}(x^*, \lambda^*, \mu^*) & = \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda^*, \mu^*), \end{aligned}$$

was die Sattelpunkteigenschaft zeigt.  $\square$

Als unmittelbare Konsequenz aus Satz 10.3, Lemma 10.9 und Proposition 10.20 erhalten wir den folgenden

**Satz 10.21 (Starke Dualität)**

Die Zielfunktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  sei differenzierbar und konvex. Darüber hinaus sei  $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$  affin-linear und  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  bestehe aus einem konvexen Teil  $\Phi$ , der differenzierbar sei, und einem affin-linearen Anteil  $\Psi$ . Ist das primale Problem lösbar und existiert ein Slater-Punkt  $\hat{x} \in \mathbb{R}^n$ , so dass

$$h(\hat{x}) = 0, \quad \Psi(\hat{x}) \leq 0, \quad \Phi(\hat{x}) < 0,$$

dann gilt  $\inf(P) = \sup(D)$  und das duale Problem besitzt ebenfalls eine Lösung.

Die Slater-Bedingung ist in der Tat entscheidend für Satz 10.21, wie das folgende Beispiel zeigt:

**Beispiel 10.22**

Wir betrachten die Aufgabe

$$\left. \begin{array}{l} \text{minimiere} \quad -x \\ \text{s.t.} \quad x^2 \leq 0. \end{array} \right\}$$

Mit  $f(x) = -x$  und  $g(x) = x^2$  passt die Aufgabe in das allgemeine Setting von (CLP). Die zulässige Menge ist durch  $\mathcal{F} = \{0\}$  gegeben, so dass  $\inf(P) = f(0) = 0$  ist. Die Slater-Bedingung ist natürlich nicht erfüllt. Für die Lagrange-Funktion erhalten wir

$$\mathcal{L}(x, \lambda) = -x + \lambda x^2.$$

Diese nimmt für positive  $\lambda$  ihr globales Minimum bei  $x_\lambda = 1/(2\lambda)$  an. Die duale Zielfunktion ist daher

$$q(\lambda) = \begin{cases} \mathcal{L}(x_\lambda, \lambda) = -\frac{1}{4\lambda}, & \lambda > 0 \\ -\infty, & \lambda \leq 0. \end{cases}$$

Daher gilt zwar  $\sup_{\lambda \geq 0} q(\lambda) = 0 = \inf(P)$ , aber dieses Supremum wird nicht angenommen. Es existieren also keine Lagrange-Multiplikatoren für diese Aufgabe. (Das liegt allerdings an der selten dämlichen Beschreibung der zulässigen Menge ..)  $\diamond$

Einen Überblick über den Zusammenhang zwischen Sattelpunkteigenschaft, KKT-Bedingungen, Optimalität und Dualität gibt die Abbildung 10.1.

**Bemerkung 10.23**

Sind *alle* Nebenbedingungen affin-linear, dann wird die Slater-Bedingung zum Beweis von Satz 10.3 nicht benötigt, wie bereits in Bemerkung 10.5 erwähnt. Dementsprechend gilt auch der starke Dualitätssatz im Fall affin-linearer Nebenbedingungen, *ohne* dass die Slater-Bedingung vorausgesetzt werden muss.  $\diamond$

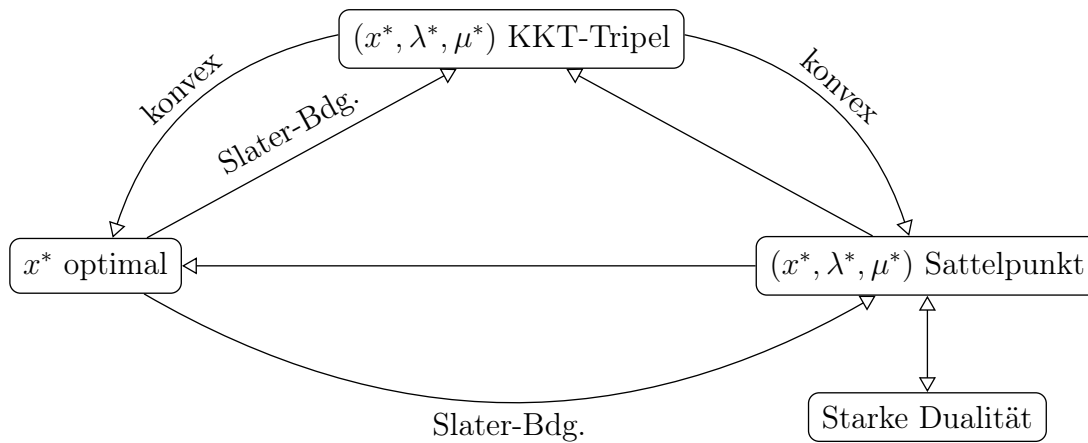


ABBILDUNG 10.1. Übersicht über Optimalitätsbedingungen in der konvexen Optimierung. Mit „konvex“ ist hierbei gemeint, dass das Optimierungsproblem die Form (CLP) mit konvexem  $f$  und  $g$  und linear-affinem  $h$  haben muss. „Slater-Bdg.“ steht abkürzend für die Voraussetzungen von Satz 10.3. Mit „Starker Dualität“ ist gemeint, dass primales und duales Problem lösbar sind und  $\inf(P) = \sup(D)$  gilt. Pfeile ohne Beschriftung gelten auch im Fall von (NLP).

## § 11 Subgradientenverfahren

Ziel dieses Abschnitts ist es, einen einfachen, aber sehr allgemein verwendbaren Algorithmus zur Minimierung konvexer Funktionen ohne Nebenbedingungen zu beschreiben. Wir betrachten also das Problem

$$\left. \begin{array}{l} \text{minimiere } f(x) \\ \text{s.t. } x \in \mathbb{R}^n \end{array} \right\}$$

für eine konvexe Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ .

Die folgende Definition verallgemeinert den Begriff des Gradienten auf konvexe, aber nicht notwendigerweise differenzierbare Funktionen.

### Definition 11.1

Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  konvex und  $x_0 \in \mathbb{R}^n$ . Dann heißt  $s \in \mathbb{R}^n$  *Subgradient* von  $f$  in  $x_0$ , falls gilt:

$$f(x) - f(x_0) \geq s^\top(x - x_0) \quad \forall x \in \mathbb{R}^n.$$

Die Menge aller Subgradienten von  $f$  in  $x_0$  heißt *Subdifferential* von  $f$  in  $x_0$  und wird mit  $\partial f(x_0)$  bezeichnet.

Definition 11.1 wird in Abbildung ?? illustriert.

### Satz 11.2

Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  konvex und  $x_0 \in \mathbb{R}^n$ . Dann ist  $x_0$  ein globaler Minimierer von  $f$  genau dann, wenn  $0 \in \partial f(x_0)$ .

*Beweis:* Nach Definition ist  $x_0$  globaler Minimierer von  $f$  genau dann, wenn

$$f(x) - f(x_0) \geq 0 \quad \forall x \in \mathbb{R}^n$$

gilt, also genau dann, wenn  $0$  Subgradient von  $f$  in  $x_0$  ist.  $\square$

Die Existenz von Subgradienten ist etwas schwieriger zu beweisen, sie basiert auf dem Trennungssatz für konvexe Mengen.

### Satz 11.3

Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  konvex und  $x_0 \in \mathbb{R}^n$ . Dann gilt  $\partial f(x_0) \neq \emptyset$ .

*Beweis:* Nach Lemma 5.3 ist der Epigraph  $\text{epi}(f)$  eine konvexe Menge. Offensichtlich gilt  $\text{int}(\text{epi}(f)) \neq \emptyset$ ,  $\text{epi}(f) \neq \mathbb{R}^{n+1}$ , und  $(x_0, f(x_0)) \notin \text{int}(\text{epi}(f))$ . Nach Lemma 7.4 existiert ein  $c \in \mathbb{R}^{n+1} \setminus \{0\}$  mit

$$c^\top(x_0, f(x_0)) \geq \sup_{z \in \text{epi}(f)} c^\top z.$$

Mit  $s := (c_1, \dots, c_n) \in \mathbb{R}^n$  gilt also

$$s^\top(x - x_0) + c_{n+1}(y - f(x_0)) \leq 0 \quad \forall (x, y) \in \text{epi}(f). \quad (11.1)$$

Aus  $(x_0, f(x_0) + 1) \in \text{epi}(f)$  folgt insbesondere  $c_{n+1} \leq 0$ . Falls  $c_{n+1} = 0$  gilt, liefert (11.1) auch  $s = 0$ , also einen Widerspruch zu  $c \neq 0$ . Wir können nach Skalierung von  $c$  also  $c_{n+1} = -1$  annehmen. Dann impliziert (11.1)

$$s^\top(x - x_0) \leq f(x) - f(x_0) \quad \forall x \in \mathbb{R}^n$$

und somit  $s \in \partial f(x_0)$  nach Definition.  $\square$

**Bemerkung 11.4**

Man kann auch die folgende Umkehrung von Satz 11.3 zeigen: Falls für eine Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  an jeder Stelle mindestens ein Subgradient existiert, so ist  $f$  konvex. Zum Beweis kann man wiederum Lemma 5.3 zusammen mit Satz 4.5 (1) verwenden.

**Satz 11.5**

Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  konvex und diffbar in  $x_0 \in \mathbb{R}^n$ . Dann gilt  $\partial f(x_0) = \{\nabla f(x_0)\}$ .

*Beweis:* Sei  $s \in \partial f(x_0)$ . Für alle  $d \in \mathbb{R}^n$  gilt dann

$$\nabla f(x_0)^\top d = \lim_{t \rightarrow 0^+} \frac{f(x_0 + td) - f(x_0)}{t} \geq \lim_{t \rightarrow 0^+} \frac{s^\top (x_0 + td - x_0)}{t} = s^\top d$$

und analog  $\nabla f(x_0)^\top (-d) \geq s^\top (-d)$ . Es folgt  $\nabla f(x_0)^\top d = s^\top d$ , insgesamt ergibt sich also  $\nabla f(x_0) = s$  und damit  $\partial f(x_0) \subseteq \{\nabla f(x_0)\}$ . Der Rest der Aussage folgt nun aus Satz 11.3 (oder direkter aus Satz 5.4).  $\square$

**Algorithmus 11.6 (Subgradientenverfahren)**

- 1: Gegeben sei eine konvexe Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ .
- 2: Wähle  $x^0 \in \mathbb{R}^n$  und setze  $k := 0$ .
- 3: Bestimme ein  $s^k \in \partial f(x^k)$ .
- 4: **if**  $s^k = 0$  **then**
- 5:     **stop:**  $x^k$  ist Minimierer von  $f$ .
- 6: **else**
- 7:     Wähle Schrittweite  $\sigma_k > 0$ .
- 8:     Setze  $x^{k+1} := x^k - \frac{\sigma_k}{\|s^k\|} s^k$  und  $k := k + 1$ .
- 9:     **goto** 3
- 10: **end if**

**Satz 11.7**

Die Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  sei konvex und besitze einen Minimierer  $x^*$ . Angenommen, Algorithmus 11.6 stoppt nicht und die in Schritt 7 gewählten Schrittweiten erfüllen

$$\sum_{k=0}^{\infty} \sigma_k^2 < \infty, \quad \sum_{k=0}^{\infty} \sigma_k = \infty.$$

Außerdem gelte  $\|s^k\| \leq S$  für alle  $k \in \mathbb{N}$ . Dann folgt

$$\lim_{k \rightarrow \infty} \min\{f(x^0), \dots, f(x^k)\} = f(x^*).$$

*Beweis:* Wegen  $s^k \in \partial f(x^k)$  und  $\|s^k\| \leq S$  gilt

$$\begin{aligned} \|x^{k+1} - x^*\|^2 &= \left\| x^k - \frac{\sigma_k}{\|s^k\|} s^k - x^* \right\|^2 \\ &= \|x^k - x^*\|^2 - 2 \frac{\sigma_k}{\|s^k\|} (s^k)^\top (x^k - x^*) + \frac{\sigma_k^2}{\|s^k\|^2} \|s^k\|^2 \\ &\leq \|x^k - x^*\|^2 - \frac{2}{S} \sigma_k (f(x^k) - f(x^*)) + \sigma_k^2. \end{aligned}$$

Rekursive Anwendung ergibt

$$0 \leq \|x^{k+1} - x^*\|^2 \leq \|x^0 - x^*\|^2 - \frac{2}{S} \sum_{i=0}^k \sigma_i (f(x^i) - f(x^*)) + \sum_{i=0}^k \sigma_i^2$$

und damit

$$\sum_{i=0}^k \sigma_i (f(x^i) - f(x^*)) \leq \frac{S}{2} \left( \|x^0 - x^*\|^2 + \sum_{i=0}^k \sigma_i^2 \right).$$

Es gilt ferner

$$\left( \min\{f(x^0), \dots, f(x^k)\} - f(x^*) \right) \sum_{i=0}^k \sigma_i \leq \sum_{i=0}^k \sigma_i (f(x^i) - f(x^*)),$$

also

$$\min\{f(x^0), \dots, f(x^k)\} - f(x^*) \leq \frac{S(\|x^0 - x^*\|^2 + \sum_{i=0}^k \sigma_i^2)}{2 \sum_{i=0}^k \sigma_i}.$$

Nach Voraussetzung konvergiert der Zähler und der Nenner gegen unendlich, also konvergiert die linke Seite gegen Null.  $\square$

### Bemerkung 11.8

Die Folge  $\sigma_k := \frac{1}{k+1}$  erfüllt die Voraussetzungen von Satz 11.7.

### Bemerkung 11.9

Die Aussage von Satz 11.7 bleibt auch ohne die Voraussetzung gültig, dass die Subgradienten beschränkt sind.

### Beispiel 11.10

Betrachte die Betragsfunktion  $f(x) = |x|$ . Es gilt

$$\partial f(x_0) = \begin{cases} \{-1\}, & x_0 < 0 \\ \{1\}, & x_0 > 0 \\ [-1, 1], & x_0 = 0. \end{cases}$$

Die Verfahren "pendelt" also um den Nullpunkt. Es stoppt nur dann, wenn  $x^k = 0$  für ein  $k$  gilt und aus  $\partial f(x^k) = [-1, 1]$  die Null als Subgradient gewählt wird. Andernfalls konvergiert  $x^k$  gegen Null.

Teil 2

Lineare Optimierung





## Polytope und Polyeder

### Inhalt

12	<b>Darstellungen von Polyedern</b>	49
13	<b>Seitenflächen, Facetten und Ecken</b>	52
14	<b>Satz von Weyl-Minkowski</b>	57

### § 12 Darstellungen von Polyedern

In diesem Teil der Vorlesung behandeln wir spezielle konvexe Optimierungsprobleme, nämlich *lineare Programme* (LP), d.h. Probleme der Form

$$\left. \begin{array}{l} \text{minimiere } c^\top x \\ \text{s.t. } Ax = b, \quad Gx \leq d \end{array} \right\} \quad (12.1)$$

mit  $c \in \mathbb{R}^n$ ,  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $G \in \mathbb{R}^{p \times n}$  und  $d \in \mathbb{R}^p$ . Derartige Aufgaben finden sich in zahlreichen Anwendungen wieder, wie beispielsweise beim Transportproblem aus Beispiel 2.1.

Wir untersuchen zunächst die zulässige Menge von (12.1) genauer. Dazu definieren wir:

#### Definition 12.1

Eine Menge  $P \subset \mathbb{R}^n$  heißt *Polyeder*, falls sie der Durchschnitt endlich vieler abgeschlossener Halbräume ist. Ein beschränktes Polyeder heißt *Polytop*.  $\diamond$

Beispiele für Polyeder zeigt Abbildung ???. Da Halbräume nach Definition 4.3 konvex sind, folgt aus Satz 4.5 sofort:

#### Korollar 12.2

Jedes Polyeder ist konvex.  $\diamond$

#### Bemerkung 12.3

Nach Korollar 7.3 ist jede konvexe abgeschlossene Menge Durchschnitt abgeschlossener Halbräume. Polyeder zeichnen sich per Definition dadurch aus, dass der Durchschnitt nur von *endlich vielen* solchen Halbräumen gebildet wird.

#### Lemma 12.4

Eine Menge  $P \subset \mathbb{R}^n$  ist genau dann ein Polyeder, falls  $m \in \mathbb{N}$ ,  $A \in \mathbb{R}^{m \times n}$  und  $b \in \mathbb{R}^m$  existieren, so dass

$$P = P(A, b) := \{x \in \mathbb{R}^n : Ax \leq b\}. \quad (12.2)$$

Die Matrix  $A$  und der Vektor  $b$  sind i.A. nicht eindeutig bestimmt.

*Beweis:* Wegen

$$P(A, b) = \{x \in \mathbb{R}^n : A_{(i)}^\top x \leq b_i \forall i = 1, \dots, m\} = \bigcap_{i=1}^m H(A_{(i)}, b_i)_{\leq} \quad (12.3)$$

ist  $P(A, b)$  ein Polyeder. Hierbei bezeichnen  $A_{(i)}$  wieder die Zeilen von  $A$  und  $H(A_{(i)}, b_i)_{\leq}$  den entsprechenden Halbraum, vgl. Definition 4.3. Ist andersrum  $P$  ein Polyeder, dann lässt  $P$  sich als Schnitt endlich vieler Halbräume schreiben, was wegen (12.3) auf die Darstellung (12.2) führt. Dass  $A$  und  $b$  nicht eindeutig sind, folgt beispielsweise aus  $P = P \cap \{x : 0^\top x \leq 1\}$ .  $\square$

### Lemma 12.5

Ein Polyeder  $P \subset \mathbb{R}^n$  ist genau dann ein Kegel, falls ein  $m \in \mathbb{N}$  und ein  $A \in \mathbb{R}^{m \times n}$  existieren, so dass  $P = P(A, 0)$ .

*Beweis:* Dass  $P(A, 0) = \{x \in \mathbb{R}^n : Ax \leq 0\}$  ein Kegel ist, sieht man sofort. Um „ $\Rightarrow$ “ zu zeigen, sei der Polyeder  $P$  ein Kegel. Nach Lemma 12.4 existieren  $A \in \mathbb{R}^{m \times n}$  und  $b \in \mathbb{R}^m$ , so dass sich  $P$  in der Form  $P(A, b)$  schreiben lässt. Sei nun  $x \in P$  beliebig. Dann ist  $tx \in P$  für alle  $t \geq 0$ , so dass

$$Ax \leq \frac{1}{t} b \xrightarrow{t \rightarrow \infty} 0, \quad (12.4)$$

und daher  $P \subset P(A, 0) = \{x \in \mathbb{R}^n : Ax \leq 0\}$ . Da  $P$  aber ein Kegel ist, folgt  $0 \in P$  und deshalb  $b \geq A0 = 0$ . Daher gilt für jedes  $x \in P(A, 0)$ , dass  $Ax \leq 0 \leq b$ , was  $x \in P$  impliziert. Daraus folgt  $P = P(A, 0)$  und damit die Aussage.  $\square$

Dieses Lemma führt auf folgende

### Definition 12.6

Ein Polyeder der Form  $P(A, 0)$  mit einer Matrix  $A \in \mathbb{R}^{m \times n}$  heißt *polyedrischer Kegel*.  $\diamond$

Das folgende Resultat zeigt, dass die zulässige Menge von (12.1) ebenfalls ein Polyeder ist.

### Lemma 12.7

Eine Menge der Form

$$P = \{x \in \mathbb{R}^n : Ax = b, Gx \leq d\}$$

mit  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $G \in \mathbb{R}^{p \times n}$  und  $d \in \mathbb{R}^p$  ist ein Polyeder.

*Beweis:* Wir definieren

$$\tilde{A} := \begin{pmatrix} A \\ -A \\ G \end{pmatrix} \in \mathbb{R}^{(2m+p) \times n} \quad \text{und} \quad \tilde{b} := \begin{pmatrix} b \\ -b \\ d \end{pmatrix} \in \mathbb{R}^{2m+p}. \quad (12.5)$$

Dann ist  $P = \{x \in \mathbb{R}^n : \tilde{A}x \leq \tilde{b}\}$  und nach Lemma 12.4 somit ein Polyeder.  $\square$

### Bemerkung 12.8

Das obige Lemma zeigt also, dass jede Menge, die durch eine endliche Zahl affiner Gleichungen und Ungleichungen beschrieben wird, ein Polyeder ist.  $\diamond$

Das bedeutet, LPs entsprechen Optimierungsproblemen mit einer polyedrischen zulässigen Menge und einer linearen Zielfunktion. Die Lemmata 12.4 und 12.7 zeigen, dass man einen Polyeder und damit auch zugehörige LPs auf unterschiedliche Weise darstellen kann. Eine spezielle und oft sehr nützliche Darstellung ist die so genannte *Standardform*.

**Definition 12.9 (LP in Standardform)**

Ein LP der Form

$$\left. \begin{array}{l} \text{minimiere } c^\top x \\ \text{s.t. } Ax = b, \quad x \geq 0 \end{array} \right\} \quad (\text{LP})$$

heißt *LP in Standardform*. ◇

Diese Darstellung eignet sich sehr gut für numerische Verfahren wie den Simplex-Algorithmus, den wir in Kapitel 6 näher untersuchen. Das folgende Lemma zeigt, dass man jedes LP in ein LP in Standardform überführen kann.

**Lemma 12.10**

Jedes LP der (allgemeinen) Form (12.1) kann in ein LP in Standardform überführt werden.

*Beweis:* Wir wenden zunächst die Transformation der zulässigen Menge aus dem Beweis von Lemma 12.7 an, so dass (12.1) äquivalent zu

$$\left. \begin{array}{l} \text{minimiere } c^\top x \\ \text{s.t. } \tilde{A}x \leq \tilde{b} \end{array} \right\} \quad (12.6)$$

ist, wobei  $\tilde{A} \in \mathbb{R}^{q \times n}$  und  $\tilde{b} \in \mathbb{R}^q$  mit  $q := 2m + p$  wie in (12.5) definiert sind. Sei nun  $x \in P := \{x \in \mathbb{R}^n : \tilde{A}x \leq \tilde{b}\}$  beliebig. Dann definieren wir

$$\hat{A} := (\tilde{A}, -\tilde{A}, I), \quad y_1 := x^{(+)}, \quad y_2 := x^{(-)} \quad \text{und} \quad y_3 := \tilde{b} - \tilde{A}x. \quad (12.7)$$

Hierbei ist  $I \in \mathbb{R}^{q \times q}$  die Einheitsmatrix und  $x^{(+)} := (\max\{0, x_i\})_{i=1}^n$  und  $x^{(-)} := (\max\{0, -x_i\})_{i=1}^n$ . Dann folgt für  $y := (y_1, y_2, y_3) \in \mathbb{R}^{2n+q}$ , dass

$$y \in \hat{P} := \{y \in \mathbb{R}^{2n+q} : \hat{A}y = \tilde{b}, y \geq 0\}.$$

Sei andersrum  $y \in \hat{P}$  beliebig. Dann folgt für  $x := y_1 - y_2$ , dass  $\tilde{A}x = \tilde{A}y_1 - \tilde{A}y_2 = \tilde{b} - y_3 \leq \tilde{b}$ , was  $x \in P$  zeigt, so dass

$$P = \{x \in \mathbb{R}^n : x = y_1 - y_2, y = (y_1, y_2, y_3) \in \hat{P}\}.$$

Daraus folgt

$$\inf_{x \in P} c^\top x = \inf_{(y_1, y_2, y_3) \in \hat{P}} c^\top (y_1 - y_2),$$

d.h. mit  $\hat{c} = (c, -c, 0)$  ist (12.1) äquivalent zu

$$\left. \begin{array}{l} \text{minimiere } \hat{c}^\top y \\ \text{s.t. } \hat{A}y = \tilde{b}, \quad y \geq 0, \end{array} \right\}$$

was gerade ein LP in Standardform ist. □

**Bemerkung 12.11**

Die in (12.7) eingeführte Variable  $y_3$  wird auch *Schlupfvariable* genannt. ◇

**Bemerkung 12.12**

Bei der im Beweis von Lemma 12.10 verwendeten Transformation wird die Variablenzahl von  $n$  auf  $2n + q = 2n + 2m + p$  erhöht. Natürlich kann (12.1) oft einfacher in Standardform überführt werden, ohne dass die Variablenzahl so drastisch vergrößert wird. Ist  $G$  z.B. einfach zu invertieren, dann kann man (12.1) durch die Setzung  $z := G^{-1}x$ ,  $\bar{A} := AG$  und  $\bar{c} := G^T c$  in ein LP in Standardform (bzgl.  $z$ ) der Dimension  $p$  umgewandelt werden.  $\diamond$

**Definition 12.13**

In Anlehnung an Definition 12.9 nennen wir einen Polyeder in der Form

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

*Polyeder in Standardform.* Eine Polyeder, der in Form

$$P = P(A, b) = \{x \in \mathbb{R}^n : Ax \leq b\}$$

gegeben ist, heißt *Polyeder in natürlicher Form*.  $\diamond$

Eine weitere Darstellung von Polyedern behandelt das folgende Lemma:

**Lemma 12.14**

Sei  $P = P(A, b)$  ein beliebiges Polyeder mit  $A \in \mathbb{R}^{m \times n}$  und  $b \in \mathbb{R}^m$ . Ist  $0 \in P$ , dann existieren  $k, \ell \in \mathbb{N} \cup \{0\}$  mit  $k + \ell = m$  und Matrizen  $G \in \mathbb{R}^{k \times n}$  und  $H \in \mathbb{R}^{\ell \times n}$ , so dass

$$P = \{x \in \mathbb{R}^n : Gx \leq \mathbf{1}, Hx \leq 0\}.$$

*Beweis:* Der Beweis erfolgt über eine einfache Skalierung. Wegen  $0 \in P$  folgt  $b \geq A0 = 0$ . Wir definieren  $\mathcal{I} := \{i \in \{1, \dots, m\} : b_i = 0\}$  und setzen  $\ell := |\mathcal{I}|$  und  $H := A_{(\mathcal{I})}$ . Weiter definieren wir  $\mathcal{I}^c = \{1, \dots, m\} \setminus \mathcal{I}$ ,  $k := |\mathcal{I}^c|$  und

$$\bar{A}_{(j)} := \frac{1}{b_j} A_{(j)}, \quad j \in \mathcal{I}^c, \quad G := (\bar{A}_{(j)})_{j \in \mathcal{I}^c}.$$

Damit ist  $Ax \leq b$  nach Konstruktion äquivalent zu

$$\begin{pmatrix} G \\ H \end{pmatrix} x \leq \begin{pmatrix} \mathbf{1} \\ 0 \end{pmatrix},$$

was die Behauptung beweist.  $\square$

**§ 13 Seitenflächen, Facetten und Ecken****Definition 13.1**

Sei  $P \subset \mathbb{R}^n$  ein Polyeder,  $c \in \mathbb{R}^n$  und  $\delta \in \mathbb{R}$ . Falls  $c^\top x = \delta$  für alle  $x \in P$  gilt, dann heißt  $c^\top x = \delta$  für  $P$  *gültige Gleichung*. Entsprechend nennen wir eine Ungleichung  $c^\top x \leq \delta$  *gültig für  $P$* , falls  $c^\top x \leq \delta$  für alle  $x \in P$ .  $\diamond$

**Definition 13.2 (Seitenfläche)**

Sei  $P \subset \mathbb{R}^n$  ein Polyeder. Eine Menge  $F \subset P$  heißt *Seitenfläche* von  $P$ , falls eine für  $P$  gültige Ungleichung  $c^\top x \leq \delta$  existiert mit

$$F = P \cap \{x \in \mathbb{R}^n : c^\top x = \delta\} = \{x \in P : c^\top x = \delta\}.$$

Die Seitenfläche  $F$  heißt dann *von  $c^\top x \leq \delta$  induzierte Seitenfläche*. Eine Seitenfläche  $F$  heißt *echt*, falls  $F \neq \emptyset$  und  $F \neq P$ .  $\diamond$

Man beachte, dass eine Seitenfläche selber wieder ein Polyeder ist. Wie schon die Darstellung eines Polyeders selbst ist auch die einer Seitenfläche nicht eindeutig, d.h. eine Seitenfläche kann durch unterschiedliche Ungleichungen induziert werden.

**Lemma 13.3**

Sei  $P \subset \mathbb{R}^n$  ein beliebiger Polyeder. Dann ist die leere Menge eine Seitenfläche von  $P$ .

*Beweis:* Die Ungleichung  $0^\top x \leq 1$  ist natürlich gültig für  $P$ . Außerdem gilt  $P \cap \{x : 0^\top x = 1\} = \emptyset$ , was die Aussage beweist.  $\square$

**Satz 13.4**

Sei  $P \neq \emptyset$  ein Polyeder. Dann gilt

- (1) Für jede Seitenfläche  $F \neq \emptyset$  gibt es ein  $c \in \mathbb{R}^n$ , so dass  $F$  gleich der Menge der Optimallösungen des linearen Programms

$$\left. \begin{array}{l} \text{maximiere } c^\top x \\ \text{s.t. } x \in P \end{array} \right\} \quad (13.1)$$

ist.

- (2) Für beliebiges  $c \in \mathbb{R}^n$  ist die Menge der Optimallösungen von (13.1) eine Seitenfläche von  $P$ .

*Beweis:* zu (1):

Da  $F$  eine nichtleere Seitenfläche ist, existieren  $c \in \mathbb{R}^n$  und  $\delta \in \mathbb{R}$  mit

$$c^\top x \leq \delta \quad \forall x \in P \quad \text{und} \quad c^\top x = \delta \quad \forall x \in F \neq \emptyset,$$

was die Aussage wegen  $F \subset P$  liefert.

zu (2):

Sei  $c \in \mathbb{R}^n$  beliebig und  $\mathbb{L}$  die Menge der Optimallösungen von (13.1), d.h.

$$\mathbb{L} = \{x \in P : c^\top x \geq c^\top y \quad \forall y \in P\}.$$

Im Fall  $\mathbb{L} = \emptyset$  folgt aus Lemma 13.3 sofort, dass  $\mathbb{L}$  eine Seitenfläche ist. Ist hingegen  $\mathbb{L} \neq \emptyset$ , dann gibt es ein  $x^* \in P$  mit

$$\delta^* := c^\top x^* \geq c^\top x \quad \forall x \in P.$$

Die Ungleichung  $c^\top x \leq \delta^*$  ist damit gültig für  $P$  und  $\mathbb{L} = \{x \in P : c^\top x = \delta^*\}$  somit eine Seitenfläche.  $\square$

**Bemerkung 13.5**

Die Aussage von Satz 13.4(2) hat weitreichende Konsequenzen. Insbesondere folgt daraus, dass man bei der Suche nach der Lösung eines LPs nicht alle zulässigen Punkte betrachten muss, sondern nur Seitenflächen des Polyeders berücksichtigen muss. Das ist eine der wesentlichen Grundideen des Simplex-Verfahrens aus Kapitel 6 unten.  $\diamond$

**Definition 13.6**

Sei  $M \neq \emptyset$  eine Menge im  $\mathbb{R}^n$ . Die *Dimension*  $\dim(M)$  von  $M$  ist definiert durch

$$\dim(M) = \max\{d \in \mathbb{N} \cup \{0\} : \exists x_0, \dots, x_d \in M$$

mit  $x_1 - x_0, x_2 - x_0, \dots, x_d - x_0$  linear unabhängig\}.  $\diamond$

**Bemerkung 13.7**

Die Dimension einer Menge  $M$  kann äquivalent auch mit Hilfe der so genannten *affinen Hülle* definiert werden. Die affine Hülle ist definiert als der Durchschnitt aller affinen Unterräume (= affin verschobene lineare Unterräume), die  $M$  enthalten. Ähnlich wie bei der konvexen und der konischen Hülle kann man zeigen, dass die affine Hülle der kleinste affine Unterraum ist, der  $M$  enthält. Die Dimension von  $M$  ist dann durch die Dimension des Unterraums gegeben, der zur affinen Hülle gehört. Man kann zeigen, dass diese Definition mit der obigen übereinstimmt, s. z.B. [?, Abschnitt 7.1].  $\diamond$

**Lemma 13.8**

Ist  $F \subset P$  eine Seitenfläche eines Polyeders  $P \in \mathbb{R}^n$ , dann gilt entweder  $F = P$  oder  $\dim(F) < \dim(P)$ .

*Beweis:* Wegen  $P = P \cap \{x : 0^\top x = 0\}$  ist  $P$  selbst eine Seitenfläche. Sei nun  $F \subset P$  mit echter Inklusion. Dann existiert ein  $\tilde{x} \in P$  mit  $\tilde{x} \notin F$ . Sei weiter  $c^\top x \leq \delta$  die Ungleichung, die  $F$  induziert, und  $d = \dim(F)$  die Dimension der Seitenfläche. Dann gibt es (maximal)  $d + 1$  Vektoren  $x^0, \dots, x^d \in F$ , so dass  $x^i - x^0$ ,  $i = 1, \dots, d$  linear unabhängig sind. Wir nehmen an,  $\tilde{x} - x^0$  sei linear abhängig von  $x^i - x^0$ ,  $i = 1, \dots, d$ , d.h. es existieren Koeffizienten  $\alpha_i$ , so dass

$$\begin{aligned} \tilde{x} - x_0 &= \sum_{i=1}^d \alpha_i (x^i - x^0) \\ \Rightarrow c^\top (\tilde{x} - x^0) &= \sum_{i=1}^d \alpha_i (\underbrace{c^\top x^i}_{=\delta} - \underbrace{c^\top x^0}_{=\delta}) = 0 \\ \Rightarrow c^\top \tilde{x} &= c^\top x^0 = \delta \quad \Rightarrow \quad \tilde{x} \in F \end{aligned}$$

im Widerspruch zur Voraussetzung. Daraus folgt, dass  $\tilde{x} - x^0$  von  $x^i - x^0$ ,  $i = 1, \dots, d$  linear unabhängig ist, und wegen  $\tilde{x}, x^0 \in P$  folgt daraus, dass  $\dim(P)$  mindestens um Eins größer ist als  $d$ .  $\square$

Aus Definition 13.6 folgt sofort, dass  $M$  einpunktig ist, falls  $\dim(M) = 0$ . Ist  $\dim(M) = 1$ , dann ist  $M$  in einer Geraden enthalten. Entsprechend definiert man:

**Definition 13.9**

Eine Seitenfläche  $F \subset \mathbb{R}^n$  eines Polyeders heißt:

- *Ecke*, falls  $\dim(F) = 0$ ,
- *Kante*, falls  $\dim(F) = 1$ ,
- *Facette*, falls  $\dim(F) = \dim(P) - 1$ .

Eine in eine Richtung unbeschränkte Kante heißt *Extremstrahl*, d.h. ist  $F$  Extremstrahl, dann existieren  $x \in P$  und  $a \neq 0$ , so dass  $F = \{x\} + \text{cone}\{a\}$ .  $\diamond$

**Lemma 13.10**

Ist  $F$  eine Facette, dann ist  $F$  in keiner anderen echten Seitenfläche von  $P$  enthalten.

*Beweis:* Sei  $F$  eine Facette, d.h.  $d := \dim(F) = \dim(P) - 1$ . Dann gibt es  $d + 1$  Vektoren  $x^0, \dots, x^d \in F$ , so dass  $x^i - x^0$ ,  $i = 1, \dots, d$  linear unabhängig sind. Nehmen wir an es gäbe eine Seitenfläche  $G \neq P$  mit  $F \subset G$  mit echter Inklusion. Dann

existiert ein  $\tilde{x} \in G$  mit  $\tilde{x} \notin F$ . Völlig analog zum Beweis von Lemma 13.8 zeigt man, dass dann  $\tilde{x} - x^0$  linear unabhängig von  $x^i - x^0$ ,  $i = 1, \dots, d$  ist, so dass  $\dim(G) = d + 1 = \dim(P)$ . Wegen  $G \neq P$  ist dies ein Widerspruch zu Lemma 13.8.  $\square$

### Bemerkung 13.11

Mit Hilfe von [?, 7.2(e)] kann man zeigen, dass auch die umgekehrte Aussage gilt, d.h. jede echte Seitenfläche, die nicht in einer anderen echten Seitenfläche enthalten ist, ist eine Facette.  $\diamond$

Wir wollen nun alternative Charakterisierungen von Ecken des Polyeders angeben. Dazu sei der Polyeder zunächst in natürlicher Form gegeben.

### Satz 13.12

Sei  $P = P(A, b)$  mit  $A \in \mathbb{R}^{m \times n}$  und  $b \in \mathbb{R}^m$  ein Polyeder und  $z \in P$  gegeben. Sei ferner die Indexmenge  $\mathcal{E}$  durch

$$\mathcal{E} = \{i \in \{1, \dots, m\} : A_{(i)}^\top z = b_i\}$$

definiert. Dann sind äquivalent:

- (1)  $z$  ist eine Ecke von  $P$ .
- (2) Es gibt ein  $c \in \mathbb{R}^n$ ,  $c \neq 0$ , so dass  $z$  der eindeutige Optimierer von  $\max_{x \in P} c^\top x$  ist.
- (3) Aus  $z = \lambda y + (1 - \lambda)x$  mit  $x, y \in P$  und  $\lambda \in (0, 1)$  folgt  $x = y = z$ .
- (4)  $\text{rang}(A_{(\mathcal{E})}) = n$ .

*Beweis:* Wir beweisen den Ringschluss (1)  $\Rightarrow$  (2)  $\Rightarrow$  (3)  $\Rightarrow$  (4)  $\Rightarrow$  (1).

zu (1)  $\Rightarrow$  (2):

analog zum Beweis von Satz 13.4(1): da  $\{z\}$  nach Definition einer Ecke eine einpunktige Seitenfläche ist, gibt es eine für  $P$  gültige Ungleichung  $c^\top x \leq \delta$  mit

$$c^\top z > c^\top x \quad \forall x \in P, x \neq z, \quad (13.2)$$

was die eindeutige Optimalität von  $z$  liefert. Ist  $P = \{z\}$ , dann kann man  $c \neq 0$  wählen, andernfalls folgt  $c \neq 0$  sofort aus (13.2).

zu (2)  $\Rightarrow$  (3):

Sei  $z$  die eindeutige Optimallösung von  $\max_{x \in P} c^\top x$  mit Optimalwert  $\gamma$ . Gilt  $z = \lambda y + (1 - \lambda)x$  mit  $x, y \in P$ ,  $x \neq y$  und  $\lambda \in (0, 1)$ , dann folgt

$$\begin{aligned} \gamma &= c^\top z = c^\top (\lambda y + (1 - \lambda)x) \\ &= \lambda c^\top y + (1 - \lambda)c^\top x \leq \lambda \gamma + (1 - \lambda)\gamma = \gamma. \end{aligned}$$

Daraus folgt aber  $c^\top y = c^\top x = \gamma$  im Widerspruch zur eindeutigen Optimalität von  $z$ .

zu (3)  $\Rightarrow$  (4):

Wir argumentieren per Kontraposition. Sei also (4) falsch, d.h.  $\text{rang}(A_{(\mathcal{E})}) < n$ . Dann existiert ein  $v \neq 0$  mit  $A_{(\mathcal{E})}v = 0$ . Nach Definition von  $\mathcal{E}$  folgt daraus

$$A_{(\mathcal{E})}(z \pm tv) = b_{\mathcal{E}} \quad \forall t \in \mathbb{R}.$$

Ist  $i \in \{1, \dots, m\} \setminus \mathcal{E}$ , dann gilt  $(Ax)_i < b_i$ , so dass

$$A_{(i)}^\top(z \pm tv) \leq b_i \quad \forall t \in [0, t_i],$$

falls  $t_i > 0$  hinreichend klein gewählt wird. Daraus folgt

$$A(z \pm tv) \leq b \quad \forall t \in [0, \bar{t}] \quad \text{mit} \quad \bar{t} := \min_{i \notin \mathcal{E}} t_i.$$

Daher liegen  $x := z + \bar{t}v$  und  $y = z - \bar{t}v$  in  $P$ . Andererseits gilt  $z = (1/2)y + (1/2)x$ , so dass sich  $z$  wegen  $x, y \in P$  und  $x, y \neq z$  als echte Konvexkombination zweier Elemente aus  $P$  darstellen lässt und somit (3) nicht gilt.

zu (4)  $\Rightarrow$  (1):

Wir definieren die Seitenfläche

$$F := \{x \in P : c^\top x = \delta\} \quad \text{mit} \quad c := \sum_{i \in \mathcal{E}} A_{(i)} \quad \text{und} \quad \delta := \sum_{i \in \mathcal{E}} b_i.$$

Aus  $A_{(\mathcal{E})}z = b_{\mathcal{E}}$  folgt sofort  $z \in F$ . Sei andererseits  $x \in F$  beliebig. Dann folgt

$$\sum_{i \in \mathcal{E}} \underbrace{(Ax - b)_i}_{\leq 0, \text{ da } x \in P} = 0 \quad \Rightarrow \quad (Ax - b)_i = 0 \quad \forall i \in \mathcal{E} \quad \Rightarrow \quad A_{(\mathcal{E})}x = b_{\mathcal{E}}.$$

Da nach Voraussetzung  $\text{rang}(A_{(\mathcal{E})}) = n$ , ist die Abbildung  $x \mapsto A_{(\mathcal{E})}x$  injektiv, so dass  $x = z$ . Weil  $x \in F$  beliebig war, folgt daraus  $F = \{z\}$ , so dass  $\{z\}$  in der Tat eine nulldimensionale Seitenfläche, also Ecke, ist.  $\square$

### Bemerkung 13.13

Satz 13.12(3) zeigt, dass man Ecken alternativ als Punkte des Polyeders definieren kann, die sich nicht als echte Konvexkombination von anderen Punkten in  $P$  schreiben lassen. Diese Eigenschaft einer Ecke ist unabhängig von der Darstellung des Polyeders, gilt also auch für Polyeder in Standardform. Daher werden Ecken in der Literatur oft auch auf diese Weise definiert. Diese Definition lässt sich auch auf allgemeine konvexe Mengen übertragen, s. ?  $\diamond$

Um die Ecken eines Polyeders in Standardform zu charakterisieren, benötigen wir die folgende

### Definition 13.14

Der Träger  $\text{supp}(x)$  eines Vektors  $x \in \mathbb{R}^n$  ist die Indexmenge definiert durch

$$\text{supp}(x) := \{i \in \{1, \dots, n\} : x_i \neq 0\} \quad \diamond$$

### Satz 13.15

Sei

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

ein Polyeder in Standardform und  $z \in P$ ,  $z \neq 0$ . Dann sind äquivalent:

- (1)  $z$  ist eine Ecke von  $P$ .
- (2)  $\text{rang}(A_{[\text{supp}(z)]}) = |\text{supp}(z)|$ , d.h. die Spaltenvektoren von  $A_{[\text{supp}(z)]} \in \mathbb{R}^{m \times |\text{supp}(z)|}$  sind linear unabhängig.

*Beweis:* Für den Beweis kürzen wir ab:  $\mathcal{S} := \text{supp}(z)$ ,  $s := |\mathcal{S}|$  und  $\mathcal{S}^c := \{1, \dots, n\} \setminus \mathcal{S}$ . Der Beweis verläuft ganz ähnlich zu dem von Satz 13.12. Diesmal zeigen wir (1)  $\Rightarrow$  (2)  $\Rightarrow$  (1).

zu (1)  $\Rightarrow$  (2):

Wie im Beweis von Satz 13.12 zeigt man, dass die Ecke  $z$  sich nicht als echte Konvexkombination von zwei Elementen aus  $P$  darstellen kann, d.h.  $z = \lambda y + (1 - \lambda)x$ ,



$x, y \in P$ ,  $\lambda \in (0, 1)$  impliziert  $z = x = y$ .<sup>1</sup> Wir nehmen wieder an (2) sei falsch, d.h.  $\text{rang}(A_{[\mathcal{S}]}) < s$ . Genau wie im Beweis von Satz 13.12 können wir dann zeigen, dass ein  $v \in \mathbb{R}^n$  und ein  $t > 0$  existieren mit  $z \pm tv \in P$ , s. Übung. Daraus folgt  $z = (1/2)(z+tv) + (1/2)(z-tv)$ , so dass  $z$  als echte Konvexkombination von Elementen aus  $P$  dargestellt werden kann, was den gewünschten Widerspruch liefert.

zu (2)  $\Rightarrow$  (1):

Auch der Beweis verläuft analog zu dem entsprechenden Teil vom Beweis von Satz 13.12. Wir betrachten die Seitenfläche

$$F := \{x \in P : -(\mathbf{1}^c)^\top x = 0\},$$

wobei  $\mathbf{1}^c \in \mathbb{R}^n$  durch  $\mathbf{1}_i^c = 0$ , falls  $i \in \mathcal{S}$ , und  $\mathbf{1}_i^c = 1$ , falls  $i \notin \mathcal{S}$ , definiert ist. Da  $-(\mathbf{1}^c)^\top x \leq 0$  eine gültige Ungleichung ist, ist  $F$  per Konstruktion eine Seitenfläche und  $z \in F$ . Sei  $x \in F$  beliebig. Wegen  $x \in P$  folgt dann  $x_{\mathcal{S}^c} = 0$  und

$$b = Ax = A_{[\mathcal{S}]}x_{\mathcal{S}}.$$

Nun ist  $\text{rang}(A_{[\mathcal{S}]}) = s$ , d.h. die Spalten von  $A_{[\mathcal{S}]} \in \mathbb{R}^{m \times s}$  sind linear unabhängig. Daraus folgt, dass  $x_{\mathcal{S}} \mapsto A_{\mathcal{S}}x_{\mathcal{S}}$  injektiv ist, und daher  $x_{\mathcal{S}} = z_{\mathcal{S}}$ . Wegen  $x_{\mathcal{S}^c} = z_{\mathcal{S}^c} = 0$  ist  $x = z$ , so dass  $F = \{z\}$ . Dies beweist, dass  $z$  in der Tat eine Ecke von  $P$  ist.  $\square$

### Bemerkung 13.16

Wegen  $\text{supp}(0) = \emptyset$  wird  $z = 0$  nicht durch Satz 13.15 erfasst. Ist aber  $0 \in P$ , dann ist  $0$  immer eine Ecke, denn wegen  $P \subset \mathbb{R}_+^n$  kann man  $0$  nicht als echte Konvexkombination von Elementen aus  $P$  darstellen.  $\diamond$

## § 14 Satz von Weyl-Minkowski

Polyeder wurden in Definition 12.1 als Schnitte endlich vieler affiner Halbräume definiert. Die Beschreibung mithilfe endlich vieler affin-linearer Ungleichungen wird auch als *äußere Beschreibung* bezeichnet. Im Folgenden zeigen wir, dass Polyeder ebensogut durch eine *innere Beschreibung* definiert werden können, diese besteht aus den Ecken und Extremalstrahlen des Polyeders. In der ganzzahligen linearen Optimierung ist diese Äquivalenz von grundlegender Bedeutung, da für konkrete Optimierungsprobleme in der Regel die innere Beschreibung bekannt ist, jedoch für die meisten Optimierungsverfahren die äußere Beschreibung benötigt wird. Der folgende Satz, den wir in diesem Abschnitt konstruktiv beweisen, zeigt, dass die beiden Beschreibungsformen prinzipiell ineinander überführt werden können.

### Satz 14.1 (Weyl-Minkowski)

Eine Menge  $P \subseteq \mathbb{R}^n$  ist ein Polyeder genau dann, wenn endliche Punkt Mengen  $X, Y \subseteq \mathbb{R}^n$  existieren mit

$$P = \text{conv}(X) + \text{cone}(Y). \quad \diamond$$

Der Beweis benötigt einige Vorbereitungen.

<sup>1</sup>Man beachte, dass der zugehörige Teil des Beweises von Satz 13.12 unabhängig von der konkreten Beschreibung von  $P$  ist und deshalb auch auf Polyeder in Standardform angewendet werden kann.

**Definition 14.2 (Projektion)**

Sei  $X \subseteq \mathbb{R}^n$  und  $k \leq n$ . Die *Projektion* von  $X$  auf die ersten  $k$  Dimensionen ist definiert durch

$$\pi_k(X) := \{x \in \mathbb{R}^k : \exists y \in \mathbb{R}^{n-k} \text{ mit } (x, y) \in X\}. \quad \diamond$$

**Satz 14.3 (Fourier-Motzkin-Elimination)**

Jede Projektion eines Polyeders ist wieder ein Polyeder.

*Beweis:* Sei  $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ . Es reicht zu zeigen, dass

$$P' = \{(x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1} : \exists x_n \in \mathbb{R} \text{ mit } x = (x_1, \dots, x_n) \in P\}$$

ein Polyeder ist. Setze dazu

$$\begin{aligned} I_+ &:= \{j \in \{1, \dots, m\} : a_{jn} > 0\} \\ I_- &:= \{j \in \{1, \dots, m\} : a_{jn} < 0\} \\ I_0 &:= \{j \in \{1, \dots, m\} : a_{jn} = 0\}. \end{aligned}$$

Nach Division der zu  $I_+$  und  $I_-$  gehörigen Ungleichungen durch  $|a_{jn}|$  wird  $P$  durch die folgenden Ungleichungen beschrieben:

$$\begin{aligned} \sum_{i=1}^{n-1} \bar{a}_{ji} x_i + x_n &\leq \bar{b}_j \text{ für } j \in I_+ \\ \sum_{i=1}^{n-1} \bar{a}_{ji} x_i - x_n &\leq \bar{b}_j \text{ für } j \in I_- \\ \sum_{i=1}^{n-1} a_{ji} x_i &\leq b_j \text{ für } j \in I_0 \end{aligned}$$

mit  $\bar{a}_{ji} := a_{ji}/a_{jn}$  und  $\bar{b}_j = b_j/a_{jn}$  für  $j \in I_+$  und  $\bar{a}_{ji} := -a_{ji}/a_{jn}$  und  $\bar{b}_j = -b_j/a_{jn}$  für  $j \in I_-$ . Also gilt  $(x_1, \dots, x_n) \in P$  genau dann, wenn

- (i)  $\sum_{i=1}^{n-1} a_{ji} x_i \leq b_j$  für alle  $j \in I_0$
- (ii)  $-\bar{b}_{j_2} + \sum_{i=1}^{n-1} \bar{a}_{j_2,i} x_i \leq x_n \leq \bar{b}_{j_1} - \sum_{i=1}^{n-1} \bar{a}_{j_1,i} x_i$  für alle  $j_1 \in I_+, j_2 \in I_-$ .

Damit ist  $(x_1, \dots, x_{n-1}) \in P'$  genau dann, wenn (i) gilt und

- (iii)  $\sum_{i=1}^{n-1} (\bar{a}_{j_1,i} + \bar{a}_{j_2,i}) x_i \leq \bar{b}_{j_1} + \bar{b}_{j_2}$  für alle  $j_1 \in I_+, j_2 \in I_-$ .

Da alle Ungleichungen in (i) und (iii) affin-linear sind, ist  $P'$  ein Polyeder, vgl. Bemerkung 12.8.  $\square$

**Bemerkung 14.4**

Die Idee aus Satz 14.3 kann im Prinzip zur Lösung eines linearen Optimierungsproblems verwendet werden: Das Problem

$$\begin{aligned} \text{minimiere} \quad & c^\top x \\ \text{s.t.} \quad & Ax \leq b \\ & x \in \mathbb{R}^n \end{aligned}$$

ist äquivalent zu

$$\begin{aligned} & \text{minimiere } z \\ & \text{s.t. } c^\top x \leq z \\ & \quad Ax \leq b \\ & \quad (z, x) \in \mathbb{R}^{n+1}. \end{aligned}$$

Setze dann  $P = \{(z, x) \in \mathbb{R}^{n+1} : -z + c^\top x \leq 0, Ax \leq b\}$  und bestimme die Projektion  $P' = \pi_1(P)$  auf die erste Dimension. Dann ist  $P' \subset \mathbb{R}$  ein Polyeder nach Satz 14.3. Als Polyeder in  $\mathbb{R}$  ist  $P'$  also ein abgeschlossenes Intervall in  $\mathbb{R}$  und der gesuchte Optimalwert ist  $\min_{z \in P'} z$ , also das linke Intervallende.

Dieses Verfahren ist jedoch sehr langsam und daher in der Praxis nicht zur Lösung linearer Optimierungsprobleme geeignet.  $\diamond$

#### Lemma 14.5

Seien  $X, Y \subset \mathbb{R}^n$  endliche Punktmenge. Dann ist  $P = \text{conv}(X) + \text{cone}(Y)$  ein Polyeder.

*Beweis:* Wir setzen  $k = |X|$  und  $\ell = |Y|$ , d.h.  $X = \{x^1, \dots, x^k\}$  und  $Y = \{y^1, \dots, y^\ell\}$ . Nach Definition von konvexer und konischer Hülle gilt  $z \in P$  genau dann, wenn  $\lambda \in \mathbb{R}^k$  und  $\mu \in \mathbb{R}^\ell$  existieren mit

$$z - \sum_{i=1}^k \lambda_i x^i - \sum_{j=1}^{\ell} \mu_j y^j = 0 \quad (14.1a)$$

$$\sum_{i=1}^k \lambda_i = 1 \quad (14.1b)$$

$$\lambda, \mu \geq 0. \quad (14.1c)$$

Definieren wir den Polyeder

$$\tilde{P} = \{w \in \mathbb{R}^{n+k+\ell} : w = (z, \lambda, \mu) \text{ erfüllt (14.1)}\},$$

dann ist

$$\pi_n(\tilde{P}) = \{z \in \mathbb{R}^n : \exists \lambda \in \mathbb{R}^k, \mu \in \mathbb{R}^\ell \text{ mit } (z, \lambda, \mu) \text{ erfüllen (14.1)}\} = P.$$

Also ist  $P$  die Projektion eines Polyeders auf die ersten  $n$  Dimensionen und damit nach Satz 14.3 selbst ein Polyeder.  $\square$

#### Definition 14.6

Sei  $X \subseteq \mathbb{R}^n$  mit  $0 \in X$ . Dann heißt

$$X^* = \{y \in \mathbb{R}^n : x^\top y \leq 1 \forall x \in X\}$$

die *polare Menge* zu  $X$ .

#### Lemma 14.7

Für alle  $X \subseteq \mathbb{R}^n$  mit  $0 \in X$  ist  $X^*$  konvex.

*Beweis:* Klar, da  $X^*$  per Definition Schnitt affiner Halbräume ist.  $\square$

#### Lemma 14.8

Sei  $C \subset \mathbb{R}^n$  eine abgeschlossene, konvexe Menge mit  $0 \in C$ . Dann gilt  $C^{**} := (C^*)^* = C$ .

*Beweis:* Sei  $x \in C$  beliebig. Dann gilt nach Definition der polaren Menge

$$x^\top y \leq 1 \quad \forall y \in C^*,$$

und daher  $C \subset C^{**} = \{w \in \mathbb{R}^n : w^\top y \leq 1 \quad \forall y \in C^*\}$ . Um  $C = C^{**}$  zu zeigen, argumentieren wir per Widerspruch. Nehmen wir also an,  $C \neq C^{**}$ . Wegen  $C \subset C^{**}$  existiert dann ein  $z \in C^{**} \setminus C$ . Da  $C$  nichtleer, konvex und abgeschlossen ist, existieren wegen des strikten Trennungssatzes aus Lemma 7.2 ein  $s \in \mathbb{R}^n$ ,  $s \neq 0$ , und  $\alpha \in \mathbb{R}$ , so dass

$$s^\top z > \alpha > s^\top x \quad \forall x \in C.$$

Wegen  $0 \in C$  folgt daraus  $\alpha > 0$ , so dass wir diese Ungleichung durch  $\alpha$  dividieren können. Mit  $\tilde{s} := s/\alpha$  ergibt sich dann

$$\tilde{s}^\top z > 1 > \tilde{s}^\top x \quad \forall x \in C. \quad (14.2)$$

Aus der zweiten Ungleichung folgt  $\tilde{s} \in C^*$ . Wegen  $z \in C^{**}$  gilt aber  $z^\top y \leq 1$  für alle  $y \in C^*$  im Widerspruch zur ersten Ungleichung in (14.2).  $\square$

### Lemma 14.9

Sei  $P = \{x \in \mathbb{R}^n : Ax \leq \mathbf{1}, Bx \leq 0\}$  mit  $A \in \mathbb{R}^{m \times n}$  und  $B \in \mathbb{R}^{p \times n}$  gegeben. Dann gilt  $P^* = \text{conv}(A^\top \cup \{0\}) + \text{cone}(B^\top)$ , wobei die Matrizen  $A^\top$  und  $B^\top$  als Mengen von Spaltenvektoren betrachtet werden.

*Beweis:* Nach Definition gilt

$$\begin{aligned} P^* &= \{b \in \mathbb{R}^n : b^\top x \leq 1 \text{ für alle } x \in \mathbb{R}^n \text{ mit } Ax \leq \mathbf{1}, Bx \leq 0\} \\ &= \{b \in \mathbb{R}^n : b^\top x \leq 1 \text{ für alle } x \in \mathbb{R}^n \text{ mit } \tilde{A}x \leq c\} \end{aligned}$$

mit

$$\tilde{A} := \begin{pmatrix} A \\ B \end{pmatrix} \quad \text{und} \quad c := \begin{pmatrix} \mathbf{1} \\ 0 \end{pmatrix}.$$

Aus der inhomogenen Version des Farkas-Lemmas in Satz 8.6 folgt nun, dass  $b \in \mathbb{R}^n$  genau dann in  $P^*$  liegt, falls ein  $u \in \mathbb{R}^{m+p}$ ,  $u \geq 0$ , existiert, so dass

- (a) entweder  $\tilde{A}^\top u = b$  und  $c^\top u \leq 1$
- (b) oder  $\tilde{A}^\top u = 0$  und  $c^\top u < 0$ .

Wegen  $c = (\mathbf{1}, 0) \geq 0$  und  $u \geq 0$  kommt der Fall (b) nicht in Frage. Daher ist  $b \in P^*$  genau dann, falls  $u = (y, z) \in \mathbb{R}^m \times \mathbb{R}^p$  existiert mit

$$y, z \geq 0, \quad (A^\top, B^\top) \begin{pmatrix} y \\ z \end{pmatrix} = b, \quad \begin{pmatrix} \mathbf{1} \\ 0 \end{pmatrix}^\top \begin{pmatrix} y \\ z \end{pmatrix} \leq 1.$$

Daraus folgt

$$\begin{aligned} P^* &= \{b \in \mathbb{R}^n : \exists y \in \mathbb{R}^m, z \in \mathbb{R}^p \text{ mit } b = A^\top y + B^\top z, y, z \geq 0, \mathbf{1}^\top y \leq 1\} \\ &= \left\{ \sum_{i=1}^m A_{(i)} y_i + \sum_{j=1}^p B_{(j)} z_j : z_j \geq 0, j = 1, \dots, p, y_i \geq 0, i = 1, \dots, m+1, \sum_{i=1}^{m+1} y_i = 1 \right\} \\ &= \text{conv}(A^\top \cup \{0\}) + \text{cone}(B^\top) \end{aligned}$$

wie behauptet.  $\square$

### Lemma 14.10

Sei  $P \subseteq \mathbb{R}^n$  ein Polyeder mit  $0 \in P$ . Dann ist auch  $P^*$  ein Polyeder mit  $0 \in P^*$ .

*Beweis:* Nach Lemma 12.14 kann jeder Polyeder, der die Null enthält, auf die Form  $P = \{x \in \mathbb{R}^n : Ax \leq \mathbf{1}, Bx \leq 0\}$  gebracht werden. Dann folgt die Behauptung mit Lemma 14.5 sofort aus Lemma 14.9. Man beachte hierbei, dass  $0 \in \text{cone}(B^\top)$  wegen der Kegeleigenschaft.  $\square$

Wir kommen nun zum Beweis vom Satz von Weyl-Minkowski.

*Beweis von Satz 14.1:* Die Richtung „ $\Leftarrow$ “ entspricht gerade der Aussage von Lemma 14.5. Wir zeigen also „ $\Rightarrow$ “ und unterscheiden drei Fälle:

- (1)  $P = \emptyset$   
Wähle dann  $X = Y = \emptyset$ .
- (2)  $0 \in P$   
Nach Lemma 14.10 ist  $P^*$  ein Polyeder mit  $0 \in P^*$ , also gilt nach Lemma 12.14, dass

$$P^* = \{x \in \mathbb{R}^n : Ax \leq \mathbf{1}, Bx \leq 0\}$$

für geeignete  $A, B$ . Wegen Lemma 14.8 ist  $P = (P^*)^*$ , so dass  $P$  auf Grund von Lemma 14.9 die gewünschte Form hat.

- (3)  $0 \notin P \neq \emptyset$   
Wähle  $t \in P$  und betrachte  $\bar{P} := P + \{-t\}$ . Dann gilt  $0 \in \bar{P}$  und damit  $\bar{P} = \text{conv}(\bar{X}) + \text{cone}(\bar{Y})$  für endliche Mengen  $\bar{X}, \bar{Y} \subseteq \mathbb{R}^n$  wie in (2). Also gilt  $P = \text{conv}(X) + \text{cone}(Y)$  mit  $X := \bar{X} + \{t\}$  und  $Y := \bar{Y}$ , vgl. Übung.  $\square$

Da Polytope nach Definition 12.1 beschränkte Polyeder sind, die konische Hülle einer nichtleeren Menge wegen der Kegeleigenschaft aber immer unbeschränkt ist, folgt aus Satz 14.1 sofort:

#### **Korollar 14.11**

Eine Menge  $P \subseteq \mathbb{R}^n$  ist ein Polytop genau dann, wenn eine endliche Punktmenge  $X \subseteq \mathbb{R}^n$  existiert mit  $P = \text{conv}(X)$ .

#### **Beispiel 14.12**

Die Kardinalität der Punktmenge  $X \subset \mathbb{R}^n$  zur Darstellung eines Polytops durch  $P = \text{conv}(X)$  ist nicht notwendigerweise polynomiell in  $n$  oder  $m$ , wobei  $m$  die Anzahl der Halbräume ist, die zur Beschreibung von  $P$  benötigt wird. Betrachte z.B. den Hyperwürfel

$$P := [0, 1]^n = \{x \in \mathbb{R}^n : 0 \leq x_i \leq 1 \text{ für } i = 1, \dots, n\}.$$

Für jedes  $X$  mit  $P = \text{conv}(X)$  gilt  $|X| \geq 2^n$ , da  $P$   $2^n$  Ecken besitzt (aber  $m = 2n$ ).

Umgekehrt gilt für

$$X := \{e_i, -e_i : i = 1, \dots, n\},$$

dass  $\text{conv}(X)$   $2^n$  Facetten besitzt, also keine äußere Beschreibung mit weniger als  $2^n$  Ungleichungen existiert, obwohl  $|X| = 2n$  gilt.



## Grundlagen der linearen Optimierung

### Inhalt

15	<b>Existenzaussagen</b>	63
16	<b>Dualitätssätze</b>	65

### § 15 Existenzaussagen

Wir kehren zurück zu linearen Optimierungsproblemen der Form (12.1). O.B.d.A. nehmen wir hier und im Folgenden an, das LP habe Standardform, also

$$\left. \begin{array}{l} \text{minimiere } c^\top x \\ \text{s.t. } Ax = b, \quad x \geq 0 \end{array} \right\} \quad (\text{LP})$$

mit  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  und  $c \in \mathbb{R}^n$ . Wie in Lemma 12.10 gesehen, kann jedes LP auf diese Form gebracht werden.

#### Satz 15.1

Die zulässige Menge  $P$  von (LP) sei nichtleer. Dann besitzt (LP) genau dann eine Lösung, wenn  $f$  auf  $P$  nach unten beschränkt ist.

*Beweis:* Die Zielfunktion sei auf  $P$  nach unten beschränkt, d.h.

$$f^* := \inf_{x \in P} c^\top x > -\infty.$$

Dann gibt es eine Infimalfolge  $\{x^n\} \subset P$  mit  $c^\top x^n \rightarrow f^*$  für  $n \rightarrow \infty$ . Wir betrachten die Folge  $\{(c^\top x^n, Ax^n - b)\} \subset \mathbb{R} \times \mathbb{R}^m$ . Da die  $x^n$  zulässig sind, gilt

$$(c^\top x^n, Ax^n - b) = (c^\top x^n, 0) \rightarrow (f^*, 0) \quad \text{für } n \rightarrow \infty.$$

Darüber hinaus gilt für alle  $n \in \mathbb{N}$ , dass

$$(c^\top x^n, Ax^n - b) \in \{(c^\top x, Ax - b) \in \mathbb{R} \times \mathbb{R}^m : x \geq 0\} = \text{cone} \left\{ \begin{pmatrix} c^\top \\ A \end{pmatrix} \right\} - \left\{ \begin{pmatrix} 0 \\ b \end{pmatrix} \right\} =: K.$$

Nach Lemma 8.1 ist die konische Hülle einer endlichen Punktmenge abgeschlossen, also auch die um  $(0, b)$  verschobene konische Hülle. Daraus folgt  $(f^*, 0) \in K$ , d.h. es existiert ein  $x^* \in \mathbb{R}^n$  mit  $x^* \geq 0$ ,  $Ax^* - b = 0$  (und somit  $x^* \in P$ ) und  $c^\top x^* = f^*$ . Da  $f^*$  der Infimalwert war, ist  $x^*$  optimal.

Die umgekehrte Implikation ist klar. □

Als nächstes zeigen wir, dass das Optimum –falls existent– u.a. in einer Ecke angenommen wird. Dazu zeigen wir zunächst:

**Lemma 15.2**

Ein nichtleerer Polyeder  $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$  in Standardform hat mindestens eine Ecke.

*Beweis:* Ist  $0 \in P$ , dann ist  $0$  nach Bemerkung 13.16 eine Ecke. Im Fall  $0 \notin P$ , wählen wir ein  $x^* \in P$  mit minimalem Träger  $\mathcal{S}^* := \text{supp}(x^*) = \{i \in \{1, \dots, n\} : x_i^* > 0\}$ , d.h.  $|\mathcal{S}^*|$  sei minimal. (Wegen  $0 \notin P$  ist  $|\mathcal{S}^*| > 0$ .) Wir zeigen, dass  $\text{rang}(A_{[\mathcal{S}^*]}) = |\mathcal{S}^*|$ , woraus nach Satz 13.15 folgt, dass  $x^*$  eine Ecke von  $P$  ist. Dazu argumentieren wir per Widerspruch und nehmen an, die Spalten von  $A_{[\mathcal{S}^*]}$  seien linear abhängig. Dann existieren Zahlen  $\{\alpha_i\}_{i \in \mathcal{S}^*}$ , von denen mindestens eine ungleich Null ist, mit

$$\sum_{i \in \mathcal{S}^*} \alpha_i A_{[i]} = 0. \quad (15.1)$$

Wir nehmen an, dass mindestens eines der  $\alpha_i$  echt kleiner Null ist (sonst multipliziere (15.1) mit  $-1$ ). Weiter setzen wir

$$\delta := \max_{i \in \mathcal{S}^*} \{-x_i^*/\alpha_i : \alpha_i < 0\} = \max_{1 \leq i \leq n} \{-x_i^*/\alpha_i : x_i^* > 0, \alpha_i < 0\}$$

Mit dieser Konstruktion folgt  $x_i^* + \delta \alpha_i \geq 0$  für alle  $i \in \mathcal{S}^*$  und die Existenz mindestens eines Index  $\ell \in \mathcal{S}^*$ , so dass

$$x_\ell^* + \delta \alpha_\ell = 0. \quad (15.2)$$

Dann definieren wir

$$\hat{x}_i := \begin{cases} x_i^* + \delta \alpha_i, & i \in \mathcal{S}^* \\ x_i^*, & i \in \{1, \dots, m\} \setminus \mathcal{S}^*. \end{cases}$$

Für das so konstruierte  $\hat{x}$  gilt

$$\hat{x} \geq 0, \quad \text{und} \quad A\hat{x} = Ax^* + \delta \sum_{i \in \mathcal{S}^*} \alpha_i A_{[i]} = b,$$

so dass  $\hat{x} \in P$ . Wegen (15.2) ist aber  $|\text{supp}(\hat{x})| < |\text{supp}(x^*)|$  im Widerspruch zur Voraussetzung.  $\square$

**Lemma 15.3**

Sei  $P$  ein Polyeder in Standardform. Dann ist die Anzahl der Ecken von  $P$  endlich.

*Beweis:* Nach Satz 13.15 wird jede Ecke  $x$  von  $P$  durch

$$\text{rang}(A_{[\mathcal{S}]}) = |\mathcal{S}| \quad \text{mit} \quad \mathcal{S} = \text{supp}(x) \quad (15.3)$$

charakterisiert. Nach der Definition des Trägers ist  $x_i$  für  $i \in \{1, \dots, n\} \setminus \mathcal{S}$  eindeutig festgelegt (nämlich gleich Null). Da  $y \mapsto A_{[\mathcal{S}]}y$  wegen (15.3) injektiv ist, ist  $x_{\mathcal{S}}$  ebenfalls eindeutig bestimmt, so dass die Ecke eindeutig durch den Träger bestimmt wird. Da die Anzahl der möglichen Träger als Teilmengen von  $\{1, \dots, n\}$  endlich ist, ist somit auch die Zahl der Ecken endlich.  $\square$

**Satz 15.4**

Es sei ein LP in Standardform (LP) gegeben. Ist das LP lösbar, dann wird das Minimum auch in mindestens einer Ecke angenommen.

*Beweis:* Nach Voraussetzung ist

$$f^* := \inf_{x \in P} c^\top x > -\infty.$$



Wir definieren den Polyeder

$$P^* = P \cap \{x \in \mathbb{R}^n : c^\top x = f^*\} = \left\{ x \in \mathbb{R}^n : \begin{pmatrix} c^\top \\ A \end{pmatrix} x = \begin{pmatrix} f^* \\ b \end{pmatrix}, x \geq 0 \right\}.$$

Da das LP nach Voraussetzung lösbar ist, ist  $P^* \neq \emptyset$ . Nach Lemma 15.2 existiert daher eine Ecke  $x^* \in P^*$ . Wir zeigen per Widerspruchsbeweis, dass  $x^*$  dann auch eine Ecke von  $P$  ist. Angenommen, dass sei falsch. Nach Satz 13.12 existieren dann  $z, y \in P$  und  $\lambda \in (0, 1)$  mit

$$x^* = \lambda y + (1 - \lambda)z, \quad z, y \neq x^*. \quad (15.4)$$

Da  $y, z \in P$  zulässig sind und  $f^*$  der Optimalwert des LPs ist, folgt

$$\lambda \underbrace{(c^\top y - f^*)}_{\geq 0} + (1 - \lambda) \underbrace{(c^\top z - f^*)}_{\geq 0} = c^\top (\lambda y + (1 - \lambda)z) - f^* = c^\top x^* - f^* = 0$$

Wegen  $0 < \lambda < 1$  impliziert dies  $c^\top y = c^\top z = f^*$ , also  $y, z \in P^*$ . Da  $x^*$  eine Ecke von  $P^*$  ist, folgt wiederum nach Satz 13.12, dass  $x^* = y = z$  im Widerspruch zu (15.4). Daher ist  $x^*$  auch eine Ecke von  $P$ , und wegen  $x^* \in P^* \Rightarrow c^\top x^* = f^*$  ist  $x^*$  auch optimal.  $\square$

Satz 15.4 zeigt, dass die Optimallösung eines LPs –wenn sie existiert– immer in einer Ecke angenommen wird. Um das LP zu lösen, könnte man also alle Ecken berechnen und ausprobieren, wo der beste Zielfunktionswert auftritt. Da die Anzahl der Ecken eines Polyeders nach Lemma 15.3 endlich ist, terminiert dieses Verfahren zwar endlich, ist aber nicht sehr effizient und in der Praxis nicht einsetzbar. Dennoch steckt darin der Kern eines der wichtigsten Verfahren zur Lösung linearer Probleme, des *Simplexverfahrens*. Man startet dabei mit einer zulässigen Ecke, die der Durchschnitt von  $n$  Hyperebenen ist, und versucht dann systematisch, die Lösung zu verbessern, indem man entlang einer Kante zu einer benachbarten Ecke wandert. Dieses Verfahren werden wir in Kapitel 6 noch genauer untersuchen. Zuvor jedoch wenden wir die Ergebnisse aus Abschnitt 10 auf den Spezialfall eines LPs an, um u.a. geeignete Abbruchbedingungen für das Simplexverfahren zu erhalten.

## § 16 Dualitätssätze

Bevor wir die Dualitätsergebnisse aus Abschnitt 10 auf ein LP in Standardform anwenden, wollen wir uns überlegen, wie das zugehörige Dualproblem aussieht. Dazu betrachten wir wieder (LP) mit gegebenen  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  und  $c \in \mathbb{R}^n$ .<sup>1</sup>

### Satz 16.1

Das Dualproblem zu einem LP in Standardform ist äquivalent zu:

$$\left. \begin{array}{l} \text{maximiere } b^\top \nu \\ \text{s.t. } A^\top \nu \leq c. \end{array} \right\} \quad (\text{DLP})$$

### Bemerkung 16.2

Wir beobachten, dass die zulässige Menge des Dualproblems durch einen Polyeder in natürlicher Form gegeben ist. Ein LP dieser Struktur nennt man auch *LP in natürlicher Form*.  $\diamond$

<sup>1</sup>Im Gegensatz zu Abschnitt 10 ist die Dimension der Gleichungsnebenbedingungen hier  $m$  und nicht  $p$ .

*Beweis von Satz 16.1:* Nach Definition 11.1 lautet das duale Problem zu (LP):

$$\sup_{\substack{\mu \in \mathbb{R}^m \\ \lambda \geq 0}} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu). \quad (16.1)$$

mit der Lagrange-Funktion

$$\begin{aligned} \mathcal{L}(x, \lambda, \mu) &= c^\top x - \lambda^\top x + (b - Ax)^\top \mu \\ &= (c - \lambda - A^\top \mu)^\top x + b^\top \mu. \end{aligned} \quad (16.2)$$

Daraus folgt

$$\inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu) = -\infty,$$

falls  $c - \lambda - A^\top \mu \neq 0$ . Bei der äußeren Maximierung in (16.1) müssen also nur die  $\lambda$  und  $\mu$  berücksichtigt werden, für die  $\lambda = c - A^\top \mu$  gilt. Für diese aber ist das Infimum von  $\mathcal{L}$  über alle  $x$  durch  $b^\top \mu$  gegeben, so dass

$$(16.1) \Leftrightarrow \left\{ \begin{array}{l} \text{maximiere } b^\top \mu \\ \text{s.t. } \lambda \geq 0, \quad A^\top \mu + \lambda = c. \end{array} \right\} \quad (16.3)$$

Da  $\lambda$  gar nicht mehr im Zielfunktional auftaucht, können wir es aus dem Problem eliminieren und erhalten mit der Substitution  $\nu := \mu$  gerade (DLP).  $\square$

### Bemerkung 16.3

Die Äquivalenz aus Satz 16.1 ist wie folgt gemeint:

Ist  $\nu^* \in \mathbb{R}^m$  eine Lösung von (DLP), dann lösen  $\mu^* = \nu^*$  und  $\lambda^* = c - A^\top \nu^*$  das Problem (16.1). Sind  $(\mu^*, \lambda^*) \in \mathbb{R}^m \times \mathbb{R}^n$  eine Lösung von (16.1), dann gilt  $\lambda^* = c - A^\top \mu^*$ , und  $\nu^* = \mu^*$  löst (DLP).  $\diamond$

### Bemerkung 16.4

Wie im Beweis von Satz 16.1 gesehen gilt:

$$\sup(D) = \sup_{\substack{\mu \in \mathbb{R}^m \\ \lambda \geq 0}} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu) = \sup_{\substack{A^\top \mu + \lambda = c \\ \lambda \geq 0}} b^\top \mu = \sup_{A^\top \nu \leq c} b^\top \nu = \sup(\text{DLP}),$$

was die Gleichheit der Optimalwerte von (16.1) und (DLP) liefert.  $\diamond$

### Bemerkung 16.5

Gelegentlich wird in der Literatur auch das Problem auf der rechten Seite von (16.3) als duales LP eingeführt. Wir werden es im Zusammenhang mit dem dualen Simplex-Verfahren benötigen, bezeichnen die Optimierungsvariablen dann aber mit  $\nu$  und  $s$ , d.h.

$$\left. \begin{array}{l} \text{maximiere } b^\top \nu \\ \text{s.t. } A^\top \nu + s = c, \quad s \geq 0, \end{array} \right\} \quad (16.4)$$

da  $s$  die Schlupfvariable zu  $\nu$  in der Darstellung (DLP) ist. Diese Form des Dualproblems wird auch als „duale Standardform“ bezeichnet (obwohl es eigentlich kein Problem in Standardform ist).  $\diamond$

Das folgende Resultat zeigt, dass man im Fall der linearen Optimierung bei erneuter „Dualisierung“ wieder beim primalen Problem landet:

### Proposition 16.6

Das zu (DLP) duale Problem ist äquivalent zu (LP).

*Beweis:* Die Argumentation ist dieselbe wie im Beweis von Satz 16.1. Zunächst schreiben wir (DLP) um in ein Minimierungsproblem:

$$\left. \begin{array}{l} \text{minimiere} \quad -b^\top \nu \\ \text{s.t.} \quad A^\top \nu \leq c. \end{array} \right\} \quad (16.5)$$

(vgl. auch Bemerkung 1.2). Das zugehörige duale Problem lautet

$$\sup_{\omega \geq 0} \inf_{\nu \in \mathbb{R}^n} -b^\top \nu + (A^\top \nu - c)^\top \omega. \quad (16.6)$$

Das innere Infimum ist gleich  $-\infty$ , falls  $A\omega - b \neq 0$ . Daher kommen nur  $\omega \in \mathbb{R}^n$  mit  $A\omega = b$  für die äußere Maximierung in Betracht. Für diese ist die Zielfunktion in (16.6) gleich  $-c^\top \omega$ , so dass

$$\begin{aligned} (16.6) &\Leftrightarrow \left\{ \begin{array}{l} \text{maximiere} \quad -c^\top \omega \\ \text{s.t.} \quad A\omega = b, \quad \omega \geq 0 \end{array} \right. \\ &\Leftrightarrow \left\{ \begin{array}{l} \text{minimiere} \quad c^\top x \\ \text{s.t.} \quad Ax = b, \quad x \geq 0, \end{array} \right. \end{aligned}$$

was gerade dem primalen Problem (LP) entspricht.  $\square$

Primale und duale Lösung hängen eng mit den KKT-Bedingungen für ein LP in Standardform zusammen:

**Satz 16.7 (KKT-Bedingungen für LPs)**

Ein Punkt  $x^* \in \mathbb{R}^n$  ist genau dann optimal für (LP), falls ein  $\nu^* \in \mathbb{R}^m$  existiert mit

$$\begin{aligned} Ax^* &= b, \quad x^* \geq 0 \\ A^\top \nu^* &\leq c \\ (c - A^\top \nu^*)^\top x^* &= 0. \end{aligned} \quad (16.7)$$

Umgekehrt ist  $\nu^* \in \mathbb{R}^m$  genau dann optimal für das duale Problem (DLP), falls es ein  $x^* \in \mathbb{R}^n$  gibt, so dass (16.7) erfüllt ist.

Zusammengefasst gilt also:

$$x^* \text{ löst (LP)} \Leftrightarrow (16.7) \Leftrightarrow \nu^* \text{ löst (DLP)}.$$

*Beweis:* Sei  $x^*$  eine optimale Lösung von (LP). Wegen  $x^* \in P$  ist die zulässige Menge nicht leer. Da die Nebenbedingungen linear-affin sind, ist die Slater-Bedingung (10.10) damit erfüllt, vgl. auch Bemerkung 10.5. Nach Satz 10.3 folgt daraus die Existenz von  $\lambda^* \in \mathbb{R}^n$  und  $\mu^* \in \mathbb{R}^m$ , so dass die folgenden KKT-Bedingungen erfüllt sind:

$$\begin{aligned} Ax^* &= b, \quad x^* \geq 0 \quad (\text{Zulässigkeit}) \\ c - A^\top \mu^* - \lambda^* &= 0 \quad (\text{Gradientengl.}) \\ \lambda^* &\geq 0, \quad (\lambda^*)^\top x^* = 0 \quad (\text{kompl. Schlupfbdg.}) \end{aligned} \quad (16.8)$$

Wieder können wir  $\lambda^*$  aus dem System eliminieren und erhalten

$$\begin{aligned} Ax^* &= b, \quad x^* \geq 0 \\ c - A^\top \mu^* &\geq 0 \\ (c - A^\top \mu^*)^\top x^* &= 0. \end{aligned}$$

Die Substitution  $\nu^* = \mu^*$  liefert das gewünschte Ergebnis. Ist umgekehrt  $(x^*, \nu^*) \in \mathbb{R}^n \times \mathbb{R}^m$  eine Lösung von (16.7), dann erfüllt  $x^*$  zusammen mit  $\mu^* = \nu^*$  und  $\lambda^* = c - A^\top \nu^*$  das KKT-System (16.8). Nach Korollar 10.12 ist  $x^*$  damit optimal für (LP).

Die KKT-Bedingungen des dualen Problems (DLP), also

$$\left. \begin{array}{l} \text{minimiere} \quad -b^\top \nu \\ \text{s.t.} \quad A^\top \nu \leq c, \end{array} \right\}$$

sind durch

$$\begin{aligned} A^\top \nu^* &\leq c && \text{(Zulässigkeit)} \\ -b + Ax^* &= 0 && \text{(Gradientengl.)} \\ x^* &\geq 0, \quad (x^*)^\top (A^\top \nu^* - c) = 0 && \text{(kompl. Schlupfbdg.)} \end{aligned}$$

gegeben, was gerade (16.7) entspricht. Nach Satz 10.3 und Korollar 10.12 sind sie notwendig und hinreichend für Optimalität, wie im Satz behauptet.  $\square$

Der obige Satz zeigt auch, dass das primale Problem genau dann lösbar ist, wenn das duale eine Lösung besitzt.

### **Satz 16.8 (Starke Dualität für LPs)**

Folgende Aussagen sind äquivalent:

- (1) Die zulässige Menge des primalen Problems ist nichtleer und es gilt

$$\inf_{Ax=b, x \geq 0} c^\top x > -\infty.$$

- (2) Die zulässige Menge des dualen Problems (DLP) ist nichtleer und es gilt

$$\sup_{A^\top \nu \leq c} b^\top \nu < \infty.$$

- (3) Die zulässigen Mengen von (LP) und (DLP) sind nichtleer.

- (4) Das primale und das duale Problem sind lösbar und es gilt  $\inf(\text{LP}) = \sup(\text{DLP})$ , d.h. die Optimalwerte stimmen überein.

*Beweis:* Es ist klar, dass (4) die ersten drei Aussagen impliziert. Wir reicht deshalb aus zu zeigen, dass aus den ersten drei Aussagen jeweils (4) folgt.

zu (1)  $\Rightarrow$  (4):

Gilt (1), dann ist (LP) nach Satz 15.1 lösbar. Da die Nebenbedingungen linear-affin sind, ist die Slater-Bedingung automatisch erfüllt, so dass das zugehörige duale Problem nach Satz 10.21 ebenfalls lösbar ist und keine Dualitätslücke auftritt, vgl. auch Bemerkung 10.23. Da das Dualproblem nach Satz 16.1 äquivalent zu (DLP) ist, folgt die Behauptung. Man beachte hierbei, dass nach Satz 10.21 und Bemerkung 16.4

$$\inf_{Ax=b, x \geq 0} c^\top x = \inf(P) = \sup(D) = \sup_{A^\top \nu \leq c} b^\top \nu$$

gilt, was die Gleichheit der Optimalwerte liefert.

zu (3)  $\Rightarrow$  (4):

Wir zeigen, dass (3) Aussage (1) impliziert, woraus mit dem bisher Gezeigten folgt, dass auch (4) gilt. Ist (3) erfüllt, dann existieren für (LP) bzw. (DLP) zulässige

$\hat{x} \in \mathbb{R}^n$  und  $\hat{\nu} \in \mathbb{R}^m$ . Damit ist die zulässige Menge von (LP) nichtleer. Aus der schwachen Dualität in 10.16 folgt außerdem

$$\inf_{Ax=b, x \geq 0} c^\top x = \inf(P) \geq \sup(D) = \sup_{A^\top \nu \leq c} b^\top \nu \geq b^\top \hat{\nu} > -\infty.$$

Daraus folgt (1) und damit (4).

zu (2)  $\Rightarrow$  (4):

Nach Lemma 12.10 lässt sich das duale Problem (DLP) in ein äquivalentes Problem in Standardform überführen, welches wegen (2) nach Satz 15.1 lösbar ist, vgl. Übung. Nach Satz 16.7 ist damit auch (LP) lösbar, so dass erst recht (1) erfüllt ist, was wie bereits bewiesen (4) impliziert.  $\square$



## Simplex-Algorithmus

### Inhalt

17	<b>Simplex-Algorithmus in primaler Form</b>	71
18	<b>Praktische Umsetzung des Simplex-Algorithmus</b>	78
19	<b>Reoptimierung</b>	80

### § 17 Simplex-Algorithmus in primaler Form

Wir wollen wie bereits angedeutet auf Basis von Satz 15.4 einen Algorithmus zur Lösung von LPs in Standardform designen, indem wir gezielt die Ecken des Polyeders, der die zulässige Menge beschreibt, auf Optimalität untersucht. Dazu benötigen wir:

- (1) eine effiziente Beschreibung der Ecken,
- (2) einen Update der aktuellen Ecke, der nicht planlos eine neue Ecke sucht, sondern eine Ecke wählt, die den Wert des Zielfunktional verbessert,
- (3) eine Abbruchbedingung, die die Optimalität einer Ecke erkennt oder registriert, dass die Zielfunktion auf der zulässigen Menge unbeschränkt und das Problem somit nicht lösbar ist.

Im Folgenden sei  $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$  mit  $A \in \mathbb{R}^{m \times n}$ ,  $m \leq n$ , und  $b \in \mathbb{R}^m$  immer ein Polyeder in Standardform. Im Fall  $m > n$  ist entweder  $P = \emptyset$  (falls  $b$  nicht im Bild von  $A$  liegt), oder das Gleichungssystem  $Ax = b$  enthält mindestens  $m - n$  redundante Gleichungen, die wir bei der Beschreibung von  $P$  weglassen können.

Wir kümmern uns zunächst um eine einfache Beschreibung der Ecken. Dazu definieren wir:

#### Definition 17.1 (Basis)

Sei  $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$  mit  $A \in \mathbb{R}^{m \times n}$ ,  $m \leq n$ , und  $b \in \mathbb{R}^m$  ein Polyeder in Standardform. Ferner sei  $\mathcal{B} \subset \{1, \dots, n\}$  eine Indexmenge mit  $|\mathcal{B}| = m$ . Die „Restindexmenge“ bezeichnen wir mit  $\mathcal{N} = \{1, \dots, n\} \setminus \mathcal{B}$ . Die Indexmenge  $\mathcal{B}$  heißt *Basis*, falls  $A_{[\mathcal{B}]} \in \mathbb{R}^{m \times m}$  regulär ist. Die Restmenge  $\mathcal{N}$  nennen wir *Nichtbasis*. Die Matrix  $A_{[\mathcal{B}]}$  heißt *Basismatrix*. Ist  $\mathcal{B}$  eine Basis, dann heißt der Vektor  $x \in \mathbb{R}^n$  mit  $x_{\mathcal{B}} = A_{[\mathcal{B}]}^{-1}b$  und  $x_{\mathcal{N}} = 0$  *Basisvektor*. Die Komponenten  $x_j$ ,  $j \in \mathcal{B}$ , heißen *Basisvariablen*, die Komponenten  $x_k$ ,  $k \in \mathcal{N}$ , nennen wir *Nichtbasisvariablen*. Eine Basis heißt *zulässig*, falls  $x_{\mathcal{B}} \geq 0$ . Der zugehörige Basisvektor heißt dann ebenfalls zulässig. Eine zulässige Basis wird als *nicht degeneriert* bezeichnet, falls  $x_{\mathcal{B}} > 0$ , andernfalls *degeneriert*.  $\diamond$

#### Bemerkung 17.2

Zu einer gegebenen Basis ist der Basisvektor nach Konstruktion eindeutig bestimmt.

Umgekehrt kann es zu einem gegebenen Vektor unterschiedliche Basen geben, s. Übung.  $\diamond$

**Satz 17.3**

Sei  $P$  ein Polyeder in Standardform, dessen zugehörige Matrix  $A \in \mathbb{R}^{m \times n}$  vollen Zeilenrang, d.h.  $\text{rang}(A) = m \leq n$ . Dann ist  $x \in P$  genau dann eine Ecke von  $P$ , falls  $x$  ein zulässiger Basisvektor ist.

*Beweis:* Sei  $x \in P$  eine Ecke und  $\mathcal{I} = \text{supp}(x)$ . Nach Satz 13.15 ist dann  $\text{rang}(A_{[\mathcal{I}]}) = |\mathcal{I}|$ , d.h. die Spalten von  $A_{[\mathcal{I}]}$  sind linear unabhängig. Wegen  $\text{rang}(A) \leq m$  muss  $|\mathcal{I}| \leq m$  sein. Ist  $|\mathcal{I}| = m$ , dann wählen wir die Basis  $\mathcal{B} = \mathcal{I}$ . Andernfalls gibt es wegen  $\text{rang}(A) = m$  eine Indexmenge  $\mathcal{J} \subset \{1, \dots, n\} \setminus \mathcal{I}$  mit  $|\mathcal{J}| = m - |\mathcal{I}|$ , so dass mit  $A_{[\mathcal{B}]} \in \mathbb{R}^{m \times m}$  mit  $\mathcal{B} := \mathcal{I} \cup \mathcal{J}$  regulär ist. (Das bedeutet, wir füllen  $A_{[\mathcal{I}]}$  mit linear unabhängigen Spalten  $A_{[j]}$ ,  $j \in \mathcal{J}$ , zu einer invertierbaren Matrix auf.)<sup>1</sup> Damit ist  $\mathcal{B}$  in beiden Fällen nach Definition eine Basis und wegen

$$b = Ax = (A_{[\mathcal{B}]}, A_{[\mathcal{N}]}) \begin{pmatrix} x_{\mathcal{B}} \\ x_{\mathcal{N}} \end{pmatrix} = (A_{[\mathcal{B}]}, A_{[\mathcal{N}]}) \begin{pmatrix} x_{\mathcal{B}} \\ 0 \end{pmatrix} = A_{[\mathcal{B}]}x_{\mathcal{B}}$$

ist  $x$  der zugehörige Basisvektor. Hierbei ist  $\mathcal{N} := \{1, \dots, n\} \setminus \mathcal{B} \subset \{1, \dots, n\} \setminus \mathcal{I}$ , so dass  $x_{\mathcal{N}} = 0$ . Wegen  $x \in P$  ist  $\mathcal{B}$  zulässig.

Sei umgekehrt  $x$  ein zulässiger Basisvektor mit Basis  $\mathcal{B}$ . Dann hat  $A_{[\mathcal{B}]}$  vollen Rang, also auch  $A_{[\text{supp}(x)]}$  (wegen  $\text{supp}(x) \subset \mathcal{B}$  nach Definition des Basisvektors), woraus nach Satz 13.15 folgt, dass  $x$  eine Ecke von  $P$  ist.  $\square$

**Bemerkung 17.4**

Die Bedingung  $\text{rang}(A) = m$  ist eigentlich nicht sonderlich einschränkend, da im Fall  $\text{rang}(A) < m$  entweder  $P = \emptyset$  gilt oder redundante Gleichungen bei der Beschreibung von  $P$  auftreten. In der Praxis ist  $m$  jedoch häufig sehr groß, so dass die Bestimmung des Ranges von  $A$  bzw. eine Aussortierung redundanter Gleichungen schwierig ist.  $\diamond$

Als nächstes überlegen wir uns, wie man ausgehend von einer Ecke zu einer besseren Ecke kommt, d.h. einer Ecke mit einem kleineren Zielfunktionswert. Betrachten wir dazu das Zielfunktional für eine gegebene zulässige Basis  $\mathcal{B}$  mit Basisvektor  $x$ . Dann gilt für einen beliebigen zulässigen Punkt  $y \in P$ , dass

$$\begin{aligned} c^{\top} y &= c_{\mathcal{B}}^{\top} y_{\mathcal{B}} + c_{\mathcal{N}}^{\top} y_{\mathcal{N}} \\ &= c_{\mathcal{B}}^{\top} (A_{[\mathcal{B}]}^{-1} b - A_{[\mathcal{B}]}^{-1} A_{[\mathcal{N}]} y_{\mathcal{N}}) + c_{\mathcal{N}}^{\top} y_{\mathcal{N}} \quad (\text{wegen } b = Ay = A_{[\mathcal{B}]} y_{\mathcal{B}} + A_{[\mathcal{N}]} y_{\mathcal{N}}) \\ &= c_{\mathcal{B}}^{\top} x_{\mathcal{B}} + (c_{\mathcal{N}}^{\top} - c_{\mathcal{B}}^{\top} A_{[\mathcal{B}]}^{-1} A_{[\mathcal{N}]}) y_{\mathcal{N}} \quad (\text{wegen } b = Ax = A_{[\mathcal{B}]} x_{\mathcal{B}} + A_{[\mathcal{N}]} \underbrace{x_{\mathcal{N}}}_{=0}) \\ &= c^{\top} x + \underbrace{(c_{\mathcal{N}}^{\top} - A_{(\mathcal{N})}^{\top} A_{(\mathcal{B})}^{-\top} c_{\mathcal{B}}^{\top})}_{=: s_{\mathcal{N}}}^{\top} y_{\mathcal{N}} \end{aligned} \quad (17.1)$$

(vgl. (3.1) für die letzte Umformung). Den Vektor  $s_{\mathcal{N}}$  bezeichnen wir als *reduzierte Kosten*. Wegen  $y_{\mathcal{N}} \geq 0$  (da zulässig) kann die Zielfunktion nur verkleinert werden, falls  $s_j < 0$  für mindestens ein  $j \in \mathcal{N}$ , so dass gilt:

<sup>1</sup>Hierdurch wird auch der Fall  $x = 0$  abgedeckt, was nach Bemerkung 13.16 eine Ecke ist. Hier gilt  $\mathcal{I} = \emptyset$ , und wir können beliebige linear unabhängige Zeilen von  $A$  als Basismatrix wählen.



**Lemma 17.5**

Sei  $\mathcal{B}$  eine zulässige Basis mit Basisvektor  $x$ . Wenn für die reduzierten Kosten

$$s_{\mathcal{N}} = c_{\mathcal{N}} - A_{(\mathcal{N})}^{\top} A_{(\mathcal{B})}^{-\top} c_{\mathcal{B}} \geq 0$$

gilt, dann ist  $x$  eine Lösung des LPs.  $\diamond$

Ist  $s_{\mathcal{N}} \not\geq 0$ , dann wählen wir ein  $j \in \mathcal{N}$  mit  $s_j < 0$ . Für die nächste Ecke (mit verkleinertem Zielfunktionswert)  $x^+$  machen wir folgenden Ansatz:

$$x_j^+ = \gamma \geq 0, \quad x_i^+ = 0 \quad \forall i \in \mathcal{N} \setminus \{j\}.$$

Jetzt sind  $\gamma$  und die verbleibenden Komponenten von  $x^+$  (also die in  $\mathcal{B}$ ) so einzustellen, dass  $x^+$  die gewünschten Eigenschaften hat, nämlich

- (1) die Zielfunktion verringert,
- (2) zulässig ist,
- (3) eine Ecke ist.

zu (1):

Nach (17.1) gilt

$$c^{\top} x^+ = c^{\top} x + s_{\mathcal{N}}^{\top} x_{\mathcal{N}}^+ = c^{\top} x + s_j \gamma \quad (17.2)$$

und damit eine Verbesserung der Zielfunktion für alle  $\gamma > 0$ .

zu (2):

Die Gleichungsnebenbedingungen liefern:

$$\begin{aligned} Ax^+ &\stackrel{!}{=} b \\ \Leftrightarrow A_{[\mathcal{B}]} x_{\mathcal{B}}^+ + A_{[\mathcal{N}]} x_{\mathcal{N}}^+ &\stackrel{!}{=} b \\ \Leftrightarrow x_{\mathcal{B}}^+ &\stackrel{!}{=} A_{[\mathcal{B}]}^{-1} (b - \gamma A_{[\mathcal{B}]}^{-1} A_{[\mathcal{N}]} A_{[j]}) \\ \Leftrightarrow x_{\mathcal{B}}^+ &\stackrel{!}{=} x_{\mathcal{B}} - \underbrace{\gamma A_{[\mathcal{B}]}^{-1} A_{[j]}}_{=: w_{\mathcal{B}}}. \end{aligned} \quad (17.3)$$

Damit werden alle Komponenten von  $x^+$  durch eine Wahl von  $\gamma$  festgelegt. In Abhängigkeit des Vorzeichens von  $w_{\mathcal{B}}$  verfährt man nun weiter.

**Lemma 17.6**

Sei  $s_j < 0$  für ein  $j \in \mathcal{N}$  und  $w_{\mathcal{B}} \leq 0$ . Dann ist das LP unbeschränkt, also nicht lösbar.

*Beweis:* Wegen  $x_i^+ = 0, i \in \mathcal{N} \setminus \{j\}, x_j^+ = \gamma \geq 0$  und  $x_{\mathcal{B}}^+ = x_{\mathcal{B}} - \gamma w_{\mathcal{B}} \geq x_{\mathcal{B}} \geq 0$  erfüllt  $x^+$  für jedes  $\gamma \geq 0$  die Ungleichungsnebenbedingungen. Wegen der Konstruktion von  $x_{\mathcal{B}}^+$  in (17.3) sind auch die Gleichungsnebenbedingungen erfüllt, d.h.  $x^+$  ist für jedes  $\gamma \geq 0$  zulässig. Für die Zielfunktion in  $x^+$  gilt nach (17.2) aber

$$c^{\top} x^+ = c^{\top} x + \gamma s_j \rightarrow -\infty \quad \text{für } \gamma \rightarrow \infty,$$

so dass die Zielfunktion auf  $P$  in der Tat nach unten unbeschränkt ist.  $\square$

Ist also  $w_{\mathcal{B}} \leq 0$  brauchen wir nicht weiter nach einer neuen Ecke zu suchen, das LP ist nicht lösbar. Andernfalls ist  $w_k > 0$  für mindestens ein  $k \in \mathcal{B}$ . Um in diesem Fall  $x^+ \geq 0$  zu gewährleisten, müssen wir  $\gamma$  so wählen, dass

$$\begin{aligned} \gamma \geq 0 \quad \text{und} \quad x_i^+ = x_i - \gamma w_i &\geq 0 \quad \text{für alle } i \in \mathcal{B} \text{ mit } w_i > 0 \\ \Leftrightarrow 0 \leq \gamma \leq \frac{x_i}{w_i} &\quad \text{für alle } i \in \mathcal{B} \text{ mit } w_i > 0. \end{aligned} \quad (17.4)$$

zu (3):

Um überhaupt eine Veränderung zu erzielen müssen wir  $\gamma > 0$  wählen. Dann aber verlässt der Index  $j$  wegen  $x_j^+ = \gamma > 0$  die Nichtbasis und wird zu einem neuen Basisindex. Da eine Basis nur maximal  $m$  Elemente enthalten darf, muss ein Index aus  $\mathcal{B}$  die Basis verlassen. Wir müssen  $\gamma > 0$  daher so einstellen, dass  $x_i^+ = 0$  für ein  $i \in \mathcal{B}$ . Der zugehörige Index  $i$  verlässt dann die Basis und wird neuer Nichtbasisindex. In Anbetracht von (17.4) wählen wir  $\gamma$  daher als

$$\gamma := \min_{k \in \mathcal{B}, w_k > 0} \frac{x_k}{w_k}. \quad (17.5)$$

Dann ist  $x^+$  wegen (17.4) zulässig, und für einen Index  $\ell \in \mathcal{B}$ , in dem das Minimum in (17.5) angenommen wird, gilt  $x_\ell^+ = x_\ell - \gamma w_\ell = 0$ , so dass er in die Nichtbasis wechseln kann. Die neue Basis ist dann

$$\mathcal{B}^+ = (\mathcal{B} \cup \{j\}) \setminus \{\ell\}. \quad (17.6)$$

Damit es sich wirklich um eine neue Basis handelt, müssen wir noch zeigen, dass  $A_{[\mathcal{B}^+]}$  regulär ist.

**Lemma 17.7**

Sei  $s_j < 0$  für ein  $j \in \mathcal{N}$  und  $w_{\mathcal{B}} \not\leq 0$ . Das Minimum in (17.5) werde in  $\ell \in \mathcal{B}$  angenommen. Dann ist  $\mathcal{B}^+ = (\mathcal{B} \cup \{j\}) \setminus \{\ell\}$  eine zulässige Basis mit Basisvektor

$$x_i^+ = \begin{cases} x_i - \gamma w_i, & i \in \mathcal{B} \\ \gamma, & i = j \\ 0, & i \in \mathcal{N} \setminus \{j\}, \end{cases} \quad (17.7)$$

wobei  $\gamma$  wie in (17.5) gewählt wird. Daraus folgt insbesondere, dass  $x_\ell^+ = 0$ .

*Beweis:* Seien  $(\alpha_i)_{i \in \mathcal{B}^+}$  so gewählt, dass

$$\begin{aligned} 0 &= \sum_{i \in \mathcal{B}^+} \alpha_i A_{[i]} = \sum_{i \in \mathcal{B}, i \neq \ell} \alpha_i A_{[i]} + \alpha_j A_{[j]} = \sum_{i \in \mathcal{B}, i \neq \ell} \alpha_i A_{[i]} + \alpha_j A_{[\mathcal{B}]} A_{[\mathcal{B}]}^{-1} A_{[j]} \\ &= \sum_{i \in \mathcal{B}, i \neq \ell} \alpha_i A_{[i]} + \alpha_j A_{[\mathcal{B}]} w_{\mathcal{B}} \\ &= \sum_{i \in \mathcal{B}, i \neq \ell} \alpha_i A_{[i]} + \alpha_j \sum_{i \in \mathcal{B}} A_{[i]} w_i \\ &= \sum_{i \in \mathcal{B}, i \neq \ell} (\alpha_i + \alpha_j w_i) A_{[i]} + \alpha_j w_\ell A_{[\ell]}. \end{aligned}$$

Da  $\mathcal{B}$  eine Basis ist, sind die  $A_{[i]}$ ,  $i \in \mathcal{B}$ , linear unabhängig, so dass  $\alpha_i + \alpha_j w_i = 0$  für alle  $i \in \mathcal{B} \setminus \{\ell\}$  und  $\alpha_j w_\ell = 0$  gelten muss. Aus  $w_\ell > 0$  folgt aber  $\alpha_j = 0$  und damit auch  $\alpha_i = 0$  für alle  $i \in \mathcal{B} \setminus \{\ell\}$ . Also ist  $\alpha_i = 0$  für alle  $i \in \mathcal{B}^+$ , so dass  $A_{[\mathcal{B}^+]} \in \mathbb{R}^{m \times m}$  regulär und  $\mathcal{B}^+$  somit Basis ist. Aus (17.3) folgt sofort, dass

$$A_{[\mathcal{B}^+]} x_{\mathcal{B}^+}^+ = A_{[\mathcal{B}]} x_{\mathcal{B}}^+ + A_{[j]} \underbrace{x_j^+}_{=\gamma} - A_{[\ell]} \underbrace{x_\ell^+}_{=0} = b.$$

Wegen  $x_{\mathcal{N}^+}^+ = 0$  mit  $\mathcal{N}^+ := \{1, \dots, n\} \setminus \mathcal{B}^+$  ist  $x^+$  der zu  $\mathcal{B}^+$  gehörige Basisvektor. Wie in (17.3) und (17.4) gesehen, ist dieser zulässig, was schließlich die Zulässigkeit von  $\mathcal{B}^+$  liefert.  $\square$

Zusammenfassend haben wir mit dieser Konstruktion Folgendes erreicht:

- Sind die reduzierten Kosten  $s_{\mathcal{N}} \geq 0$ , dann ist der aktuelle Basisvektor eine Lösung des LPs.
- Sind  $s_{\mathcal{N}} \not\geq 0$  und  $w_{\mathcal{B}} \leq 0$ , dann besitzt das LP keine Lösung.
- Sind  $s_{\mathcal{N}} \not\geq 0$  und  $w_{\mathcal{B}} \not\leq 0$ , dann erhält man mit den in (17.5)–(17.6) beschriebenen Setzungen eine neue Basis mit zugehörigem Basisvektor (also Ecke), der einen kleineren Zielfunktionswert liefert.

Dies führt auf folgenden Algorithmus:

**Algorithmus 17.8** (Primale Version des Simplex-Verfahrens)

- 1: Gegeben sei eine zulässige Basis  $\mathcal{B}$  mit zugehörigem Basisvektor  $x$  und Nichtbasis  $\mathcal{N} = \{1, \dots, n\} \setminus \mathcal{B}$ .
- 2: Löse das lineare Gleichungssystem

$$A_{(\mathcal{B})}^{\top} \nu = c_{\mathcal{B}}. \quad (17.8)$$

- 3: Berechne die reduzierten Kosten

$$s_{\mathcal{N}} = c_{\mathcal{N}} - A_{(\mathcal{N})}^{\top} \nu. \quad (17.9)$$

- 4: **if**  $s_{\mathcal{N}} \geq 0$  **then**
- 5:     **stop**:  $\mathcal{B}$  ist optimal, zugehöriger Basisvektor  $x$  ist Lösung des LPs.
- 6: **else**
- 7:     Wähle ein  $j \in \mathcal{N}$  mit  $s_j < 0$ .
- 8:     Löse das lineare Gleichungssystem

$$A_{[\mathcal{B}]} w_{\mathcal{B}} = A_{[j]}. \quad (17.10)$$

- 9:     **if**  $w_{\mathcal{B}} \leq 0$  **then**
- 10:         **stop**: LP ist unbeschränkt, also nicht lösbar.
- 11:     **else**
- 12:         Bestimme  $\gamma \in \mathbb{R}$  und  $\ell \in \mathcal{B}$ , so dass

$$\gamma = \frac{x_{\ell}}{w_{\ell}} = \min_{k \in \mathcal{B}, w_k > 0} \frac{x_k}{w_k}.$$

- 13:         **Update**: Setze

$$\begin{aligned} x_{\mathcal{B}} &= x_{\mathcal{B}} - \gamma w_{\mathcal{B}} \\ x_j &= \gamma \\ \mathcal{B} &= (\mathcal{B} \cup \{j\}) \setminus \{\ell\} \\ \mathcal{N} &= \{1, \dots, n\} \setminus \mathcal{B} \end{aligned}$$

- 14:     **end if**
- 15: **end if**
- 16: **goto** 2

**Bemerkung 17.9**

Schritt 2 wird oft auch als *Backward Transformation* bezeichnet. Die Schritte 3–7 sind als *Pricing* bekannt. Schritt 8 heißt auch *Forward Transformation*, und Schritte 9–12 nennt man *Ratio-Test*.  $\diamond$

**Satz 17.10**

Terminiert Algorithmus 17.8 in endlich vielen Schritten, dann liefert er das korrekte Ergebnis.

*Beweis:* Eigentlich haben wir dieses Resultat in den vorangegangenen Lemmas schon bewiesen. Während Lemma 17.6 die Abbruchbedingung des Ratio-Tests in Schritt 9 bzw. 10 liefert, folgt die Abbruchbedingung des Princings in den Schritten 4 bzw. 5 direkt aus Lemma 17.5.  $\square$

**Bemerkung 17.11**

Alternativ kann man die Abbruchbedingung  $s_{\mathcal{N}} \geq 0$  auch mit Hilfe von Dualitätsargumenten begründen:

Für  $\nu \in \mathbb{R}^m$  aus Schritt 2 gilt  $A_{(\mathcal{B})}^{\top} \nu = c_{\mathcal{B}}$  und  $A_{(\mathcal{N})}^{\top} \nu = c_{\mathcal{N}} - s_{\mathcal{N}} \leq c_{\mathcal{N}}$ , falls  $s_{\mathcal{N}} \geq 0$ , d.h.  $\nu$  ist zulässig für das duale Problem (DLP). Zudem gilt

$$(c - A^{\top} \nu)^{\top} x = \underbrace{(c_{\mathcal{B}} - A_{(\mathcal{B})}^{\top} \nu)^{\top}}_{=0} x_{\mathcal{B}} + (c_{\mathcal{N}} - A_{(\mathcal{N})}^{\top} \nu)^{\top} \underbrace{x_{\mathcal{N}}}_{=0} = 0,$$

so dass insgesamt das KKT-System (16.7) erfüllt ist und  $x$  nach Satz 16.7 damit eine Lösung des LPs ist. Dieser Beweis zeigt auch, dass ebenfalls eine Lösung des dualen Problems zu (LP) gefunden ist, wenn der Algorithmus in Schritt 4 bzw. 5 terminiert. Dabei handelt es sich um das in Schritt 2 berechnete  $\nu$ . Man kann zeigen, dass  $\nu$  (zusammen mit den Schlupfvariablen  $s = c - A^{\top} \nu$ ) sogar eine optimale *Basislösung* für das (in Standardform gebrachte) Problem (DLP) ist.  $\diamond$

**Satz 17.12**

Alle zulässigen Basen seien nicht degeneriert. Dann bricht Algorithmus 17.8 nach endlich vielen Iterationen ab.

*Beweis:* Sei  $\mathcal{B}$  die aktuelle Basis mit zugehörigem Basisvektor  $x$ . Terminiert der Algorithmus nicht mit dieser Basis, dann folgt für  $\gamma$  aus Schritt 12 wegen Nicht-Degeneriertheit, d.h.  $x_{\mathcal{B}} > 0$ , dass

$$\gamma = \min_{k \in \mathcal{B}, w_k > 0} \frac{x_k}{w_k} > 0.$$

Für die nächste Iterierte  $x^+$  ergibt sich nach (17.2) damit, dass

$$c^{\top} x^+ = c^{\top} x + \gamma s_j < c^{\top} x \tag{17.11}$$

Nehmen wir an, der Algorithmus terminiere nicht endlich und erzeuge so eine Folge  $\{\mathcal{B}_k\}_{k \in \mathbb{N}}$  an Basen. Wegen (17.11) ist die zugehörige Folge der Zielfunktionswerte streng monoton fallend. Deshalb kann keine Basis in der Folge zweimal auftreten. (Beachte hierbei, dass der zu einer Basis gehörige Basisvektor eindeutig definiert ist, s. Bemerkung 17.2.) Da es nach Definition einer Basis nur endlich viele Basen gibt, ist dies ein Widerspruch.  $\square$

Leider trifft dieses Resultat auf den degenerierten Fall nicht mehr zu. In diesem Fall gibt es Beispiele, bei denen bestimmte Basen unendlich oft auftreten. Man spricht dann vom *Kreiseln des Simplex-Verfahrens*. Ein solches Beispiel wird in der Übung vorgestellt. Die Grundversion des Simplex-Verfahrens in Algorithmus 17.8 lässt jedoch bei der Wahl der Indizes  $j$  und  $\ell$  in Schritt 7 bzw. 12 einige Freiheit. Wählt man diese geschickt, so kann das Kreiseln vermieden werden:

**Proposition 17.13**

Die folgenden Auswahlregeln für die Schritte 7 und 12 garantieren die Endlichkeit des Simplex-Verfahrens:

- *Regel von Bland*

Wähle in den Schritten 7 und 12 jeweils die kleinsten in Frage kommenden Indizes  $j \in \mathcal{N}$  und  $\ell \in \mathcal{B}$ .

- *Lexikographische Regel*

Schritt 7 bleibt unverändert, d.h. wir wählen einen beliebigen Index  $j \in \mathcal{N}$  mit  $s_j < 0$ . Der Ratio-Test in Schritt 12 wird dagegen wie folgt modifiziert. Zunächst bezeichnen wir die Menge der Indizes  $i \in \mathcal{B}$  mit  $x_i/w_i = \min_{k \in \mathcal{B}, w_k > 0} (x_k/w_k)$  mit  $\mathcal{M}$ . Dann wählt man den Index  $\ell \in \mathcal{M}$  so, dass  $(A_{[\mathcal{B}]^{-1}A})_{\ell}/w_{\ell}$  lexikographisch minimal wird. Um diesen Begriff zu erklären, kürzen wir wie folgt ab  $v^i = (A_{[\mathcal{B}]^{-1}A})_i/w_i$ ,  $i \in \mathcal{M}$ . Dann ist  $v^{\ell}$  „lexikographisch minimal“, falls für jedes  $i \in \mathcal{M}$  mit  $v^i \neq v^{\ell}$  gilt:

$$v_{\iota}^{\ell} < v_{\iota}^i \quad \text{mit} \quad \iota := \min\{m \in \mathcal{M} : v_m^{\ell} \neq v_m^i\}. \quad \diamond$$

Beweise dieser Aussagen sind in [?, Abschnitt 3.4] und [?, Abschnitt 3.4] zu finden. Neben den hier erwähnten gibt zahlreiche weitere Auswahlregeln, s. z.B. ?.

Aus praktischer Sicht stellt sich natürlich nicht nur die Frage der Endlichkeit, sondern man benötigt auch Aussagen über das Worst-Case Verhalten des Simplex-Algorithmus. Eine negative Antwort hierfür liefert das bekannte Beispiel von Klee und Minty:

**Beispiel 17.14 (Klee-Minty-Würfel)**

Sei  $\varepsilon \in (0, 1/2)$  gegeben. Wir betrachte folgendes LP:

$$\left. \begin{array}{ll} \text{minimiere} & -x_n \\ \text{s.t.} & \varepsilon \leq x_1 \leq 1 \\ & \varepsilon x_{j-1} \leq x_j \leq 1 - \varepsilon x_{j-1}, \quad j = 2, 3, \dots, n. \end{array} \right\} \quad (17.12)$$

Die zulässige Menge entspricht einem (für kleines  $\varepsilon$ ) leicht verzerrten Hyperwürfel. Man kann zeigen, dass der Simplex-Algorithmus (mit der Regel von Bland) zur Lösung dieser Aufgabe *alle* Ecken dieses verzerrten Würfels abläuft. Er benötigt also  $2^n$  Iterationen, das heißt der Simplex-Algorithmus hat in diesem Beispiel keine polynomiale sondern exponentielle Laufzeit.  $\diamond$

In der Praxis verhält sich der Simplex-Algorithmus allerdings häufig sehr viel besser als im Beispiel von Klee und Minty. Bei vielen Anwendungen zeigt er ein polynomiales Laufzeitverhalten.

**Bemerkung 17.15**

Man kann auch das zu (LP) duale Problem (DLP) mit dem Simplex-Algorithmus lösen, nachdem man es in die Standardform gebracht hat.

Es gibt zahlreiche Beispiele, in denen das duale Simplex-Verfahren sehr viel schneller konvergiert als das primale und umgekehrt. Beispielsweise durchläuft das duale Simplex-Verfahren im Fall des Klee-Minty-Würfels aus Beispiel 17.14 nicht alle Ecken und konvergiert daher sehr viel schneller als die primale Version. Leider kann man einem LP a priori aber kaum ansehen, welche Version des Simplex-Verfahrens sich besser zu seiner Lösung eignet.

Die Verwendung des dualen Simplex-Algorithmus ist außerdem vorzuziehen, wenn nach dem Hinzufügen einer neuen Nebenbedingung (im primalen LP) reoptimiert werden soll. Die dual optimale Basis bleibt dann nämlich dual zulässig und kann als Startbasis für weitere Iterationen verwendet werden. Umgekehrt sollte bei einer Veränderung der Zielfunktion primal reoptimiert werden. Siehe dazu § 19.

## § 18 Praktische Umsetzung des Simplex-Algorithmus

Der Simplex-Algorithmus lässt noch einige wesentliche Fragen offen, insbesondere:

- Wie erhält man eine zulässige Startbasis?
- Wie löst man möglichst effizient die linearen Gleichungssysteme in den Schritten 2 und 8?

Insbesondere eine effiziente Lösung der Gleichungssysteme ist sehr wichtig, da sie den aufwendigsten Teil einer Iteration im Simplex-Algorithmus darstellt. Wir wenden uns aber zunächst der Frage einer zulässigen Startbasis zu. Dazu betrachten wir das folgende Hilfsproblem:

$$\left. \begin{array}{l} \text{minimiere } \mathbf{1}^\top y \\ \text{s.t. } Ax + y = b, \quad x, y \geq 0, \end{array} \right\} \quad (18.1)$$

wobei  $A \in \mathbb{R}^{m \times n}$  und  $b \in \mathbb{R}^m$  die Daten von (LP) und  $\mathbf{1} = (1, 1, \dots, 1) \in \mathbb{R}^m$ .

### Satz 18.1

Sei  $b \geq 0$ . Dann gilt:

- (1)  $\mathcal{B} = \{n+1, n+2, \dots, n+m\}$  ist eine zulässige Basis für (18.1) mit Basisvektor  $(x, y)^\top = (0, b)^\top$ .
- (2) Das Hilfs-LP (18.1) ist lösbar.
- (3) Sei  $(x^*, y^*)^\top$  ein optimaler Basisvektor von (18.1). Ist  $y^* \neq 0$ , dann ist die zulässige Menge des Ausgangsproblems (LP) leer. Ist hingegen  $y^* = 0$  und  $\text{rang}(A) = m$ , dann ist  $x^*$  eine zulässige Basisvektor für (LP).

### Bemerkung 18.2

Da das Hilfs-LP (18.1) lösbar ist und man zudem eine zulässige Basis für (18.1) kennt nach Satz 18.1(1) bzw. (2), kann man das Hilfs-LP mittels des Simplex-Verfahrens lösen. Als Lösung erhält man die gewünschte zulässige Basis für (LP), mit der man den „eigentlichen“ Simplex-Algorithmus zur Lösung des Ausgangsproblems (LP) starten kann (falls die zulässige Menge von (LP) nichtleer ist). Man beachte, dass die Bedingung  $b \geq 0$  keine wirkliche Einschränkung darstellt, denn ist  $b_i < 0$  für ein  $i \in \{1, \dots, m\}$ , so kann man die entsprechende Gleichung einfach mit  $-1$  multiplizieren.

Die Anwendung des Simplex-Verfahrens auf das Hilfsproblem (18.1) wird oft als *Phase I des Simplex-Algorithmus* bezeichnet. Die Lösung des eigentlichen Problems (LP) mit Hilfe des Simplex-Verfahrens heißt dann *Phase II*.  $\diamond$

*Beweis von Satz 18.1:* zu (1):

Der Vektor  $(x, y) = (0, b)$  erfüllt trivialerweise die Gleichungsnebenbedingung von

(18.1). Wegen  $b \geq 0$  sind zudem die Ungleichungsnebenbedingungen erfüllt, so dass  $(0, b)$  zulässig ist. Die Gleichungsnebenbedingung in (18.1) lässt sich auch als

$$(A, I) \begin{pmatrix} x \\ y \end{pmatrix} = b$$

schreiben, wobei  $I \in \mathbb{R}^{m \times m}$  die Einheitsmatrix bezeichnet. Mit  $\mathcal{B} = (n+1, n+2, \dots, n+m)$  erhalten wir  $(A, I)_{[\mathcal{B}]} = I$ , so dass es sich bei  $\mathcal{B}$  tatsächlich eine Basis ist. Zudem ist  $(0, b)$  trivialerweise der passende Basisvektor. Da er wie bereits gesehen zulässig ist, ist  $\mathcal{B}$  ebenfalls zulässig.

zu (2):

Wegen der Nebenbedingung  $y \geq 0$  gilt für die Zielfunktion von (18.1), dass  $\mathbf{1}^\top y \geq 0$  für alle zulässigen  $y$ . Damit ist das Zielfunktional auf der zulässigen Menge nach unten beschränkt, was nach Satz 15.1 wegen der Zulässigkeit die Existenz einer Lösung impliziert.

zu (3):

Sei  $(x^*, y^*)$  ein optimaler Basisvektor mit  $y^* \neq 0$ . Der Optimalwert von (18.1) ist dann  $f^* = \mathbf{1}^\top y^* > 0$ . Gäbe es einen zulässigen Punkt  $\hat{x} \in P$  des Ausgangsproblems (LP), dann wäre  $(\hat{x}, 0)$  zulässig für (18.1), würde aber den Zielfunktionswert  $\mathbf{1}^\top 0 = 0 < f^*$  liefern, im Widerspruch zur Optimalität von  $(x^*, y^*)$ .

Sei nun  $y^* = 0$  und  $\text{rang}(A) = m$ . Da  $(x^*, y^*)$  ein Basisvektor ist, gibt es eine zugehörige Basis  $\mathcal{B}^* \subset \{1, \dots, n+m\}$  mit  $|\mathcal{B}^*| = m$  und  $(x^*, y^*)_{\mathcal{B}^*}^\top = (A, I)_{[\mathcal{B}^*]}^{-1} b$ . Da  $(A, I)_{[\mathcal{B}^*]}$  regulär ist, sind die Spalten

$$A_{[i]}, \quad i \in \mathcal{M} := \{k \in \mathcal{B}^* \cap \{1, \dots, n\} : x_k^* > 0\} = \mathcal{B}^* \cap \text{supp}(x^*)$$

linear unabhängig. Ist  $|\mathcal{M}| = m$ , so setzen wir  $\hat{\mathcal{B}} = \mathcal{M}$ . Andernfalls kann  $\mathcal{M}$  wegen  $\text{rang}(A) = m$  zu einer Indexmenge  $\hat{\mathcal{B}}$  ergänzt werden, so dass  $|\hat{\mathcal{B}}| = m$  und  $A_{[\hat{\mathcal{B}}]}$  regulär ist. Damit ist  $\hat{\mathcal{B}}$  eine Basis für (LP). Für  $\hat{\mathcal{B}}$  gilt dann nach Konstruktion  $x_{\hat{\mathcal{N}}}^* = 0$  mit  $\hat{\mathcal{N}} = \{1, \dots, n\} \setminus \hat{\mathcal{B}}$  und wegen  $y^* = 0$ , dass

$$b = (A, I) \begin{pmatrix} x^* \\ y^* \end{pmatrix} = A_{[\mathcal{M}]} x_{\mathcal{M}}^* = A_{[\hat{\mathcal{B}}]} x_{\hat{\mathcal{B}}}^* \quad \Leftrightarrow \quad x_{\hat{\mathcal{B}}}^* = A_{[\hat{\mathcal{B}}]}^{-1} b.$$

Wegen  $Ax^* = b$  und  $x^* \geq 0$  ist die Basis zulässig. □

### Bemerkung 18.3

Im Fall  $\mathcal{M} \neq \hat{\mathcal{B}}$  gibt es Indizes  $i \in \hat{\mathcal{B}}$  mit  $x_i^* = 0$ , so dass  $\hat{\mathcal{B}}$  degeneriert ist, was zu den oben erwähnten Problemen führen kann (Kreiseln).

Wir kommen nun zu einer effizienten Methode, die linearen Gleichungssysteme in den Schritten 2 und 8 zu lösen. Dazu nehmen wir an, es sei gerade eine Iteration im Simplex-Verfahren mit Basis  $\mathcal{B}$  durchgeführt worden, und der Algorithmus sei in dieser Iteration nicht terminiert. Die zugehörige Basismatrix ist

$$A_{[\mathcal{B}]} = (A_{[\mathcal{B}_1]}, \dots, A_{[\mathcal{B}_m]}),$$

wobei  $\mathcal{B}_1, \dots, \mathcal{B}_m \in \{1, \dots, n\}$  die Indizes in der Basis seien. Der Index  $\ell$  aus Schritt 12, der nach dem Update in Schritt 13 die Basis verlässt, sei das  $i$ -te Element in  $\mathcal{B}$ , d.h.  $\ell = \mathcal{B}_i$ ,  $1 \leq i \leq m$ . An seine Stelle tritt nun der Index  $j$  aus 7, d.h. die neue Basis  $\mathcal{B}^+$  für die nächste Iteration ist gegeben durch

$$\mathcal{B}^+ = \{\mathcal{B}_1, \dots, \mathcal{B}_{i-1}, j, \mathcal{B}_{i+1}, \dots, \mathcal{B}_m\}.$$

Unter Beachtung der Forward-Transformation in Schritt 8 ergibt sich dann für die neue Basismatrix:

$$\begin{aligned} A_{[\mathcal{B}^+]} &= (A_{[\mathcal{B}_1]}, \dots, A_{[\mathcal{B}_{i-1}]}, A_{[j]}, A_{[\mathcal{B}_{i+1}]}, A_{[\mathcal{B}_m]}) \\ &= (A_{[\mathcal{B}_1]}, \dots, A_{[\mathcal{B}_{i-1}]}, A_{[\mathcal{B}]}w_{\mathcal{B}}, A_{[\mathcal{B}_{i+1}]}, A_{[\mathcal{B}_m]}) \\ &= (A_{[\mathcal{B}]}e_1, \dots, A_{[\mathcal{B}]}e_{i-1}, A_{[\mathcal{B}]}w_{\mathcal{B}}, A_{[\mathcal{B}]}e_{i+1}, \dots, A_{[\mathcal{B}]}e_m) \\ &= A_{[\mathcal{B}]} \underbrace{(e_1, \dots, e_{i-1}, w_{\mathcal{B}}, e_{i+1}, \dots, e_m)}_{=: F}, \end{aligned}$$

wobei  $e_k$  der  $k$ -te Einheitsvektor ist. Sind  $w_{\mathcal{B}_1}, \dots, w_{\mathcal{B}_m}$  die Komponenten von  $w_{\mathcal{B}}$ , dann definieren wir den Vektor  $v \in \mathbb{R}^m$  durch

$$v_k = \begin{cases} \frac{1}{w_{\mathcal{B}_i}}, & k = i \\ -\frac{w_{\mathcal{B}_k}}{w_{\mathcal{B}_i}}, & k \neq j, \end{cases} \quad k = 1, \dots, m,$$

wobei  $i$  wie oben durch  $\ell = \mathcal{B}_i$  festgelegt ist. Man beachte hierbei, dass  $v$  wohldefiniert ist, da  $w_{\mathcal{B}_i} = w_{\ell} > 0$  nach Schritt 12 des Algorithmus gilt. Die Inverse von  $F \in \mathbb{R}^{m \times m}$  ist dann durch

$$F^{-1} = (e_1, \dots, e_{i-1}, v, e_{i+1}, \dots, e_m)$$

gegeben, wie man durch einfaches Nachrechnen leicht bestätigt. Damit folgt für die Inverse von  $A_{[\mathcal{B}^+]}$ , dass

$$A_{[\mathcal{B}^+]}^{-1} = F^{-1}A_{[\mathcal{B}]}^{-1}.$$

Ist also die Inverse von  $A_{[\mathcal{B}]}$  bekannt, so kann die Inverse der folgenden Basismatrix durch eine Matrix-Multiplikation mit der dünn besetzten Matrix  $F^{-1}$  berechnet werden. In der Praxis berechnet man häufig eine LU-Zerlegung der Startbasismatrix. Die Inversen der folgenden Basismatrizen ergeben sich dann durch sukzessive Matrixmultiplikation mit dem jeweiligen  $F^{-1}$ . (Man beachte, dass sich  $w$  und damit auch  $F$  in jeder Iteration des Algorithmus ändert.) Um numerische Instabilitäten zu vermeiden, sollte allerdings ab und an eine neue LU-Faktorisierung vorgenommen werden, z.B. in jeder 100sten Iteration.

Es gibt einige weitere Methoden, die Matrizen im Simplex-Algorithmus effizient zu invertieren. Für Probleme mit kleiner Dimension eignet sich das so genannte *Simplex-Tableau*, welches beispielsweise in [?, Kapitel 3] näher erläutert wird. Da dem Simplex-Tableau das Gaußsche Eliminationsverfahren zugrunde liegt, eignet es sich nicht für große  $n$  und  $m$ . Für Probleme mit großer Dimension kann man alternativ zum oben beschriebenen Verfahren auch die so genannte *Sherman-Morrison-Formel* verwenden, welche in [?, Kapitel 3] näher beschrieben ist.

## § 19 Reoptimierung

In vielen Anwendungen verändert sich die Struktur des LPs, indem beispielsweise eine neue Nebenbedingung hinzukommt. Im Fall des Transportproblems aus Beispiel 2.1 könnte z.B. eine neue Filiale eröffnet werden, die ebenfalls beliefert werden muss. Dann vergrößert sich die Anzahl der Variablen und eine neue Nebenbedingung muss beachtet werden. Hat man bereits das LP ohne Berücksichtigung der neuen Filiale gelöst, stellt sich die Frage, inwieweit die alte Lösung für das neue Problem verwendet werden kann. Wir untersuchen im Folgenden die Fälle:

- (1) Änderung des Vektors  $c$  in der Zielfunktion,



- (2) Änderung der rechten Seite  $b$ ,
- (3) Hinzufügen einer neuen Nebenbedingung.

Dabei gehen wir davon aus, wir hätten bereits eine Lösung von (LP), d.h.

$$\left. \begin{array}{l} \text{minimiere } c^\top x \\ \text{s.t. } Ax = b, \quad x \geq 0 \end{array} \right\} \quad (\text{LP})$$

mit Hilfe des Simplex-Verfahrens bestimmt. Die zugehörige optimale Basis sei  $\mathcal{B}^*$ , den Basisvektor bezeichnen wir mit  $x^*$ . Die entsprechende Nichtbasis bezeichnen wir mit  $\mathcal{N}^* = \{1, \dots, n\} \setminus \mathcal{B}^*$ . Die zu  $\mathcal{B}^*$  gehörige duale Optimallösung sei  $(\nu^*, s^*)$ , diese wird nach Bemerkung 17.11 ebenfalls durch den Simplex-Algorithmus berechnet.

zu (1):

Angenommen, der Vektor  $c$  ändert sich zu  $\tilde{c} \in \mathbb{R}^n$ . Da sich die Nebenbedingungen nicht ändern, bleibt  $x^*$  auch für das modifizierte Problem zulässig. Ob es auch optimal bleibt, hängt von den reduzierten Kosten ab:

$$\tilde{s}_{\mathcal{N}^*} = \tilde{c}_{\mathcal{N}^*} - A_{(\mathcal{N}^*)}^\top A_{(\mathcal{B}^*)}^{-\top} \tilde{c}_{\mathcal{B}^*}.$$

$\tilde{s}_{\mathcal{N}^*} \geq 0$ : Dann ist  $x^*$  nach Lemma 17.5 auch eine Lösung des modifizierten Problems.

$\tilde{s}_{\mathcal{N}^*} \not\geq 0$ : Jetzt ist  $x^*$  nicht mehr optimal. Da es aber zulässig ist, kann der primale Simplex-Algorithmus mit  $\mathcal{B}^*$  als zulässiger Basis zur Lösung des modifizierten Problems gestartet werden.

zu (2):

Wird  $b$  in  $\tilde{b}$  verändert, dann ist  $x^*$  i.A. nicht mehr zulässig für das modifizierte primale Problem. Das entsprechende duale Problem lautet

$$\left. \begin{array}{l} \text{maximiere } \tilde{b}^\top \nu \\ \text{s.t. } A^\top \nu + s = c, \quad s \geq 0 \end{array} \right\} \quad (19.1)$$

und besitzt damit dieselbe zulässige Menge wie das duale Ausgangsproblem. Insbesondere ist die duale Lösung  $(\nu^*, s^*)$  weiterhin eine zulässige Basislösung für (19.1). Durch Anwendung des Simplex-Algorithmus auf das duale Problem (vgl. Bemerkung 17.15) kann also analog zu (1) reoptimiert werden.

zu (3):

Nun wird dem primalen Problem eine zusätzliche Unleichung hinzugefügt, z.B.  $a^\top x \leq \delta$  mit  $a \in \mathbb{R}^n$  und  $\delta \in \mathbb{R}$ . Durch Einführung einer Schlupfvariablen  $x_{n+1}$  können wir dies äquivalent schreiben als  $a^\top x + x_{n+1} = \delta$  und  $x_{n+1} \geq 0$ . Das modifizierte LP ist dann gegeben durch:

$$\left. \begin{array}{l} \text{minimiere } \tilde{c}^\top \tilde{x} \\ \text{s.t. } \tilde{x} \in \mathbb{R}^{n+1}, \quad \tilde{A} \tilde{x} = \tilde{b}, \quad \tilde{x} \geq 0 \end{array} \right\} \quad (19.2)$$

mit  $\tilde{c} = (c, 0) \in \mathbb{R}^{n+1}$ ,  $\tilde{b} = (b, \delta) \in \mathbb{R}^{m+1}$  und

$$\tilde{A} = \begin{pmatrix} A & 0 \\ a^\top & 1 \end{pmatrix} \in \mathbb{R}^{(m+1) \times (n+1)}.$$

Wir können nun von der neuen Basis  $\mathcal{B}^* \cup \{n+1\}$  ausgehen, die dual zulässig bleibt, und wie in (2) den Simplex-Algorithmus auf das duale Problem anwenden.



Teil 3

## Diskrete Optimierung



## Ganzzahlige Programmierung

### Inhalt

20	<b>Ganzzahlige Polyeder</b>	85
21	<b>Branch-and-Bound</b>	88

### § 20 Ganzzahlige Polyeder

Wir interessieren uns im Folgenden für lineare Programme, bei denen zusätzlich verlangt wird, dass alle Variablen ganzzahlige Werte annehmen. Dabei gehen wir von der natürlichen Form eines linearen Programms aus, also

$$\left. \begin{array}{l} \text{minimiere } c^\top x \\ \text{s.t. } Ax \leq b \\ x \in \mathbb{R}^n \end{array} \right\} \quad (\text{LP}')$$

und betrachten jetzt das *ganzzahlige lineare Programm*

$$\left. \begin{array}{l} \text{minimiere } c^\top x \\ \text{s.t. } Ax \leq b \\ x \in \mathbb{Z}^n \end{array} \right\} \quad (\text{ILP})$$

Im Folgenden nehmen wir oft an, dass alle Eingabedaten ganzzahlig sind, dass also  $A \in \mathbb{Z}^{m \times n}$ ,  $b \in \mathbb{Z}^m$  und  $c \in \mathbb{Z}^n$  gilt. Durch Skalierung kann auch der Fall  $A \in \mathbb{Q}^{m \times n}$ ,  $b \in \mathbb{Q}^m$  und  $c \in \mathbb{Q}^n$  hierauf zurückgeführt werden, sodass die Ganzzahligkeit der Eingabedaten aus praktischer Sicht keine wesentliche Einschränkung darstellt.

Dennoch ist die Lösung von (ILP) im Allgemeinen sehr viel schwieriger als die von (LP'), was sich sogar theoretisch beweisen lässt (das ist Inhalt der Vorlesung "Diskrete Optimierung").

In diesem Kapitel interessiert uns jedoch zunächst die Frage, unter welchen Voraussetzungen sich das Problem (ILP) direkt auf (LP') zurückführen lässt, sodass die Algorithmen aus Kapitel 6 auch zur Lösung von (ILP) eingesetzt werden können.

Nach Satz 15.4 wird das Optimum von (LP'), falls es existiert, auch immer in einer Ecke des Polyeders

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

angenommen. Diese Aussage lässt sich wie folgt verallgemeinern:

**Satz 20.1**

Sei  $X \subseteq \mathbb{R}^n$  und  $c \in \mathbb{R}^n$ . Dann gilt  $\min_{x \in X} c^\top x = \min_{x \in \text{conv}(X)} c^\top x$ .

*Beweis:* Die Richtung “ $\geq$ ” ist klar, da  $X \subseteq \text{conv}(X)$  gilt. Wir zeigen also “ $\leq$ ”. Sei dazu  $x_0 \in \text{conv}(X)$  beliebig gewählt. Nach Lemma 4.11 gilt dann

$$x_0 = \sum_{i=1}^k \lambda_i x_i$$

für geeignete  $x_1, \dots, x_k \in X$  und  $\lambda_1, \dots, \lambda_k \geq 0$  mit  $\sum_{i=1}^k \lambda_i = 1$ . Damit folgt

$$c^\top x_0 = \sum_{i=1}^k \lambda_i c^\top x_i.$$

Wegen  $\sum_{i=1}^k \lambda_i = 1$  folgt daraus, dass ein  $i \in \{1, \dots, k\}$  existiert mit  $c^\top x_i \leq c^\top x_0$ . Mit  $x_i \in X$  ergibt sich

$$\min_{x \in X} c^\top x \leq c^\top x_i \leq c^\top x_0.$$

Daraus folgt die Behauptung.  $\square$

### Korollar 20.2

Sei  $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ . Dann stimmt der Optimalwert von Problem (ILP) mit dem Optimalwert von

$$\left. \begin{array}{ll} \min & c^\top x \\ \text{s.t.} & x \in \text{conv}(P \cap \mathbb{Z}^n) \end{array} \right\} \quad (\text{LPC})$$

überein.

*Beweis:* Die Aussage folgt sofort aus Satz 20.1 mit  $X := P \cap \mathbb{Z}^n$ .  $\square$

### Satz 20.3

Sei  $P$  ein Polytop. Dann ist auch  $P' := \text{conv}(P \cap \mathbb{Z}^n)$  ein Polytop. Insbesondere gilt  $P' = \{x \in \mathbb{R}^n \mid A'x \leq b'\}$  für geeignete  $A' \in \mathbb{R}^{m' \times n}$  und  $b' \in \mathbb{R}^{m'}$ ,  $m' \in \mathbb{N}$ .

*Beweis:* Als Polytop ist  $P$  nach Definition beschränkt. Damit ist die Menge  $P \cap \mathbb{Z}^n$  endlich, nach Korollar 14.11 ist  $\text{conv}(P \cap \mathbb{Z}^n)$  also ein Polytop.  $\square$

Mit Korollar 20.2 und Satz 20.3 reduziert sich das Problem (ILP) – zumindest im beschränkten Fall – *im Prinzip* auf (LPC), also ein lineares Optimierungsproblem über  $P'$ . Dabei stellen sich aber die folgenden Fragen:

- (1) Wie können  $A'$  und  $b'$  konkret *bestimmt* werden?
- (2) Wie *groß* kann (muss)  $m'$  werden?
- (3) Wann gilt  $P' = P$ ?

Die ersten beiden Fragen sind im Allgemeinen schwer zu beantworten und werden in der Vorlesung “Diskrete Optimierung” vertieft. An dieser Stelle widmen wir uns der letzten Frage.

### Definition 20.4

Ein Polyeder  $P$  heißt *ganzzahlig*, falls  $P = \text{conv}(P \cap \mathbb{Z}^n)$  gilt.  $\diamond$

### Lemma 20.5

Ein Polytop ist gleich der konvexen Hülle seiner Ecken.

*Beweis:* Für  $P = \emptyset$  ist die Aussage klar. Sei also  $P$  ein nichtleeres Polytop. Wir nehmen zuerst an, dass  $P$  in Standardform gegeben ist. Nach Lemma 15.3 hat  $P$  endlich viele Ecken  $x_1, \dots, x_r$ . Wegen der Konvexität von  $P$  ist  $\text{conv}\{x_1, \dots, x_r\} \subseteq P$ . Für die umgekehrte Inklusion argumentieren wir per Widerspruch und nehmen an, es gäbe ein  $z \in P$  mit  $z \notin \text{conv}\{x_1, \dots, x_r\}$ . Die Menge  $\text{conv}\{x_1, \dots, x_r\}$  ist nach Korollar 14.11 ein Polytop und somit abgeschlossen und nach Lemma 15.2 nichtleer. Nach Lemma 7.2 lässt sich  $z$  von  $\text{conv}\{x_1, \dots, x_r\}$  strikt trennen, so dass ein  $s \in \mathbb{R}^n$ ,  $s \neq 0$ , existiert mit

$$s^\top z > s^\top x_i \quad \forall i = 1, \dots, r.$$

Wir betrachten das LP

$$\max_{x \in P} s^\top x. \quad (20.1)$$

Da  $P$  nichtleer und beschränkt ist, wird nach Satz 15.4 das Optimum von (20.1) in einer Ecke  $\hat{x}$  von  $P$  angenommen. Somit folgt

$$s^\top \hat{x} \geq s^\top z > s^\top x_i \quad \forall i = 1, \dots, r,$$

und damit der gewünschte Widerspruch.

Für ein beliebiges Polyeder  $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$  können wir wegen der Beschränktheit  $P \subseteq \mathbb{R}_+^n$  annehmen, also ist  $P = \{x \in \mathbb{R}^n \mid Ax \leq b, x \geq 0\}$  isomorph zum Polyeder

$$\{x \in \mathbb{R}^n \mid Ax + z = b, x, z \geq 0\}$$

in Standardform. Da sich die Ecken dabei bijektiv entsprechen, folgt die Aussage.  $\square$

### Satz 20.6

Sei  $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$  ein Polytop. Dann sind die folgenden Aussagen äquivalent:

- (i)  $P$  ist ganzzahlig
- (ii) Jede Ecke von  $P$  ist ganzzahlig.
- (iii) Für alle  $c \in \mathbb{R}^n$  stimmen die Optimalwerte von (LP') und (ILP) überein.

*Beweis:* Wir zeigen (ii)  $\Rightarrow$  (i)  $\Rightarrow$  (iii)  $\Rightarrow$  (i). Sei also jede Ecke von  $P$  ganzzahlig. Wegen  $P \cap \mathbb{Z}^n \subseteq P$  gilt

$$P' := \text{conv}(P \cap \mathbb{Z}^n) \subseteq \text{conv}(P) = P,$$

also bleibt für (i) zu zeigen:  $P \subseteq P'$ . Wähle dazu  $x \in P$ . Wegen Lemma 20.5 kann  $x$  als Konvexkombination von Ecken von  $P$  dargestellt werden, also

$$x = \sum_{i=1}^k \lambda_i x_i$$

mit  $\lambda_1, \dots, \lambda_k \geq 0$ ,  $\sum_{i=1}^k \lambda_i = 1$  und  $x_1, \dots, x_k$  Ecken von  $P$ . Wegen (ii) gilt  $x_i \in \mathbb{Z}^n$  und somit  $x_i \in P \cap \mathbb{Z}^n$  für alle  $i = 1, \dots, k$ , also  $x \in \text{conv}(P \cap \mathbb{Z}^n) = P'$ .

Die Implikation (i)  $\Rightarrow$  (iii) folgt sofort aus Satz 20.1.

Abschließend gelte also (iii), zu zeigen ist (ii). Sei dazu  $\{x_0\}$  Ecke von  $P$ . Nach Satz 13.12 (2) existiert ein  $c \in \mathbb{R}^n$ , so dass  $x_0$  eindeutige Optimallösung von

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & Ax \leq b \\ & x \in \mathbb{R}^n \end{aligned}$$

ist. Nach (iii) existiert ein  $x_1 \in \mathbb{Z}^n$  mit  $Ax_1 \leq b$  und  $c^\top x_1 = c^\top x_0$ , wegen der Eindeutigkeit ist also  $x_0 = x_1 \in \mathbb{Z}^n$ .  $\square$

### Bemerkung 20.7

Die Frage, ob ein gegebenes Polytop ganzzahlig ist oder nicht, ist im Einzelfall oft sehr schwierig zu beantworten. Bei der Lösung von kombinatorischen Optimierungsproblemen ist das Polytop  $P$  oft als konvexe Hülle gegeben, also  $P = \text{conv}(X)$  für eine endliche Menge zulässiger Lösungen  $X$ . Gesucht sind dann eine Matrix  $A'$  und eine rechte Seite  $b'$  mit  $P = \{x \in \mathbb{R}^n \mid A'x \leq b'\}$ , die nach dem Satz von Weyl-Minkowski existieren müssen (wie oben bereits erwähnt). Selbst wenn  $A'$  und  $b'$  gegeben sind, kann der Beweis der sogenannten *Vollständigkeit der polyedrischen Beschreibung*, also der Beweis von  $\text{conv}(X) = \{x \in \mathbb{R}^n \mid A'x \leq b'\}$ , sehr kompliziert sein. Dabei ist es in der Regel einfach zu zeigen, dass  $X = \{x \in \mathbb{R}^n \mid A'x \leq b'\} \cap \mathbb{Z}^n$  gilt. In diesem Fall bleibt also noch zu beweisen, dass  $\{x \in \mathbb{R}^n \mid A'x \leq b'\}$  ein ganzzahliges Polytop ist, daraus folgt dann

$$\text{conv}(X) = \text{conv}(\{x \in \mathbb{R}^n \mid A'x \leq b'\} \cap \mathbb{Z}^n) = \{x \in \mathbb{R}^n \mid A'x \leq b'\}.$$

Diese Überlegungen sind Gegenstand der Vorlesung "Diskrete Optimierung".  $\diamond$

## § 21 Branch-and-Bound

Wir betrachten das ganzzahlige lineare Programm

$$\left. \begin{array}{ll} \text{minimiere} & c^\top x \\ \text{s.t.} & Ax \leq b \\ & x \in \mathbb{Z}^n \end{array} \right\} \quad (\text{ILP})$$

und dessen sogenannte *LP-Relaxierung*

$$\left. \begin{array}{ll} \text{minimiere} & c^\top x \\ \text{s.t.} & Ax \leq b \\ & x \in \mathbb{R}^n. \end{array} \right\} \quad (\text{LPR})$$

Falls eine Optimallösung  $x^*$  von (LPR) existiert, die ganzzahlig ist, ist  $x^*$  auch optimal für (ILP). Im Allgemeinen wird  $x^*$  jedoch fraktional sein. Wir benötigen also einen Algorithmus, der auch in diesem Fall eine optimale Lösung von (ILP) berechnet.

### Definition 21.1

Sei  $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$  und  $x^*$  eine optimale Lösung von  $\min_{x \in P} c^\top x$ . Falls  $x_i^* \notin \mathbb{Z}$  für ein  $i \in \{1, \dots, n\}$ , heißt die Zerlegung von  $P$  in zwei Teilpolyeder

$$P_1 = \{x \in P \mid x_i \geq \lceil x_i^* \rceil\}, \quad P_2 = \{x \in P \mid x_i \leq \lfloor x_i^* \rfloor\}$$

ein *Branching* bezüglich  $x_i$ .

### Lemma 21.2

Sei  $x^*$  eine zulässige Lösung von  $\min_{x \in P \cap \mathbb{Z}^n} c^\top x$  und  $P_1, P_2$  ein Branching bzgl.  $x_i$ . Dann gilt  $x^* \in P_1 \cup P_2$  und

$$\min_{x \in P \cap \mathbb{Z}^n} c^\top x = \min \left\{ \min_{x \in P_1 \cap \mathbb{Z}^n} c^\top x, \min_{x \in P_2 \cap \mathbb{Z}^n} c^\top x \right\}.$$

*Beweis:* Klar, da für jedes  $x_i \in \mathbb{Z}$  gilt:  $x_i \geq \lceil x_i^* \rceil$  oder  $x_i \leq \lfloor x_i^* \rfloor$ .  $\square$



**Beispiel 21.3**

Betrachte

$$\begin{aligned} \min \quad & -x_1 - 3x_2 \\ \text{s.t.} \quad & -x_1 + 6x_2 \leq 12, \quad x_1, x_2 \geq 0 \\ & x_1 + x_2 \leq 4, \quad x_1, x_2 \in \mathbb{Z}. \end{aligned}$$

Die Optimallösung ist  $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ , die optimale Lösung der LP-Relaxierung ist  $x^* = \begin{pmatrix} \frac{12}{7} \\ \frac{16}{7} \end{pmatrix}$ .

Wähle  $x_1$  als Branching-Variable, d.h.

$$P_1 = \{x \in P \mid x_1 \geq \lceil x_1^* \rceil = 2\}$$

$$P_2 = \{x \in P \mid x_1 \leq \lfloor x_1^* \rfloor = 1\}$$

Die optimale LP-Lösung für  $P_1$  ist  $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ , also ganzzahlig!

Die optimale LP-Lösung für  $P_2$  ist  $\begin{pmatrix} 1 \\ \frac{13}{6} \end{pmatrix}$ , also weiteres Branching notwendig!  $\diamond$

Die Idee des *Boundings* ist nun die folgende: Falls eine zulässige Lösung  $\bar{x} \in P \cap \mathbb{Z}^n$  bekannt ist mit  $\min_{x \in P_j} c^\top x \geq c^\top \bar{x}$ , kann  $P_j$  keine echt bessere ganzzahlige Lösung als  $\bar{x}$  enthalten. Das Teilproblem  $P_j$  muss also bei der Suche nach der Optimallösung von (ILP) nicht weiter betrachtet werden!

Im Beispiel: Aus dem Teilproblem  $P_1$  ist  $\bar{x} = (2, 2)^\top$  als zulässige Lösung bekannt mit  $c^\top \bar{x} = -8$ . Im Teilproblem  $P_2$  ist der Optimalwert  $-\frac{15}{2} \geq -8$ . Damit ist auch das Teilproblem  $P_2$  erledigt und  $\bar{x}$  somit Optimallösung von (ILP).

**Algorithmus 21.4 (Branch & Bound)**

Gegeben:  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $c \in \mathbb{R}^n$

Gesucht: Optimallösung von (ILP)

- 1: Setze  $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ ,  $L = \{P\}$ ,  $U = \infty$ ,  $\bar{x}$  undefiniert
- 2: **if**  $L = \emptyset$  **then**
- 3:   **if**  $U = \infty$  **then**
- 4:     **stop:** (ILP) hat keine zulässige Lösung
- 5:   **else**
- 6:     **stop:**  $\bar{x}$  ist Optimallösung von (ILP) mit Wert  $U$
- 7:   **end if**
- 8: **end if**
- 9: Wähle ein  $P \in L$  und setze  $L := L \setminus \{P\}$
- 10: Löse das LP

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & x \in P \end{aligned} \tag{*}$$

- 11: **if** (\*) zulässig **then**
- 12:   Wähle Optimallösung  $x^*$  von (\*) und setze  $c^* := c^\top x^*$   
     (Falls (\*) unbeschränkt ist: hier nicht behandelt...)
- 13:   **if**  $c^* < U$  **then**
- 14:     **if**  $x^* \in \mathbb{Z}^n$  **then**
- 15:       Setze  $\bar{x} := x^*$  und  $U := c^*$
- 16:     **else**
- 17:       Wähle  $i \in \{1, \dots, n\}$  mit  $x_i^* \notin \mathbb{Z}$

```
18:     Sei  $P_1, P_2$  das Branching von  $P$  bzgl.  $x_i$ 
19:     Setze  $L := L \cup \{P_1, P_2\}$ 
20:   end if
21: end if
22: end if
23: Gehe zu 2.
```

**Bemerkung 21.5**

Die Anzahl der im Schritt 10 in Algorithmus 21.4 gelösten LPs ist im Allgemeinen *nicht* polynomiell! Wichtig für eine schnelle Laufzeit des Verfahrens sind

- (1) *primale Heuristiken*: Suche zwischendurch nach guten zulässigen Lösungen, um die Schranke  $U$  zu verbessern.
- (2) eine geschickte *Zerlegung* des Problems: Die Wahl der Branching-Variablen in Schritt 17 ist wichtig, auch andere Arten von Zerlegungen sind möglich.
- (3) eine geschickte *Enumerationsstrategie*: Die Wahl des jeweils nächsten Teilproblems in Schritt 9 hat einen großen Einfluss auf die Laufzeit.
- (4) die schnelle *Reoptimierung* nach Aufteilung des Problems mithilfe des dualen Simplexalgorithmus (siehe § 19).
- (5) die geschickte Kombination mit *Schnittebenenverfahren* (siehe dazu die Vorlesung "Diskrete Optimierung").

Teil 4

# Nichtlineare Optimierung



## Freie Optimierung

### Inhalt

---

22	Existenz optimaler Lösungen	93
23	Optimalitätsbedingungen	95
24	Algorithmen der freien Optimierung	97

---

### § 22 Existenz optimaler Lösungen

Zum Abschluss befassen wir uns mit Aufgaben der Form

$$\left. \begin{array}{l} \text{minimiere } f(x) \\ \text{s.t. } x \in \mathcal{F}, \end{array} \right\} \quad (22.1)$$

wobei  $\mathcal{F} \subseteq \mathbb{R}^n$  die zulässige Menge und  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  die Zielfunktion bezeichnet, von der wir i.A. fordern, dass sie mindestens stetig auf  $\mathbb{R}^n$  ist. Wir werden zwei Fälle von zulässigen Mengen unterscheiden, nämlich

- (1)  $\mathcal{F} = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\}$ , wobei  $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$  und  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  gegebene, i.A. nichtlineare Funktionen sind. In diesem Fall entspricht (22.1) einem (allgemeinen) *nichtlinearen Programm* (NLP).
- (2)  $\mathcal{F} = \mathbb{R}^n$ . In diesem Fall liegen keine Restriktionen vor, man spricht von einer *freien Optimierungsaufgabe*.

Bevor wir im Folgenden freie Optimierungsprobleme näher untersuchen, untersuchen wir zunächst die Frage der Existenz (global) optimaler Lösungen für den allgemeinen Fall (22.1). Es gilt der folgende aus der Analysis bekannte Satz:

**Satz 22.1 (Weierstraß)**

Sei  $\mathcal{F} \subset \mathbb{R}^n$  abgeschlossen und beschränkt und  $f$  auf  $\mathcal{F}$  stetig. Dann nimmt  $f$  sein Maximum und Minimum auf  $\mathcal{F}$  an.

*Beweis:* Wir betrachten exemplarisch das Minimum und definieren  $m := \inf_{x \in \mathcal{F}} f(x) \in \mathbb{R} \cup \{-\infty\}$ , d.h. es gibt eine Infimalfolge  $\{x^k\}_{k \in \mathbb{N}} \subset \mathcal{F}$  mit  $f(x^k) \rightarrow m$  für  $k \rightarrow \infty$ . Nach Voraussetzung ist  $\mathcal{F}$  beschränkt und abgeschlossen, also nach dem Satz von Heine-Borel kompakt. Daher existiert eine konvergente Teilfolge  $\{x_{k_\ell}\}_{\ell \in \mathbb{N}}$  mit Grenzwert  $x^* \in \mathcal{F}$ . Aus der Stetigkeit von  $f$  folgt dann

$$f(x^*) = \lim_{\ell \rightarrow \infty} f(x_{k_\ell}) = m, \quad (22.2)$$

woraus  $m > -\infty$  und wegen  $x^* \in \mathcal{F}$  außerdem  $f(x^*) = m$  und damit die Optimalität von  $x^*$  folgt. □

Es stellt sich die Frage, was passiert, wenn die zulässige Menge nicht kompakt ist, wie beispielsweise im Fall der freien Optimierung. Um dies zu klären, definieren wir

**Definition 22.2**

Sei  $z \in \mathcal{F}$  gegeben. Dann heißt

$$N_f(z) := \{x \in \mathcal{F} : f(x) \leq f(z)\}$$

Niveaumenge von  $f$  in  $z$ . ◇

Dann gilt für jedes  $z \in \mathcal{F}$ , dass

$$(22.1) \quad \Leftrightarrow \quad \left\{ \begin{array}{l} \text{minimiere } f(x) \\ \text{s.t. } x \in N_f(z), \end{array} \right\} \quad (22.3)$$

denn:

„ $\Rightarrow$ “: Löst  $x^*$  Aufgabe (22.1), dann gilt für jedes  $z \in \mathcal{F}$ , dass

$$f(x^*) \leq f(x) \quad \forall x \in N_f(z) \subset \mathcal{F}.$$

Diese Ungleichung gilt natürlich auch für  $z \in N_f(z)$ , woraus  $x^* \in N_f(z)$  folgt. Also löst  $x^*$  auch (22.3).

„ $\Leftarrow$ “: Ist  $x^*$  optimal für (22.3), dann folgt zunächst  $x^* \in N_f(z) \subset \mathcal{F}$ , also Zulässigkeit für (22.1). Außerdem gilt

$$\begin{aligned} f(x^*) &\leq f(x) \quad \forall x \in N_f(z) \\ \text{und } f(x^*) &\leq f(z) < f(x) \quad \forall x \in \mathcal{F} \setminus N_f(z), \end{aligned}$$

was die Optimalität von  $x^*$  für (22.1) impliziert.

Zusammen mit Satz 22.1 folgt daraus:

**Korollar 22.3**

Sei  $f$  auf  $\mathcal{F}$  stetig. Existiert ein  $z \in \mathcal{F}$ , so dass  $N_f(z)$  kompakt ist, dann gibt es einen globalen Minimierer von  $f$  auf  $\mathcal{F}$ . ◇

**Beispiel 22.4**

Als Beispiel für eine Aufgabe mit kompakter Niveaumenge betrachten wir das folgende freie Optimierungsproblem:

$$\text{minimiere } f(x) := \frac{1}{2} x^\top Q x + b^\top x + c \quad (22.4)$$

mit  $b \in \mathbb{R}^n$  und  $c \in \mathbb{R}$ . Zudem ist  $Q \in \mathbb{R}^{n \times n}$  eine symmetrische, positiv definite Matrix. Ist  $\lambda_{\min}$  der kleinste Eigenwert von  $Q$ , dann gilt

$$f(x) \geq \frac{\lambda_{\min}}{2} \|x\|^2 + b^\top x + c \geq \left( \frac{\lambda_{\min}}{2} \|x\| - \|b\| \right) \|x\| - |c| \quad \forall x \in \mathbb{R}^n, \quad (22.5)$$

wobei wir die Cauchy-Schwarzsche Ungleichung verwendet haben. Sei nun  $z \in \mathbb{R}^n$  beliebig. Dann folgt aus (22.5), dass

$$f(x) > \|x\| - |c| \geq f(z)$$

$$\text{für alle } x \text{ mit } \|x\| > R := \max\{f(z) + |c|, (2 + 2\|b\|)/\lambda_{\min}\}.$$

Daraus folgt  $N_f(z) \subset \overline{B_R(0)}$ , was die Beschränktheit von  $N_f(z)$  beweist. Die Abgeschlossenheit von  $N_f(z)$  folgt sofort aus der Stetigkeit von  $f$ . ◇

Verallgemeinerungen von Korollar 22.3 und allgemeine hinreichende Kriterien für kompakte Niveaumengen werden in der Vorlesung ‘‘Nichtlineare Optimierung’’ behandelt. Man beachte, dass die Kompaktheit der Niveaumenge natürlich nicht notwendig für die Existenz von Minimierern ist, wie das Beispiel  $\min_{x \in \mathbb{R}} \sin(x)$  zeigt.

### § 23 Optimalitätsbedingungen

Wir betrachten im Folgenden freie Optimierungsaufgaben der Form

$$\underset{x \in \mathbb{R}^n}{\text{minimiere}} \quad f(x), \quad (23.1)$$

d.h. die zulässige Menge ist der ganze Raum,  $\mathcal{F} = \mathbb{R}^n$ . Anders als in § 11 gehen wir aber hier in der Regel davon aus, dass die Funktion  $f$  differenzierbar ist.

Die notwendigen Optimalitätsbedingungen für (23.1) lassen sich sofort aus denen der konvexen Optimierung ableiten.

#### Lemma 23.1

Sei  $U \subseteq \mathbb{R}^n$  ein linearer Unterraum und  $x^* \in U$  ein lokaler Minimierer von

$$\left. \begin{array}{l} \text{minimiere} \quad f(x) \\ \text{s.t.} \quad x \in U. \end{array} \right\}$$

Ist  $f$  in  $x^*$  differenzierbar, dann gilt

$$\nabla f(x^*)^\top h = 0 \quad \forall h \in U. \quad (23.2)$$

*Beweis:* Sei  $h \in U$  beliebig. Da  $U$  als linearer Unterraum konvex ist, folgt die Aussage unmittelbar aus Satz 9.2, wenn man die Variationsungleichung (VI) mit  $x = \pm h + x^* \in U$  testet.  $\square$

Für lineare Unterräume als zulässige Mengen gilt also eine Variationsgleichung als notwendige Optimalitätsbedingung. Im Fall  $U = \mathbb{R}^n$  erhält man die folgende Aussage (die sich im Fall einer konvexen Funktion  $f$  auch als Spezialfall von Theorem 11.2 ergibt):

#### Satz 23.2

Sei  $x^*$  ein lokaler Minimierer von (23.1) und  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  in  $x^*$  differenzierbar. Dann gilt

$$\nabla f(x^*) = 0. \quad (23.3)$$

*Beweis:* Nach Lemma 23.1 gilt (23.2) mit  $U = \mathbb{R}^n$ . Setzen wir  $h = \nabla f(x^*)$  in (23.2) ein, dann folgt  $\|\nabla f(x^*)\|^2 = 0$  und damit die Aussage.  $\square$

#### Definition 23.3

Die Lösungen von (23.3) heißen *stationäre Punkte* von  $f$ .  $\diamond$

Natürlich müssen stationäre Punkte keine Minimierer, sondern können auch Maximierer oder Sattelpunkte sein. Lediglich im Fall einer konvexen Zielfunktion sind die notwendigen Optimalitätsbedingungen auch hinreichend, wie wir in Satz 9.4 gesehen hatten. Das bedeutet, ist  $f$  konvex, dann sind die Lösungen von (23.3) (lokale = globale) Minimierer von  $f$ . Ist  $f$  hingegen nicht konvex, benötigt man zusätzliche *hinreichende Optimalitätsbedingungen*, die sich mit Hilfe der zweiten Ableitung formulieren lassen. Um die Güte der hinreichenden Bedingungen zu beurteilen, benötigen wir zunächst *notwendige Optimalitätsbedingungen zweiter Ordnung*:

#### Proposition 23.4

Sei  $x^* \in \mathbb{R}^n$  ein lokaler Minimierer von (23.1) und  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  in einer Umgebung  $B_\rho(x^*)$ ,  $\rho > 0$ , von  $x^*$  zweimal stetig differenzierbar. Dann gilt neben  $\nabla f(x^*) = 0$ , dass

$$s^\top \nabla^2 f(x^*) s \geq 0 \quad \forall s \in \mathbb{R}^n, \quad (23.4)$$

d.h.  $\nabla^2 f(x^*)$  ist positiv semidefinit.

*Beweis:* Wir nehmen o.B.d.A. an, dass  $B_\rho(x^*)$  in der Umgebung der lokalen Optimalität von  $x^*$  enthalten ist. Dann ist  $f : B_\rho(x^*) \rightarrow \mathbb{R}$  zweimal stetig differenzierbar, und es gilt  $f(x) \geq f(x^*)$  für alle  $x \in B_\rho(x^*)$ . Sei nun  $s \in \mathbb{R}^n$  beliebig. Dann ist  $x + ts \in B_\rho(x^*)$  für  $t > 0$  hinreichend klein, und der Mittelwertsatz liefert

$$0 \leq f(x^* + ts) - f(x^*) = t \nabla f(x^*)^\top s + \frac{t^2}{2} s^\top \nabla^2 f(x^* + \tau_t s) s$$

mit einem  $\tau_t \in [0, t]$ . Nun ist  $\nabla f(x^*) = 0$  nach Satz 23.2 und daher liefert Division durch  $t^2/2$  wegen Stetigkeit von  $\nabla^2 f$  auf  $B_\rho(x^*)$ , dass

$$s^\top \nabla^2 f(x^*) s = \lim_{t \searrow 0} s^\top \nabla^2 f(x^* + \tau_t s) s \geq 0.$$

Da  $s \in \mathbb{R}^n$  beliebig war, folgt die Aussage.  $\square$

Der nächste Satz gibt hinreichende Optimalitätsbedingungen zweiter Ordnung an. Sie unterscheiden sich von den notwendigen Bedingungen zweiter Ordnung lediglich dadurch, dass  $\nabla^2 f$  nicht nur positiv semidefinit sondern positiv definit sein muss. Man kann in diesem Sinne sagen, die hinreichenden Bedingungen sind „scharf“, da sie sich nur minimal von den notwendigen unterscheiden.

### Satz 23.5

Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  zweimal stetig differenzierbar auf einer offenen Menge  $U \subset \mathbb{R}^n$ . Gilt in  $x^* \in U$  die folgende *hinreichende Optimalitätsbedingung zweiter Ordnung*

- (1)  $x^*$  ist stationär, d.h.  $\nabla f(x^*) = 0$
- (2)  $\nabla^2 f(x^*)$  ist positiv definit, d.h.  $s^\top \nabla^2 f(x^*) s > 0$  für alle  $s \in \mathbb{R}^n$ ,  $s \neq 0$ ,

dann existieren  $\varepsilon > 0$  und  $\mu > 0$ , so dass

$$f(x) - f(x^*) \geq \frac{\mu}{4} \|x - x^*\|^2 \quad \forall x \in B_\varepsilon(x^*), \quad (23.5)$$

d.h.  $x^*$  ist ein isolierter lokaler Minimierer von (23.1).

*Beweis:* Da  $\nabla^2 f(x^*)$  positiv definit ist, existiert ein  $\mu > 0$  (kleinster Eigenwert), so dass

$$s^\top \nabla^2 f(x^*) s \geq \mu \|s\|^2 \quad \forall s \in \mathbb{R}^n. \quad (23.6)$$

Für  $\varepsilon > 0$  klein genug ist  $B_\varepsilon(x^*) \subset U$ , und die Stetigkeit von  $\nabla^2 f$  impliziert

$$\|\nabla^2 f(x^*) - \nabla^2 f(y)\| \leq \frac{\mu}{2} \quad \forall y \in B_\varepsilon(x^*), \quad (23.7)$$

falls  $\varepsilon > 0$  hinreichend klein ist. Für jedes  $y \in B_\varepsilon(x^*)$  folgt damit

$$\begin{aligned} s^\top \nabla^2 f(y) s &= s^\top \nabla^2 f(x^*) s + s^\top (\nabla^2 f(y) - \nabla^2 f(x^*)) s \\ &\geq \mu \|s\|^2 - \|\nabla^2 f(y) - \nabla^2 f(x^*)\| \|s\|^2 \geq \frac{\mu}{2} \|s\|^2 \quad \forall s \in \mathbb{R}^n, \end{aligned} \quad (23.8)$$

also positive Definitheit von  $\nabla^2 f(y)$ . Sei nun  $x \in B_\varepsilon(x^*)$  beliebig. Dann folgt aus dem Mittelwertsatz die Existenz eines  $t \in [0, 1]$  mit

$$\begin{aligned} f(x) - f(x^*) &= \nabla f(x^*)^\top (x - x^*) + \frac{1}{2} (x - x^*)^\top \nabla^2 f(x^* + t(x - x^*)) (x - x^*) \\ &= \frac{1}{2} (x - x^*)^\top \nabla^2 f(x^* + t(x - x^*)) (x - x^*), \end{aligned}$$



wobei wir für die letzte Gleichheit die Stationarität von  $x^*$  benutzt haben. Aus  $x, x^* \in B_\varepsilon(x^*)$  folgt aber  $x^* + t(x - x^*) \in B_\varepsilon(x^*)$  für alle  $t \in [0, 1]$  und damit wegen (23.8) schließlich

$$f(x) - f(x^*) \geq \frac{\mu}{4} \|x - x^*\|^2.$$

Da  $x \in B_\varepsilon(x^*)$  beliebig war, folgt die Aussage.  $\square$

### Bemerkung 23.6

Die Ungleichung (23.5) wird auch als *quadratische Wachstumsbedingung* bezeichnet. Sie garantiert lokal um  $x^*$  mindestens ein quadratisches Anwachsen der Zielfunktion.  $\diamond$

## § 24 Algorithmen der freien Optimierung

Wir wollen im Folgenden einen Algorithmus zur Lösung von (23.1) konstruieren. Die Idee dabei ist, ausgehend von einer gegebenen Iterierten  $x^k \in \mathbb{R}^n$ ,  $k \in \mathbb{N}$ , eine nächste Iterierte  $x^{k+1}$  zu bestimmen, die einen besseren, d.h. kleineren Zielfunktionswert liefert. Da der Gradient  $\nabla f(x^k)$  bekanntlich in die Richtung der steilsten Steigung von  $f$  zeigt, könnte es eine gute Idee sein, in entgegengesetzter Richtung, d.h. in Richtung  $-\nabla f(x^k)$ , nach einer neuen Iterierten  $x^{k+1}$  zu suchen. Das führt auf folgenden Algorithmus, der auch als Spezialisierung des Subgradientenverfahrens (siehe § 11) für den Fall differenzierbarer Funktionen betrachtet werden kann.

### Algorithmus 24.1 (Gradientenverfahren)

- 1: Wähle Startpunkt  $x^0 \in \mathbb{R}^n$ , setze  $k = 0$
- 2: **while**  $\nabla f(x^k) \neq 0$  **do**
- 3:   Setze  $s^k = -\nabla f(x^k)$ .
- 4:   Bestimme eine *Schrittweite*  $\sigma_k > 0$ , so dass

$$f(x^k + \sigma_k s^k) < f(x^k). \quad (24.1)$$

- 5:   Update: Setze  $x^{k+1} = x^k + \sigma_k s^k$ ,  $k = k + 1$
- 6: **end while**

### Bemerkung 24.2

Die Abbruchbedingung in Schritt 2 motiviert sich über die notwendige Optimalitätsbedingung aus Satz 23.2. Terminiert Algorithmus 24.1 nach endlich vielen Schritten, hat man i.A. also nur einen stationären Punkt berechnet. Zur Überprüfung der Optimalität können die hinreichenden Bedingungen zweiter Ordnung aus Satz 23.5 herangezogen werden. In der numerischen Praxis wird natürlich eine Abbruchbedingung der Form  $\|\nabla f(x^k)\| < \varepsilon$  mit einer geeigneten Toleranz  $\varepsilon$  verwendet, z.B.  $\varepsilon \in [10^{-6}, 10^{-4}]$ .  $\diamond$

Dass sich immer eine passende Schrittweite finden lässt und Algorithmus 24.1 somit durchführbar ist, zeigt das folgende Lemma:

### Lemma 24.3

Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  in  $x \in \mathbb{R}^n$  differenzierbar und  $s = -\nabla f(x) \neq 0$ . Dann existiert ein  $\sigma > 0$ , so dass  $f(x + \sigma s) < f(x)$ , d.h. in Richtung  $s$  lässt sich ein Abstieg von  $f$  erzielen.

*Beweis:* Da  $f$  in  $x$  differenzierbar ist, folgt

$$\lim_{\sigma \searrow 0} \frac{f(x + \sigma s) - f(x)}{\sigma} = \nabla f(x)^\top s = -\|\nabla f(x)\|^2 < 0.$$

Für  $\sigma > 0$  hinreichend klein, folgt daraus  $f(x + \sigma s) < f(x)$  □

Unter verschärften Voraussetzungen an die Schrittweiten lassen sich gewisse allgemeine Konvergenzaussagen für Algorithmus 24.1 zeigen:

#### Satz 24.4

Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar. Algorithmus 24.1 terminiere nicht endlich und erzeuge Folgen  $\{x^k\}$ ,  $\{s^k\}$  und  $\{\sigma_k\}$ . Erfüllt die Schrittweitenfolge anstelle von (24.1) sogar die Bedingung

$$f(x^k + \sigma_k s^k) \leq f(x^k) - \theta \|\nabla f(x^k)\|^2 \quad \forall k \in \mathbb{N}_0 \quad (24.2)$$

mit einem  $\theta > 0$ , dann ist jeder Häufungspunkt  $x^*$  von  $\{x^k\}$  stationär, d.h.  $\nabla f(x^*) = 0$ .

#### Bemerkung 24.5

Die Bedingung, dass Algorithmus 24.1 nicht endlich terminiert, ist rein technisch, da wir sonst nur endlich viele Iterierte haben und gar nicht über Konvergenz reden können. Natürlich ist es der Idealfall, wenn der Algorithmus endlich terminiert.  $\diamond$

*Beweis von Satz 24.4:* Aus (24.2) folgt  $f(x^{k+1}) < f(x^k)$  für alle  $k \in \mathbb{N}_0$ , so dass die Folge  $\{f(x^k)\}$  streng monoton fällt. (Man beachte, dass  $\nabla f(x^k) \neq 0$ , da der Algorithmus sonst terminiert wäre.) Daher gilt

$$\lim_{k \rightarrow \infty} f(x^k) = f^* \in \mathbb{R} \cup \{-\infty\}. \quad (24.3)$$

Sei nun  $x^* \in \mathbb{R}^n$  ein beliebiger Häufungspunkt von  $\{x^k\}$  mit zugehöriger Teilfolge  $\{x^{k_\ell}\}_{\ell \in \mathbb{N}}$ , d.h.  $\lim_{\ell \rightarrow \infty} x^{k_\ell} = x^*$ . Die Stetigkeit von  $f$  und (24.3) implizieren dann

$$f^* = \lim_{k \rightarrow \infty} f(x^k) = \lim_{\ell \rightarrow \infty} f(x^{k_\ell}) = f(x^*),$$

woraus  $f^* > -\infty$  und  $\lim_{k \rightarrow \infty} f(x^k) = f(x^*)$ , also die Konvergenz der Zielfunktionswerte folgt. Als konvergente Folge ist  $\{f(x^k)\}$  eine Cauchy-Folge, so dass

$$0 \leq \|\nabla f(x^k)\|^2 = \frac{1}{\theta} (f(x^k) - f(x^k + \sigma_k s^k)) = \frac{1}{\theta} (f(x^k) - f(x^{k+1})) \rightarrow 0 \quad \text{für } k \rightarrow \infty,$$

und damit

$$\nabla f(x^k) \rightarrow 0 \quad \text{für } k \rightarrow \infty.$$

Daraus folgt aber erst recht  $\nabla f(x^{k_\ell}) \rightarrow 0$  für  $\ell \rightarrow \infty$ , so dass die stetige Differenzierbarkeit von  $f$  gerade

$$\nabla f(x^*) = \lim_{\ell \rightarrow \infty} \nabla f(x^{k_\ell}) = 0,$$

liefert. □

Schrittweiten, die (24.2) erfüllen, heißen *effizient*. Ein Verfahren, das unter geeigneten Voraussetzungen solche Schrittweiten berechnet, ist das *Armijo-Verfahren* (vgl. [?, Satz 4.5.5]).

#### Algorithmus 24.6 (Armijo-Verfahren)

- 1: Gegeben sei ein Punkt  $x \in \mathbb{R}^n$  mit  $s = -\nabla f(x)$
- 2: Wähle  $\beta \in (0, 1)$ ,  $\gamma \in (0, 1)$ ,  $\sigma^{(0)} > 0$ , setze  $m = 0$

```
3: while  $f(x) - f(x + \sigma^{(m)}s) < \gamma \sigma^{(m)} \|\nabla f(x)\|^2$  do  
4:   Setze  $\sigma^{(m+1)} = \beta \sigma^{(m)}$ ,  $m = m + 1$   
5: end while
```

In der Vorlesung “Nichtlineare Optimierung” werden das Armijo-Verfahren näher analysiert und weitere Schrittweiten-Verfahren vorgestellt.

In der numerischen Praxis stellt man schnell fest, dass das Gradientenverfahren i.d.R. nur sehr langsam konvergiert. Bessere Konvergenzeigenschaften weisen Verfahren auf, die neben der ersten auch die zweite Ableitung der Zielfunktion, also Krümmungsinformation, berücksichtigen. Beispielsweise kann man anstelle des Anti-Gradienten in Schritt 3 von Algorithmus 24.1 die so genannte *Newton-Richtung* als Suchrichtung wählen:

$$s^k = -\nabla^2 f(x^k)^{-1} \nabla f(x^k),$$

wobei  $\nabla^2 f(x^k) \in \mathbb{R}^{n \times n}$  die Hesse-Matrix von  $f$  bezeichnet. Das so entstehende Verfahren wird *globalisiertes Newton-Verfahren* genannt und konvergiert unter geeigneten Voraussetzungen lokal quadratisch, wie in der Vorlesung “Nichtlineare Optimierung” gezeigt wird. Allerdings ist die Hesse-Matrix oft sehr aufwendig zu berechnen und eventuell auch nicht invertierbar. Es gibt daher eine Vielzahl von Modifikationen dieses Verfahrens, die diese Probleme umgehen. Mehr dazu in der Vorlesung “Nichtlineare Optimierung” ...