
Optimierung 1

Peter Gritzmann

Zentrum Mathematik, Technische Universität München
D-80290 München, Germany
gritzman@ma.tum.de

Kurzkript, Technische Universität München
Sommersemester 2008

Inhaltsverzeichnis

Inhaltsverzeichnis	2
1 Einleitung	3
1.1 Optimierungsaufgaben	3
1.2 Einige Beispiele	4
1.3 Fourier-Motzkin-Elimination	9
2 Konvexitätstheorie	15
2.1 Konvexe Mengen	15
2.2 Trennungssätze, Kegel und linear-konvexe Optimierung	21
2.3 Darstellungssätze	30
3 Der Simplex-Algorithmus	41
3.1 Grundstruktur des Algorithmus	41
3.2 Finden einer Verbesserungskante	47
3.3 Über Zykel und ihre Vermeidung	55
3.4 Über die Laufzeit des Simplex-Algorithmus	66
4 LP-Dualität	72
4.1 Dualität linearer Programme	72
4.2 Dualität in ökonomischen Modellen	79
Literaturverzeichnis	80
Index	83

1 Einleitung

1.1 Optimierungsaufgaben

1.1.1 Definition. (a) Eine allgemeine **Optimierungsaufgabe** [engl.: instance of an optimization problem] (über \mathbb{R}) ist durch folgende Daten spezifiziert.

$$n \in \mathbb{N}, \quad F \subset G \subset \mathbb{R}^n, \quad \varphi : G \rightarrow \mathbb{R}, \quad \text{opt} \in \{\min, \max\}.$$

Jeder Punkt $x^* \in F$ mit

$$x \in F \quad \Rightarrow \quad \varphi(x^*) \leq \varphi(x)$$

für $\text{opt} = \min$ bzw.

$$x \in F \quad \Rightarrow \quad \varphi(x^*) \geq \varphi(x)$$

für $\text{opt} = \max$ heißt **Optimalpunkt**. Die Aufgabe besteht darin, φ über F zu optimieren (minimieren oder maximieren), d.h. zu entscheiden, ob F leer ist, ob φ nach unten (für $\text{opt} = \min$) bzw. nach oben (für $\text{opt} = \max$) beschränkt ist, ob ein Optimalpunkt existiert und, falls das der Fall ist, einen solchen zu bestimmen.

- (b) φ heißt **Zielfunktion** [engl.: objective function], F heißt **zulässiger Bereich** [engl.: feasible region] der gegebenen Aufgabe. Jeder Punkt $x \in F$ heißt **zulässig** [engl.: feasible point]. Ist der zulässige Bereich leer, so heißt die Aufgabe **unzulässig** [engl.: infeasible]. Ist $\text{opt} = \min$, so spricht man von einer **Minimierungsaufgabe**; für $\text{opt} = \max$ von einer **Maximierungsaufgabe**.
- (c) Ist $F = \mathbb{R}^n$ so spricht man von einer Aufgabe der **unrestringierten Optimierung** [engl.: unconstrained optimization], für $F \neq \mathbb{R}^n$ von einer Aufgabe der **restringierten Optimierung**.
- (d) Die Menge aller Optimierungsaufgaben heißt **Optimierungsproblem**, die Teilmengen mit $\text{opt} = \min$ bzw. $\text{opt} = \max$ heißen **Minimierungsproblem** bzw. **Maximierungsproblem**.
- (e) Besteht das Ziel nur darin, festzustellen, ob der zulässige Bereich einer gegebenen Aufgabe nicht leer ist¹, so spricht man von einer **Zulässigkeitsaufgabe**. Entsprechend ist auch das **Zulässigkeitsproblem** definiert.

1.1.2 Definition. (a) Eine **lineare Optimierungsaufgabe (in natürlicher Form)** [engl.: instance of a linear programming problem oder linear program] (über \mathbb{R}) ist durch folgende Daten spezifiziert. Seien

$$\begin{aligned} m, n &\in \mathbb{N}, \\ a_1, \dots, a_m &\in \mathbb{R}^n, \\ \beta_1, \dots, \beta_m &\in \mathbb{R}, \\ \gamma_1, \dots, \gamma_n &\in \mathbb{R}, \end{aligned}$$

¹ Das entspricht $\varphi \equiv 0$.

Ferner sei das lineare Funktional $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ für $x := (\xi_1, \dots, \xi_n)^T \in \mathbb{R}^n$ durch $\varphi(x) = \sum_{i=1}^n \gamma_i \xi_i$ definiert, und es sei F die Menge aller $x \in \mathbb{R}^n$ mit

$$a_i^T x \leq \beta_i \quad (i = 1, \dots, m).$$

Ziel ist es, φ über F zu maximieren.

(b) Ist zusätzlich eine Teilmenge

$$G \subset \{1, \dots, n\}$$

spezifiziert und die Bedingung

$$\xi_j \in \mathbb{Z} \quad (j \in G)$$

zu erfüllen, so spricht man von einer **gemischt-ganzzahligen (linearen) Optimierungsaufgabe (in natürlicher Form)** [engl.: instance of a mixed-integer (linear) programming problem]. Für $G = \{1, \dots, n\}$ heißt die Aufgabe **ganzzahlige (lineare) Optimierungsaufgabe**.

(c) Die Menge aller linearen Optimierungsaufgaben (in natürlicher Form) heißt **lineares Optimierungsproblem (in natürlicher Form)**; die Menge aller gemischt-ganzzahligen (linearen) Optimierungsaufgaben (in natürlicher Form) heißt **gemischt-ganzzahliges (lineares) Optimierungsproblem (in natürlicher Form)**; die Menge aller ganzzahligen (linearen) Optimierungsaufgaben heißt **ganzzahliges (lineares) Optimierungsproblem**. Kürzer werden auch die Bezeichnungen **LP-Problem**, von **MILP-Problem** (mixed-integer linear programming problem) bzw. **ILP-Problem** (integer linear programming problem) verwendet.

1.1.3 Definition. Sei $P \subset \mathbb{R}^n$. P heißt **Polyeder** [engl.: polyhedron], wenn es

$$n, m \in \mathbb{N}_0, \quad A \in \mathbb{R}^{m \times n}, \quad b \in \mathbb{R}^m$$

gibt mit

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

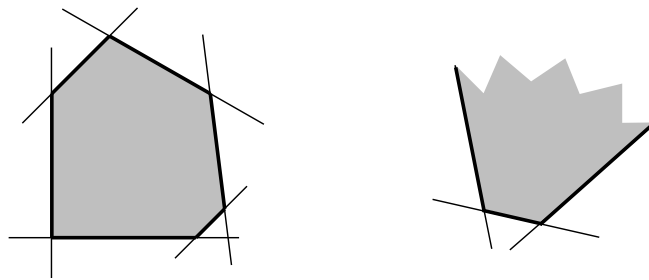
P heißt **Polytop** [engl.: polytope] genau dann, wenn P beschränkt und ein Polyeder ist.

Ist P ein Polyeder bzw. Polytop, und $n, m \in \mathbb{N}_0$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ mit $P = \{x \in \mathbb{R}^n : Ax \leq b\}$, so heißt das Tupel $(m, n; A, b)$ **\mathcal{H} -Darstellung²** [engl.: \mathcal{H} -presentation] von P . Man spricht dann von P auch als **\mathcal{H} -Polyeder** [engl.: \mathcal{H} -polyhedron] bzw. **\mathcal{H} -Polytop** [engl.: \mathcal{H} -polytope].

1.2 Einige Beispiele

Produktionsplanung:

² Diese Bezeichnung resultiert von der Interpretation der einzelnen, P darstellenden Ungleichungen als Halbräume; vgl. Definition 2.2.3.



1.1 **Abbildung** Zwei Polyeder; das links abgebildete ist ein Polytop

1.2.1 **Bezeichnung.** Eine (*lineare*) **Produktionsaufgabe** [engl.: instance of the production problem] ist spezifiziert durch folgende Daten

$$\begin{aligned} m, n &\in \mathbb{N}, \\ b &:= (\beta_1, \dots, \beta_m)^T \in [0, \infty[^m, \\ c &:= (\gamma_1, \dots, \gamma_n)^T \in [0, \infty[^n, \\ A &:= (\alpha_{i,j})_{\substack{i=1, \dots, m \\ j=1, \dots, n}} \in \mathbb{R}^{m \times n}. \end{aligned}$$

Der Vektor b quantifiziert die **verfügbare Ressourcen**, c ist der **Preisvektor**, und A heißt **Technologiematrix**.

Ferner sei das Zielfunktional $\varphi : \mathbb{R}^n \rightarrow [0, \infty[$ definiert durch $\varphi(x) := c^T x$. Ein Vektor $x \in \mathbb{R}^n$ heißt (**zulässiger**) **Produktionsplan** [engl.: production plan], wenn

$$x \geq 0 \quad \wedge \quad Ax \leq b$$

gilt. Seien F die Menge aller zulässigen Produktionspläne und $x^* \in F$. Dann heißt x^* **optimaler Produktionsplan** [engl.: optimal production plan], wenn gilt

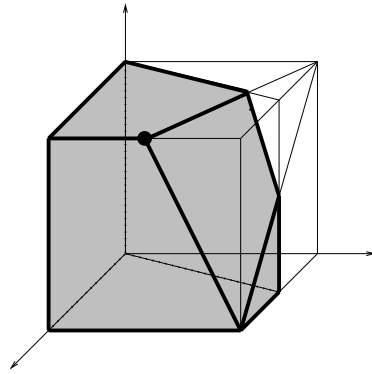
$$x \in F \quad \implies \quad \varphi(x^*) \geq \varphi(x).$$

Ziel der Produktionsaufgabe ist es zu entscheiden, ob ein optimaler Produktionsplan existiert, und falls dem so ist, einen zu finden.

Die Menge aller Produktionsaufgaben heißt (**lineares**) **Produktionsproblem** [engl.: production problem].

1.2.2 **Beispiel.** Die Produktionsaufgabe sei gegeben durch

$$\begin{array}{rcll} \max & \xi_1 + \xi_2 + \xi_3 & & \\ \xi_1 & + & 2\xi_2 & + \xi_3 \leq 3 \\ -2\xi_1 & + & \xi_2 & \leq 0 \\ \xi_1 & & & \leq 1 \\ & & \xi_2 & \leq 1 \\ & & & \xi_3 \leq 1 \\ & \xi_1, & \xi_2, & \xi_3 \geq 0 \end{array}$$



1.2 Abbildung Geometrische Darstellung des Polytops der zulässigen Produktionspläne in Beispiel 1.2.2

Standortprobleme: Ein Großhandelskonzern plant den Aufbau eines Vertriebssystems. Verträge mit Großkunden G_1, \dots, G_n sind bereits geschlossen, mögliche Standorte S_1, \dots, S_m für Versandgroßlager sind in der Beurteilung. Es soll eine optimale Auswahl der Standorte getroffen werden. Dabei treten natürlich Fixkosten κ_j für die Errichtung eines Versandlagers in Standort S_j auf. Andererseits sind für $i = 1, \dots, n$ und $j = 1, \dots, m$ die Transportkosten γ_{ij} bekannt, die auftreten, wenn man den Großkunden G_i vom Standort S_j aus beliefert, ferner ist die Gesamtliefermenge α_i an G_i vertraglich festgelegt sowie die Kapazität β_j des möglichen Lagers im Standort S_j gegeben.

Wir führen Variable $\xi_{i,j}$ ein, die angeben sollen, wieviel von Standort S_j aus an den Kunden G_i geliefert werden soll. Da der Bedarf aller Kunden gedeckt werden muss, ergeben sich die Bedingungen

$$\begin{aligned} \sum_{j=1}^m \xi_{i,j} &= \alpha_i & (i = 1, \dots, n) \\ \xi_{i,j} &\geq 0 & (i = 1, \dots, n; j = 1, \dots, m) \end{aligned}$$

Ferner modellieren wir die Entscheidung, ein Versandgroßlager am Standort S_j zu errichten, mit Hilfe einer 0-1-Variable η_j , d.h.

$$\eta_j \in \{0,1\} \quad (j = 1, \dots, m).$$

Die Kapazitätsbeschränkungen des möglichen Lagers in S_j führen dann auf die Bedingung

$$\sum_{i=1}^n \xi_{i,j} \leq \beta_j \eta_j \quad (j = 1, \dots, m).$$

Während die Variablen $\xi_{i,j}$ die potentiellen Liefermengen angeben, modellieren die 'Entscheidungsvariablen' η_j , dass nur dann eine Lieferung von S_j aus erfolgen kann, wenn an diesem Standort tatsächlich ein Lager errichtet wird. Als Zielfunktion wird die Minimierung der Kosten über den Planungshorizont angesetzt, d.h.

$$\min \sum_{i=1}^n \sum_{j=1}^m \gamma_{i,j} \xi_{i,j} + \sum_{j=1}^m \kappa_j \eta_j$$

Das Zuordnungsproblem: Formal ist für $n_1, n_2 \in \mathbb{N}$ eine Teilmenge

$$Z \subset \{1, \dots, n_1\} \times \{1, \dots, n_2\}$$

gesucht, für die einerseits

$$\begin{aligned} (i, j_1) \in Z \wedge (i, j_2) \in Z &\implies j_1 = j_2 \\ (i_1, j) \in Z \wedge (i_2, j) \in Z &\implies i_1 = i_2 \end{aligned}$$

gilt, und die andererseits

$$\sum_{(i,j) \in Z} \gamma_{i,j}$$

maximiert, wobei $\gamma_{i,j}$ für $i = 1, \dots, n_1$ und $j = 1, \dots, n_2$ den Nutzen der Zuordnung von (i, j) angibt. Die Aufgabe, eine optimale Zuordnung zu finden, läßt sich so mit Hilfe der folgenden ILP-Aufgabe beschreiben.

$$\begin{aligned} \max \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \gamma_{i,j} \xi_{i,j} \\ \sum_{j=1}^{n_2} \xi_{i,j} &\leq 1 && (i = 1, \dots, n_1) \\ \sum_{i=1}^{n_1} \xi_{i,j} &\leq 1 && (j = 1, \dots, n_2) \\ \xi_{i,j} &\in \{0, 1\} && (i = 1, \dots, n_1; j = 1, \dots, n_2). \end{aligned}$$

Das Problem des Handlungsreisenden: Das als *Traveling Salesman Problem* oder kürzer als *TSP* bekannte Rundreiseproblem ist ein Standardproblem der kombinatorischen Optimierung. Hierbei ist eine kürzeste Rundreise (Tour) zu finden, die alle vorgegebenen Städte genau einmal besucht.

1.2.3 Bemerkung. Für $n \in \mathbb{N} \setminus \{1\}$ sei τ_n die Anzahl aller verschiedenen Rundreisen durch n Städte. Dann gilt

$$\tau_n = \frac{1}{2}(n-1)! > \sqrt{\pi/2}(n-1)^{n-\frac{1}{2}}e^{-n+1}.$$

Rekonstruktion kristalliner Strukturen: Im Bereich der Nanotechnologie von Silizium-Chips sollen neue bildgebende Verfahren zur Qualitätskontrolle eingesetzt werden. Mit Hilfe moderner Methoden der hochauflösenden Transmissionselektronenmikroskopie (HRTEM) kann man für bestimmte Materialien (wie etwa Silizium; vgl. Abbildung ??) in geeigneten Richtungen bestimmen, wieviele Atome in den entsprechenden Atomsäulen enthalten sind.

Nehmen wir (der Einfachheit der Darstellung halber) an, dass die zugrunde liegende kristalline Gitterstruktur der von \mathbb{Z}^3 entspricht und die zu untersuchende Probe im Bereich $[1, q]^3$ liegt mit $q \in \mathbb{N}$. Ferner seien die Richtungen der Aufnahmen parallel zu den drei Koordinatenachsen im \mathbb{R}^3 . Dann erhält man (bis auf gewisse Messfehler) die Informationen, wieviele Atome jeweils auf jeder Geraden parallel zu einer der Koordinatenachsen liegen.

$$\begin{aligned} \sum_{i=1}^q \xi_{i,j,k} &= \alpha_{j,k} & (j,k = 1, \dots, q) \\ \sum_{j=1}^q \xi_{i,j,k} &= \beta_{i,k} & (i,k = 1, \dots, q) \\ \sum_{k=1}^q \xi_{i,j,k} &= \gamma_{i,j} & (i,j = 1, \dots, q). \end{aligned}$$

Zusammen mit 0-1-Bedingungen an die Variablen beschreibt dieses System alle kristallinen Strukturen, die mit den gegebenen Messdaten verträglich sind.

Wire Spacing: In einem Halbleiterchip sollen die ‘Streukapazitäten’ zwischen benachbarten Drähten minimiert werden. Die physikalisch-mathematische Modellierung führt somit (unter den entsprechenden Modellannahmen) zu folgendem Optimierungsproblem

1.2.4 Bezeichnung. Eine *Wire Spacing Aufgabe* ist spezifiziert durch folgende Daten

$$\begin{aligned} n &\in \mathbb{N}, \\ s_1, \dots, s_n &\in [0, \infty[, \\ r &\in]0, \infty[, \\ d &\in]0, \infty[. \end{aligned}$$

Die Aufgabe besteht dann darin, zu entscheiden, ob die Aufgabe

$$\begin{aligned} \min \sum_{i=1}^n s_i \left(\frac{1}{\xi_i} + \frac{1}{\xi_{i+1}} \right) \\ \sum_{i=1}^{n+1} \xi_i &\leq r \\ \xi_1, \dots, \xi_{n+1} &\geq d \end{aligned}$$

einen Optimalpunkt x^* besitzt und, falls das der Fall ist, einen solchen zu bestimmen.

Die Menge aller Wire Spacing Aufgabe heißt **Wire Spacing Problem**.

Eine ‘klassische Optimierungsaufgabe’: Für $n \in \mathbb{N}$ sei das folgende Optimierungsproblem im \mathbb{R}^3 gegeben

$$\begin{aligned} \min (\xi_1^n - \xi_2^n + \xi_3^n)^2 \\ \xi_1, \xi_2, \xi_3 &\geq 1 \\ \xi_1, \xi_2, \xi_3 &\in \mathbb{Z}. \end{aligned}$$

Offenbar ist $\zeta^*(n) = 0$ genau dann, wenn die Gleichung

$$\xi_1^n + \xi_3^n = \xi_2^n$$

eine positive ganzzahlige Lösung besitzt. Dass dieses für kein $n \geq 3$ der Fall ist, hat im 17. Jahrhundert bereits Pierre de Fermat behauptet; der *Satz von Fermat* konnte aber bekanntlich erst 1995 von Andrew Wiles bzw. Andrew Wiles und Richard Taylor bewiesen werden.

1.3 Fourier-Motzkin-Elimination

1.3.1 Wiederholung. Das Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn

$$\text{rang}(A) = \text{rang}(A, b)$$

gilt.

1.3.2 Beispiel. (a) Seien

$$A = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Dann gilt

$$\text{rang}(A) = 1 \neq 2 = \text{rang} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Trotzdem ist das Ungleichungssystem $Ax \leq b$ lösbar; es gilt

$$\{x \in \mathbb{R}^1 : Ax \leq b\} = [-1, 1].$$

(b) Für

$$A = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

gilt ebenfalls

$$\text{rang}(A) = 1 \neq 2 = \text{rang} \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix},$$

aber diesmal ist $Ax \leq b$ unlösbar.

1.3.3 Beispiel. (a) Seien

$$A = (1, 1) \quad b = 1.$$

Dann gilt

$$\{x \in \mathbb{R}^2 : Ax = b \wedge x \geq 0\} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + [0, 1] \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

(b) Für

$$A = (1, 1), \quad b = -1$$

ist das Ungleichungssystem $Ax = b \wedge x \geq 0$ hingegen unlösbar; vgl. Abbildung ??.

1.3.4 Beispiel. Gegeben sei das Ungleichungssystem

$$\xi_1 + \xi_2 \leq \alpha \quad \wedge \quad \xi_1 \geq 0 \quad \wedge \quad \xi_2 \geq 0,$$

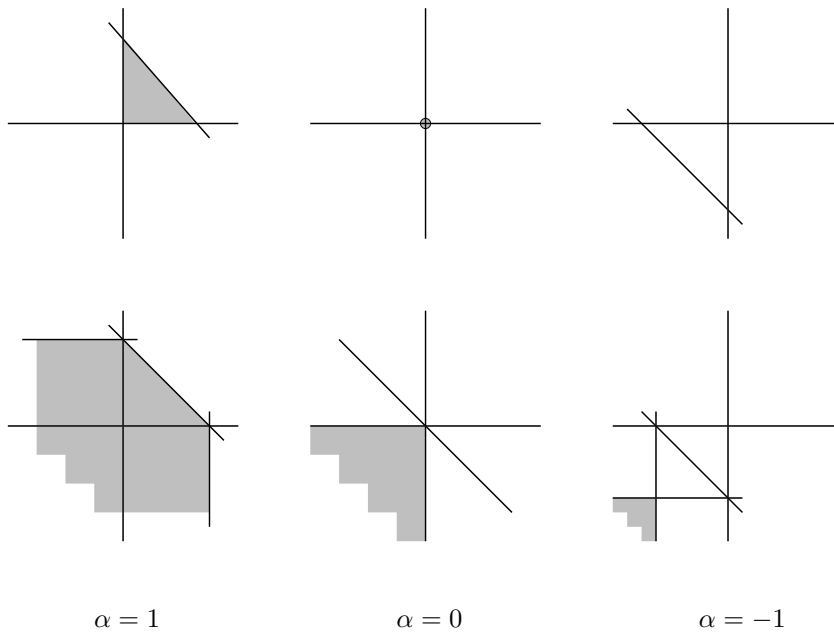
wobei der Parameter α später festgesetzt wird. Es ist offenbar äquivalent zu

$$\begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \leq \begin{pmatrix} \alpha \\ 0 \\ 0 \end{pmatrix}.$$

Addition der ersten Ungleichung zur zweiten und dritten liefert das neue System

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \leq \begin{pmatrix} \alpha \\ \alpha \\ \alpha \end{pmatrix}.$$

Natürlich ist jede Lösung des ursprünglichen Problems auch Lösung des durch ‘elementare Zeilenoperationen’ entstandenen neuen Systems. Die Skizzen in Abbildung 1.3 veranschaulichen jeweils den zulässigen Bereich des ursprünglichen sowie des neuen Systems für die Parameterwerte $\alpha = 1, 0, -1$.



1.3 Abbildung Lösungsmengen des ursprünglichen und (darunter) des neuen Systems für $\alpha = 1, 0, -1$ (jeweils von links)

Gegeben sei zunächst das lineare Gleichungssystem

$$Ax = b.$$

Es sei

$$A := (\alpha_{i,j})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \quad \wedge \quad b := (\beta_1, \dots, \beta_m)^T,$$

und wir nehmen o.B.d.A. an, dass die Gleichungen so nummeriert sind, dass

$$\alpha_{i,n} \begin{cases} < 0 & \text{für } i = 1, \dots, k \\ = 0 & \text{für } i = k + 1, \dots, l \\ > 0 & \text{für } i = l + 1, \dots, m \end{cases}$$

für $k \in \mathbb{N}_0$ und $l - k \in \mathbb{N}_0$ gilt. Durch Division der i -ten Zeile durch $|\alpha_{i,n}|$ für $i = 1, \dots, k$ bzw. durch $\alpha_{i,n}$ für $i = l + 1, \dots, m$ erhalten wir das äquivalente System

$$\begin{aligned} \sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j - \xi_n &= \hat{\beta}_i & (i = 1, \dots, k) \\ \sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j &= \hat{\beta}_i & (i = k+1, \dots, l) \\ \sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j + \xi_n &= \hat{\beta}_i & (i = l+1, \dots, m), \end{aligned}$$

wobei

$$\hat{\alpha}_{i,j} := \begin{cases} \alpha_{i,j}/|\alpha_{i,n}| & \text{für } i = 1, \dots, k \\ \alpha_{i,j} & \text{für } i = k+1, \dots, l \\ \alpha_{i,j}/\alpha_{i,n} & \text{für } i = l+1, \dots, m \end{cases}$$

und $j = 1, \dots, n-1$ sowie

$$\hat{\beta}_i := \begin{cases} \beta_i/|\alpha_{i,n}| & \text{für } i = 1, \dots, k \\ \beta_i & \text{für } i = k+1, \dots, l \\ \beta_i/\alpha_{i,n} & \text{für } i = l+1, \dots, m \end{cases}$$

gilt. Auflösen nach $-\xi_n$ bzw. ξ_n liefert somit

$$\begin{aligned} -\xi_n &= \hat{\beta}_i - \sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j & (i = 1, \dots, k) \\ \sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j &= \hat{\beta}_i & (i = k+1, \dots, l) \\ \xi_n &= \hat{\beta}_i - \sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j & (i = l+1, \dots, m). \end{aligned}$$

In diesem ersten Teil sind nur Operationen durchgeführt worden, die auch auf Ungleichungssysteme $Ax \leq b$ übertragbar sind.

Im Gauß-Algorithmus eliminiert man nun die Variable ξ_n und gelangt zu folgenden linearen Gleichungssystem mit $m-1$ Ungleichungen in $n-1$ Variablen.

$$\begin{aligned} \sum_{j=1}^{n-1} (\hat{\alpha}_{i,j} - \hat{\alpha}_{1,j}) \xi_j &= \hat{\beta}_i - \hat{\beta}_1 & (i = 2, \dots, k) \\ \sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j &= \hat{\beta}_i & (i = k+1, \dots, l) \\ \sum_{j=1}^{n-1} (\hat{\alpha}_{i,j} + \hat{\alpha}_{1,j}) \xi_j &= \hat{\beta}_i + \hat{\beta}_1 & (i = l+1, \dots, m). \end{aligned}$$

1.3.5 Bemerkung. Das System $Ax \leq b$ ist genau dann lösbar, wenn

$$\begin{aligned} \sum_{j=1}^{n-1} (\hat{\alpha}_{i,j} + \hat{\alpha}_{t,j}) \xi_j &\leq \hat{\beta}_t + \hat{\beta}_i & (i = 1, \dots, k; t = l+1, \dots, m) \\ \sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j &\leq \hat{\beta}_i & (i = k+1, \dots, l). \end{aligned}$$

lösbar ist. Ferner ist $\xi_1^*, \dots, \xi_{n-1}^*$ genau dann eine Lösung des reduzierten Systems, wenn es ein ξ_n^* gibt, so dass $\xi_1^*, \dots, \xi_{n-1}^*, \xi_n^*$ eine Lösung des Ausgangssystems ist. Für $n \geq 2$ ist die Lösungsmenge des reduzierten Systems somit die orthogonale Projektion des ursprünglichen auf den \mathbb{R}^{n-1} der ersten $n-1$ Koordinaten.

Beweis: Ist ξ_1^*, \dots, ξ_n^* eine Lösung des Ausgangssystems, so gilt

$$\begin{aligned} -\xi_n &\leq \hat{\beta}_i - \sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j & (i = 1, \dots, k) \\ \sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j &\leq \hat{\beta}_i & (i = k+1, \dots, l) \\ \xi_n &\leq \hat{\beta}_i - \sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j & (i = l+1, \dots, m), \end{aligned}$$

also insbesondere

$$\begin{aligned} \sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j^* - \hat{\beta}_i &\leq \hat{\beta}_t - \sum_{j=1}^{n-1} \hat{\alpha}_{t,j} \xi_j^* & (i = 1, \dots, k; t = l+1, \dots, m) \\ \sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j^* &\leq \hat{\beta}_i & (i = k+1, \dots, l), \end{aligned}$$

d.h. $\xi_1^*, \dots, \xi_{n-1}^*$ ist eine Lösung des reduzierten Systems. Ist umgekehrt $\xi_1^*, \dots, \xi_{n-1}^*$ eine Lösung des reduzierten Systems, so kann ξ_n^* so gewählt werden, dass gilt

$$\sum_{j=1}^{n-1} \hat{\alpha}_{i,j} \xi_j^* - \hat{\beta}_i \leq \xi_n^* \leq \hat{\beta}_t - \sum_{j=1}^{n-1} \hat{\alpha}_{t,j} \xi_j^* \quad (i = 1, \dots, k; t = l+1, \dots, m);$$

ξ_1^*, \dots, ξ_n^* erfüllt dann alle Bedingungen des Ausgangssystems, und es folgt die Behauptung. \square

1.3.6 Bezeichnung. Das beschriebene Verfahren heißt **Fourier-Motzkin-Elimination** [engl.: *Fourier-Motzkin elimination*] zur Lösung von linearen Ungleichungssystemen.

1.3.7 Prozedur: *Fourier-Motzkin-Elimination*

INPUT: Matrix $A \in \mathbb{R}^{m \times n}$, Vektor $b \in \mathbb{R}^m$
 OUTPUT: Vektor $x^* \in \mathbb{R}^n$ mit $Ax^* \leq b$, falls ein solcher existiert;
 Meldung 'Ungleichungssystem unlösbar' sonst
 BEGIN eliminiere sukzessive alle Variablen
 IF die entstehende variablenfreie Ungleichung ist unlösbar
 THEN Meldung 'Ungleichungssystem unlösbar'
 ELSE Bestimme sukzessive Komponenten ξ_1^*, \dots, ξ_n^* eines Lösungsvektors
 END

1.3.8 Beispiel. Wir wenden das Fourier-Motzkin-Eliminationsverfahren auf Beispiel 1.3.4 an

$$\begin{aligned} \left\{ \begin{array}{l} \xi_1 + \xi_2 \leq \alpha \\ \xi_1 \geq 0 \\ \xi_2 \geq 0 \end{array} \right\} &\rightarrow \left\{ \begin{array}{l} \xi_2 \leq \alpha - \xi_1 \\ \xi_1 \geq 0 \\ \xi_2 \geq 0 \end{array} \right\} \rightarrow \\ &\rightarrow \left\{ \begin{array}{l} 0 \leq \alpha - \xi_1 \\ \xi_1 \geq 0 \end{array} \right\} \rightarrow \left\{ \begin{array}{l} \xi_1 \leq \alpha \\ \xi_1 \geq 0 \end{array} \right\} \rightarrow \{0 \leq \alpha\} \end{aligned}$$

1.3.9 Satz. Das Verfahren der Fourier-Motzkin-Elimination löst ein lineares Ungleichungssystem $Ax \leq b$ in n Variablen und mit m Ungleichungsrestriktionen in n Eliminationsschritten.

Beweis: In jedem Schritt des Verfahrens wird eine Variable eliminiert. Somit erhält man nach n Eliminationsschritten eine variablenfreie Ungleichung. \square

1.3.10 Bemerkung. Seien $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ und $P := \{x \in \mathbb{R}^n : Ax \leq b\} \neq \emptyset$. Ferner sei $\nu \in \mathbb{R} \cup \{\infty\}$. Durch Fourier-Motzkin-Elimination der Variablen ξ_1, \dots, ξ_n gehe das Ungleichungssystem

$$\begin{array}{rcl} \xi_0 & - & c^T x \leq 0 \\ & & Ax \leq b \end{array}$$

über in

$$\xi_0 \leq \nu,$$

wobei $\nu = \infty$ bedeutet, dass ξ_0 keiner reellen oberen Schranke genügen muss. Dann gilt

$$\max\{c^T x : Ax \leq b\} = \nu.$$

Beweis: Für einen gegebenen Punkt $x^* \in P$ ist die Ungleichung $c^T x^* \geq \xi_0$ äquivalent zu $\xi_0 \in]-\infty, c^T x^*]$. Nach Bemerkung 1.3.5 gibt es genau zu jedem $\xi_0 \in]-\infty, \nu]$ Werte ξ_1, \dots, ξ_n , so dass das System

$$\begin{array}{rcl} \xi_0 & - & c^T x \leq 0 \\ & & Ax \leq b \end{array}$$

lösbar ist. Hieraus folgt die Behauptung. \square

1.3.11 Beispiel. Gegeben sei die folgende lineare Optimierungsaufgabe.

$$\begin{array}{rcl} \max & \xi_1 + \xi_2 & \\ 3\xi_1 & + \xi_3 & \leq 3 \\ -3\xi_1 & + \xi_3 & \leq 0 \\ & \xi_2 + \xi_3 & \leq 1 \\ & 2\xi_2 + \xi_3 & \geq 1 \\ & \xi_3 & \geq 0 \end{array}$$

Der zulässige Bereich ist in Abbildung 1.4 (links) skizziert. Gemäß Bemerkung 1.3.10 fügen wir die 'Zielfunktionsungleichung'

$$\xi_0 - \xi_1 - \xi_2 \leq 0$$

hinzu und führen anschließend die Elimination nach ξ_3 durch. Wir erhalten

$$\begin{array}{rcl} \xi_0 & - & \xi_1 - \xi_2 \leq 0 \\ & & 3\xi_1 - 2\xi_2 \leq 2 \\ & & 3\xi_1 + 2\xi_2 \geq 1 \\ & & \xi_2 \geq 0 \\ & & \xi_1 \leq 1 \\ & & \xi_1 \geq 0 \\ & & \xi_2 \leq 1. \end{array}$$

Elimination von ξ_2 liefert

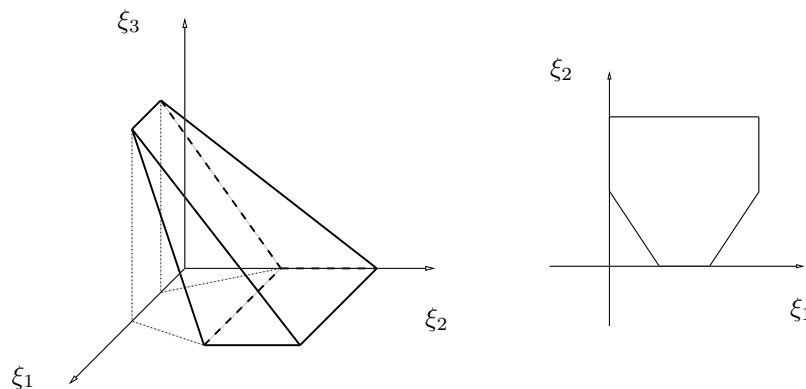
$$\begin{aligned} \xi_0 - 1 &\leq \xi_1 \\ \xi_1 &\leq \frac{4}{3} \\ -\frac{1}{3} &\leq \xi_1 \\ 0 &\leq \xi_1 \\ \xi_1 &\leq 1. \end{aligned}$$

Durch Elimination nach ξ_1 (und Weglassen der redundanten Bedingungen) erhalten wir

$$\xi_0 \leq 2.$$

Das Maximum der Zielfunktion ist somit 2. Durch Rückeinsetzen erhält man sukzessive die Koordinaten des (in diesem Beispiel eindeutig bestimmten) Optimalpunktes $\xi_1^* = 1$, $\xi_2^* = 1$ sowie $\xi_3^* = 0$.

1.3.12 Beispiel. Im zulässigen Bereich aus Beispiel 1.3.11 bestand das Ausgangssystem aus fünf Bedingungen; das durch Elimination von ξ_3 entstandene aber aus sechs Ungleichungen, die sämtlich irredundant sind.



1.4 Abbildung Die Lösungsmengen des Ausgangssystems (links) und des durch Elimination von ξ_3 entstandenen Systems (rechts).

1.3.13 Bezeichnung. u_1, \dots, u_n bezeichnen die **Standardbasisvektoren** des \mathbb{R}^n .

1.3.14 Korollar. Seien P ein Polyeder des \mathbb{R}^n , L ein linearer Teilraum des \mathbb{R}^n und Q entstehe durch orthogonale Projektion von P auf L . Dann ist Q ein Polyeder.

Beweis: Sei $k := \dim(L)$. Die Aussage ist trivial für $k = 0$; sei also im folgenden $k \in \mathbb{N}$. Wir nutzen aus, dass nach Bemerkung 1.3.5 das nach einem Schritt der Fourier-Motzkin Elimination gefundene Ungleichungssystem die orthogonale Projektion der Lösungsmenge auf den entsprechenden Koordinatenraum beschreibt, und dabei jeweils nur endlich viele Ungleichungen auftreten. Der allgemeine Fall ergibt sich durch Anwendung einer geeigneten orthogonalen Transformation. \square

2 Konvexitätstheorie

2.1 Konvexe Mengen

Zur Fixierung der verwendeten Notation beginnen wir mit einigen Standardbezeichnungen der Linearen Algebra endlich-dimensionaler reeller Vektorräume \mathbb{R}^n .

2.1.1 Bezeichnung. (a) Seien $X, Y \subset \mathbb{R}^n$. Dann heißt die durch

$$X + Y := \{x + y : x \in X \wedge y \in Y\}$$

definierte Menge **Minkowski-Summe** [engl.: Minkowski-sum] von X und Y . Ist $\lambda \in \mathbb{R}$ so sei

$$\lambda \cdot X := \{\lambda x : x \in X\}.$$

Ferner wird $0 \cdot \emptyset = \{0\}$ gesetzt. $\lambda \cdot X$ heißt **Dilat** [engl.: dilate] von X um λ .

(b) Seien $v_1, v_2, \dots, v_k \in \mathbb{R}^n$. Dann heißt v **Affinkombination** [engl.: affine combination] der v_1, \dots, v_k , falls $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ existieren mit

$$\sum_{i=1}^k \lambda_i = 1 \quad \wedge \quad v = \sum_{i=1}^k \lambda_i v_i.$$

(c) Sei $T \subset \mathbb{R}^n$. T heißt **affiner Unterraum** von \mathbb{R}^n , wenn T bezüglich Affinkombinationen abgeschlossen ist.

(d) Seien $v_1, \dots, v_k \in \mathbb{R}^n$. Die Vektoren v_1, \dots, v_k heißen **affin unabhängig** [engl.: affinely independent], wenn gilt

$$\lambda_1, \dots, \lambda_k \in \mathbb{R} \quad \wedge \quad \sum_{i=1}^k \lambda_i = 0 \quad \wedge \quad \sum_{i=1}^k \lambda_i v_i = 0 \quad \implies \quad \lambda_1 = \dots = \lambda_k = 0,$$

andernfalls **affin abhängig** [engl.: affinely dependent]. Eine Teilmenge X von \mathbb{R}^n heißt **affin abhängig**, wenn es ein $k \in \mathbb{N}$ und $v_1, \dots, v_k \in X$ gibt, so dass v_1, \dots, v_k affin abhängig sind.

(e) Ist T ein affiner Teilraum von \mathbb{R}^n und gibt es ein $k \in \mathbb{N}_0$, so dass T $k + 1$ aber keine $k + 2$ affin unabhängigen Vektoren enthält, so heißt k die **Dimension** [engl.: dimension] von T und wird mit $\dim(T)$ bezeichnet. Affine Teilräume von \mathbb{R}^n der Dimension 0, 1, 2 bzw. $n - 1$, heißen **Punkt** [engl.: point], **Gerade** [engl.: line], **Ebene** [engl.: plane] bzw. **Hyperebene** [engl.: hyperplane].

- (f) Ist $X \subset \mathbb{R}^n$, so heißt der bez. Inklusion kleinste lineare (affine) Unterraum von \mathbb{R}^n , der X enthält, die **lineare (affine) Hülle** [engl.: linear (affine) hull] von X und wird mit $\text{lin}(X)$ bzw. $\text{aff}(X)$ bezeichnet. Alternativ ergibt sich $\text{lin}(X)$ bzw. $\text{aff}(X)$ als Menge aller Linear- (Affin-) Kombinationen von Elementen aus X (wobei 0 als Linearkombination der Elemente der leeren Menge aufgefasst wird).¹
- (g) Sei $\tau : \mathbb{R}^n \rightarrow \mathbb{R}^k$. τ heißt **affine Abbildung** [engl.: affine mapping], wenn τ mit Affinkombinationen von Elementen aus \mathbb{R}^n verträglich ist, d.h. wenn

$$v_1, \dots, v_k \in \mathbb{R}^n \wedge \lambda_1, \dots, \lambda_k \in \mathbb{R} \wedge \sum_{i=1}^k \lambda_i = 1 \implies \tau\left(\sum_{i=1}^k \lambda_i v_i\right) = \sum_{i=1}^k \lambda_i \tau(v_i).$$

Ist $\tau : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine bijektive affine Abbildung, so heißt τ **affine Transformation** [engl.: affine transformation].²

2.1.2 Definition. Sei $X \subset \mathbb{R}^n$. X heißt **konvex** [engl.: convex] genau dann, wenn

$$x, y \in X \wedge \lambda \in [0, 1] \implies \lambda x + (1 - \lambda)y \in X$$

gilt.

2.1.3 Lemma. Seien \mathcal{C} eine Familie konvexer Mengen des \mathbb{R}^n , $C, C_1, C_2 \in \mathcal{C}$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $P := \{x : Ax \leq b\}$, $\mu \in \mathbb{R}$ sowie $\tau : \mathbb{R}^n \rightarrow \mathbb{R}^k$ eine affine Abbildung. Dann sind die folgenden Mengen konvex:

$$\bigcap_{C \in \mathcal{C}} C, \quad P, \quad C_1 + C_2, \quad \mu \cdot C, \quad \tau(C).$$

2.1.4 Definition. Sei $X \subset \mathbb{R}^n$.

- (a) Die **konvexe Hülle** [engl.: convex hull] $\text{conv}(X)$ von X ist definiert durch

$$\text{conv}(X) := \bigcap \{K : X \subset K \subset \mathbb{R}^n \wedge K \text{ ist konvex}\}.$$

- (b) Seien $x, x_1, \dots, x_k \in \mathbb{R}^n$. Dann heißt x **Konvexkombination** [engl.: convex combination] von x_1, \dots, x_k , falls $\lambda_1, \dots, \lambda_k$ aus \mathbb{R} existieren mit

$$\lambda_1, \dots, \lambda_k \geq 0 \wedge \sum_{i=1}^k \lambda_i = 1 \wedge \sum_{i=1}^k \lambda_i x_i = x.$$

- (c) Sei $P \subset \mathbb{R}^n$. P heißt **\mathcal{V} -Polytop** [engl.: \mathcal{V} -polytope], falls es $k \in \mathbb{N}_0$ und $v_1, \dots, v_k \in \mathbb{R}^n$ gibt mit

$$P = \text{conv}\{v_1, \dots, v_k\}.$$

2.1.5 Lemma. (a) Die konvexe Hülle einer Menge ist konvex.

¹ Eine nicht leere Teilmenge X eines Vektorraums \mathbb{R}^n ist also genau dann affin unabhängig, wenn sich jeder Vektor v aus $\text{aff}(X)$ auf genau eine Weise als Affinkombination von Vektoren aus X darstellen lässt.

² Insbesondere ist also $\tau : \mathbb{R}^n \rightarrow \mathbb{R}^k$ eine affine Abbildung, wenn es eine lineare Abbildung $\psi : \mathbb{R}^n \rightarrow \mathbb{R}^k$ und einen Vektor $w \in \mathbb{R}^k$ gibt, so dass $\tau(v) = w + \psi(v)$ für alle $v \in \mathbb{R}^n$ gilt.

(b) *Konvexkombinationen von Konvexkombinationen sind Konvexkombinationen.*

Beweis: (a) Die Behauptung folgt mit Lemma 2.1.3 (a) unmittelbar aus der Definition.

(b) Da Konvexkombinationen nur endlich viele Terme betreffen, können wir mit Hilfe von Koeffizienten 0 stets erreichen, dass in jeder der betrachteten Konvexkombinationen alle Vektoren auftreten. Seien also $k, l \in \mathbb{N}$ und

$$\begin{aligned} x_j &\in \mathbb{R}^n && (j = 1, \dots, l) \\ \lambda_{i,j} &\geq 0 && (i = 1, \dots, k; j = 1, \dots, l) \quad \text{mit} \\ \sum_{j=1}^l \lambda_{i,j} &= 1 && (i = 1, \dots, k) \\ \mu_i &\geq 0 && (i = 1, \dots, k) \quad \text{mit} \\ \sum_{i=1}^k \mu_i &= 1. \end{aligned}$$

Dann gilt

$$\sum_{i=1}^k \mu_i \left(\sum_{j=1}^l \lambda_{i,j} x_j \right) = \sum_{j=1}^l \left(\sum_{i=1}^k \mu_i \lambda_{i,j} \right) x_j.$$

Mit

$$\eta_j = \sum_{i=1}^k \mu_i \lambda_{i,j} \quad (j = 1, \dots, l)$$

ergibt sich $\eta_j \geq 0$ sowie

$$\sum_{j=1}^l \eta_j = \sum_{i=1}^k \mu_i \left(\sum_{j=1}^l \lambda_{i,j} \right) = \sum_{i=1}^k \mu_i = 1,$$

und es folgt die Behauptung. \square

2.1.6 Lemma. *Sei $X \subset \mathbb{R}^n$. Dann ist $\text{conv}(X)$ die Menge aller Konvexkombinationen von Elementen aus X .*

Beweis: Sei C die Menge aller Konvexkombinationen von Elementen aus X . Dann ist C nach Lemma 2.1.5 konvex. Wegen $X \subset C$ folgt also $\text{conv}(X) \subset C$.

Die Umkehrung ergibt sich mittels vollständiger Induktion über die ‘Länge’ der Konvexkombinationen, d.h. über die Anzahl k der involvierten Punkte von X . Sei C_k die Menge der Konvexkombinationen der Länge k . Dann gilt $C_1 = X$ und $C_2 \subset \text{conv}(X)$. Die Induktionsannahme ist nun, dass $C_k \subset \text{conv}(X)$ gilt für ein $k \in \mathbb{N}$.

Seien $x_1, \dots, x_{k+1} \in X$, $\lambda_1, \dots, \lambda_{k+1} \geq 0$ mit $\sum_{i=1}^{k+1} \lambda_i = 1$, setze $x = \sum_{i=1}^{k+1} \lambda_i x_i$ sowie $\lambda = \sum_{i=1}^k \lambda_i$. O.B.d.A. gelte $\lambda \neq 0$ (da sonst $\lambda_{k+1} = 1$ wäre, die Konvexkombination also nur die Länge 1 hätte). Aus der Induktionsvoraussetzung folgt

$$\sum_{i=1}^k \frac{\lambda_i}{\lambda} x_i \in C_k \subset \text{conv}(X),$$

somit

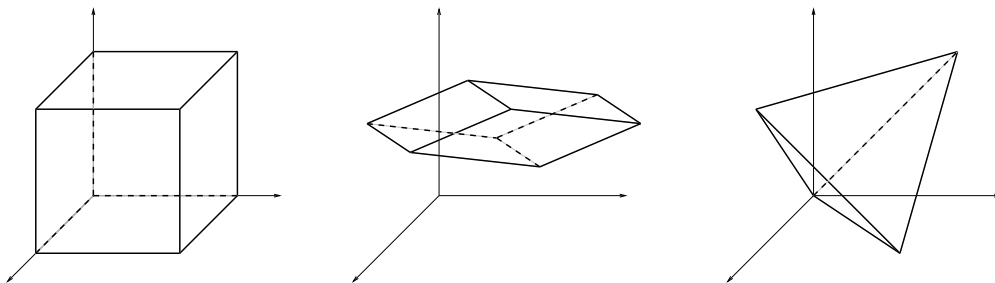
$$x = \sum_{i=1}^{k+1} \lambda_i x_i = \lambda \left(\sum_{i=1}^k \frac{\lambda_i}{\lambda} x_i \right) + (1 - \lambda) x_{k+1} \in \text{conv}(X),$$

und damit die Behauptung. \square

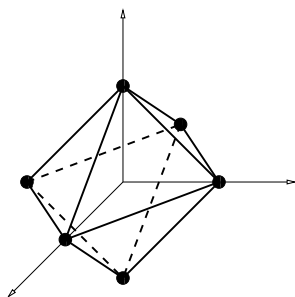
2.1.7 Bezeichnung. (a) $[-1,1]^n$ heißt (n -dimensionaler) **Einheitswürfel**; $[0,1]^n$ heißt **Standardwürfel** [engl.: unit (standard) cube]. Jedes Bild von $[-1,1]^n$ unter einer Ähnlichkeitstransformation (Translation, Dilatation, Drehung) heißt **Würfel** [engl.: cube]. Jedes Bild eines Würfels unter einer affinen Transformation heißt **Parallelotop** [engl.: parallelootope].

(b) Sei $S \subset \mathbb{R}^n$. S heißt (n -dimensionales) **Standardsimplex** [engl.: standard simplex], wenn $S = \text{conv}(\{0, u_1, \dots, u_n\})$ gilt. S heißt **Simplex** [engl.: simplex], wenn es $n+1$ affin unabhängige Punkte s_0, \dots, s_n gibt mit $S = \text{conv}(\{s_0, \dots, s_n\})$.

(c) Sei $Q_n := \text{conv}(\{\pm u_1, \dots, \pm u_n\})$. Dann heißt Q_n (n -dimensionales) **Einheits-Kreuzpolytop** bzw. **Standard-Kreuzpolytop** [engl.: unit (standard) cross-polytope]. Jedes Bild von Q_n unter einer Ähnlichkeitstransformation heißt **reguläres Kreuzpolytop** [engl.: regular cross-polytope], jedes Bild eines Standardkreuzpolytops unter einer affinen Transformation heißt **Kreuzpolytop** [engl.: cross-polytope].



2.1 Abbildung von links: Der Würfel $[0,1]^3$, ein Parallelotop, ein Simplex



2.2 Abbildung Das Oktaeder $Q_3 = \text{conv}(\{\pm u_1, \pm u_2, \pm u_3\})$.

2.1.8 Lemma. Seien $a \in \mathbb{R}^n$, $A = (a_1, \dots, a_n)^T \in \mathbb{R}^{n \times n}$ regulär, s_0, \dots, s_n affin unabhängig, sowie $B := (s_1 - s_0, \dots, s_n - s_0)$. Dann gilt

$$(a) \quad a + A[-1,1]^n = \{x \in \mathbb{R}^n : A^{-1}a - \mathbf{1} \leq A^{-1}x \leq A^{-1}a + \mathbf{1}\} \\ = a + \sum_{i=1}^n [-1,1]a_i = \text{conv} \left(\left\{ a + \sum_{i=1}^n \delta_i a_i : \delta_1, \dots, \delta_n \in \{-1,1\} \right\} \right).$$

$$(b) \operatorname{conv}(\{s_0, \dots, s_n\}) = \{x \in \mathbb{R}^n : B^{-1}x \geq B^{-1}s_0 \wedge \mathbf{1}^T B^{-1}x \leq 1 + \mathbf{1}^T B^{-1}s_0\}.$$

$$(c) a + AQ_n = \{x \in \mathbb{R}^n : y \in \{-1, 1\}^n \Rightarrow y^T A^{-1}x \leq 1 + y^T A^{-1}a\}.$$

2.1.9 Satz. (Satz von Carathéodory)

Sei $X \subset \mathbb{R}^n$. Dann ist jedes Element x aus $\operatorname{conv}(X)$ Konvexkombination von höchstens $n + 1$ Punkten von X .

Beweis: Sei $x \in \operatorname{conv}(X)$. Nach Lemma 2.1.6 ist x Konvexkombination von Punkten von X . Sei $k \in \mathbb{N}$ minimal, so dass es $x_1, \dots, x_k \in X$ und $\lambda_1, \dots, \lambda_k$ aus \mathbb{R} gibt mit

$$\lambda_1, \dots, \lambda_k \geq 0 \quad \wedge \quad \sum_{i=1}^k \lambda_i = 1 \quad \wedge \quad x = \sum_{i=1}^k \lambda_i x_i.$$

Ist $k \leq n + 1$, so ist nichts zu zeigen. Sei also $k \geq n + 2$. Dann sind x_1, \dots, x_k affin abhängig, d.h. es gibt η_1, \dots, η_k aus \mathbb{R} , nicht alle gleich 0, mit $\sum_{i=1}^k \eta_i = 0$, so dass $\sum_{i=1}^k \eta_i x_i = 0$ gilt. Seien O.B.d.A. $l \in \mathbb{N}_0$ mit $l < k$ und $\eta_1, \dots, \eta_l \leq 0, \eta_{l+1}, \dots, \eta_k > 0$, und es gelte

$$\frac{\lambda_i}{\eta_i} \geq \frac{\lambda_k}{\eta_k} \quad (i = l + 1, \dots, k).$$

Dann gilt

$$\begin{aligned} \lambda_k - \frac{\lambda_k}{\eta_k} \eta_k &= 0 \quad \wedge \quad \lambda_i - \frac{\lambda_k}{\eta_k} \eta_i \geq 0 \quad (i = 1, \dots, k - 1) \\ \wedge \quad \sum_{i=1}^{k-1} \left(\lambda_i - \frac{\lambda_k}{\eta_k} \eta_i \right) &= \sum_{i=1}^k \lambda_i - \frac{\lambda_k}{\eta_k} \sum_{i=1}^k \eta_i = 1 \end{aligned}$$

sowie

$$\sum_{i=1}^{k-1} \left(\lambda_i - \frac{\lambda_k}{\eta_k} \eta_i \right) x_i = \sum_{i=1}^k \lambda_i x_i - \frac{\lambda_k}{\eta_k} \sum_{i=1}^k \eta_i x_i = x.$$

Der Punkt x ist also bereits Konvexkombination von $k - 1$ Vektoren aus X , im Widerspruch zur Wahl von k . Hieraus folgt die Behauptung. \square

2.1.10 Bezeichnung. (Analysis)

(a) Die **euklidische Norm** im \mathbb{R}^n wird mit $\|\cdot\|$, die (euklidische) **Einheitskugel** und **Einheitssphäre** werden mit \mathbb{B}^n bzw. \mathbb{S}^{n-1} bezeichnet, d.h.

$$\mathbb{B}^n := \{x \in \mathbb{R}^n : \|x\| \leq 1\} \quad \text{und} \quad \mathbb{S}^{n-1} := \{x \in \mathbb{R}^n : \|x\| = 1\}.$$

Für $\emptyset \neq X, Y \subset \mathbb{R}^n$ sei

$$d(X, Y) := \inf\{\|x - y\| : x \in X \wedge y \in Y\}$$

der **Abstand** von X und Y . Ist X einelementig, etwa $X = \{x\}$, so schreibt man auch kürzer $d(x, Y)$ statt $d(\{x\}, Y)$.

(b) Sei $X \subset \mathbb{R}^n$. Dann bezeichnen $\operatorname{int}(X)$, $\operatorname{cl}(X)$ und $\operatorname{bd}(X)$ das **Innere**, den **Ab-schluss** bzw. den **Rand** von X .

(c) Für $X \subset \mathbb{R}^n$ bezeichne $\text{relint}(X)$ bzw. $\text{relbd}(X)$ das **relative Innere** bzw. den **relativen Rand** von X ; das ist das Innere bzw. der Rand von X bezüglich $\text{aff}(M)$, d.h. bezüglich der in $\text{aff}(X)$ induzierten Topologie (die durch die Schnitte offener euklidischer Kugeln des \mathbb{R}^n mit $\text{aff}(X)$ erzeugt wird).

2.1.11 Lemma. Ist $C \subset \mathbb{R}^n$ nicht leer und konvex, so gilt $\text{relint}(C) \neq \emptyset$.

Beweis: Seien $k := \dim(\text{aff}(C))$, $s_0, \dots, s_k \in C$ affin unabhängig und $S := \text{conv}(\{s_0, \dots, s_k\})$. Ferner seien $Q := (q_1, \dots, q_n)^T \in \mathbb{R}^{n \times n}$ orthogonal mit

$$\text{lin}(\{q_1, \dots, q_k\}) = \text{lin}(\{s_1 - s_0, \dots, s_k - s_0\})$$

und $t_0 := -Qs_0$. Dann gilt $t_0 + Qs_0 = 0$ sowie

$$t_0 + Qs_i = Q(s_i - s_0) \in \text{lin}(\{u_1, \dots, u_k\}) \quad (i = 1, \dots, k).$$

Die durch $\tau(x) := t_0 + Qx$ definierte affine Transformation überführt somit $\text{aff}(S)$ in den Koordinatenraum $\mathbb{R}^k \times \{0\}^{n-k}$. Da Bilder und Urbilder relativ offener Mengen unter τ relativ offen sind, reicht es, den Fall $k = n$ zu betrachten und zu zeigen, dass S innere Punkte besitzt. Seien

$$c := \frac{1}{n+1} \sum_{i=0}^n s_i \quad \wedge \quad B := (s_1 - s_0, \dots, s_n - s_0).$$

nach Lemma 2.1.8 gilt dann

$$S = \{x \in \mathbb{R}^n : B^{-1}x \geq B^{-1}s_0 \wedge \mathbf{1}^T B^{-1}x \leq 1 + \mathbf{1}^T B^{-1}s_0\}.$$

Da B regulär ist, kann keine der $n+1$ Gleichungen

$$\mathbf{1}^T B^{-1}s_i = 1 + \mathbf{1}^T B^{-1}s_0 \quad \wedge \quad u_j^T B^{-1}s_i = u_j^T B^{-1}s_0 \quad (j = 1, \dots, n)$$

für alle $i \in \{1, \dots, n\}$ erfüllt sein. Es folgt

$$B^{-1}c = \frac{1}{n+1} \sum_{i=0}^n B^{-1}s_i > B^{-1}s_0 \quad \wedge \quad \mathbf{1}^T B^{-1}c = \frac{1}{n+1} \sum_{i=0}^n \mathbf{1}^T B^{-1}s_i < 1 + \mathbf{1}^T B^{-1}s_0.$$

Somit gibt es ein $\rho \in]0, \infty[$, so dass für alle $y \in \mathbb{B}^n$ gilt

$$B^{-1}(c + \rho y) > B^{-1}s_0 \quad \wedge \quad \mathbf{1}^T B^{-1}(c + \rho y) < 1 + \mathbf{1}^T B^{-1}s_0,$$

d.h. $c + \rho \mathbb{B}^n \subset S \subset C$. □

2.1.12 Definition. Sei $C \subset \mathbb{R}^n$ nicht leer und konvex. Dann heißt

$$\dim(\text{aff}(C))$$

Dimension von C und wird mit $\dim(C)$ bezeichnet. Ferner wird $\dim(\emptyset) = -1$ gesetzt.

2.1.13 Satz. Die konvexe Hülle kompakter Mengen ist kompakt.

Beweis: Sei $X \subset \mathbb{R}^n$ kompakt. Ferner sei

$$S = \left\{ (\lambda_1, \dots, \lambda_{n+1})^T \in \mathbb{R}^{n+1} : 0 \leq \lambda_1, \dots, \lambda_{n+1} \in \mathbb{R} \wedge \sum_{i=1}^{n+1} \lambda_i = 1 \right\},$$

und die Funktion $\Psi : \mathbb{R}^{n+1} \times \mathbb{R}^{n(n+1)} \rightarrow \mathbb{R}^n$ sei definiert durch

$$\Psi(\lambda_1, \dots, \lambda_{n+1}, x_1, \dots, x_{n+1}) = \sum_{i=1}^{n+1} \lambda_i x_i.$$

Offenbar sind S kompakt und Ψ stetig. Aus Lemma 2.1.6 und Satz 2.1.9 (Caratheodory) folgt

$$\Psi(S \times \underbrace{X \times \dots \times X}_{n+1}) = \text{conv}(X).$$

Nach dem Satz von Tychonoff ist das kartesische Produkt der Kompakta S und X ($n+1$ mal) kompakt. Da das Bild eines Kompaktums unter einer stetigen Abbildung kompakt ist, folgt die Behauptung. \square

2.2 Trennungssätze, Kegel und linear-konvexe Optimierung

2.2.1 Definition. Eine **linear-konvexe Optimierungsaufgabe** [engl.: instance of a linear-convex optimization problem] ist durch folgende Daten spezifiziert:

$$n \in \mathbb{N}, \quad C \subset \mathbb{R}^n \text{ abgeschlossen, konvex,} \quad \varphi : \mathbb{R}^n \rightarrow \mathbb{R} \text{ linear,} \quad \text{opt} \in \{\min, \max\}.$$

Die Aufgabe besteht darin, φ über C zu optimieren. Für $\text{opt}=\max$ spricht man auch von einer **linear-konvexen Maximierungsaufgabe** [engl.: instance of a linear-convex maximization problem], für $\text{opt}=\min$ von einer **linear-konvexen Minimierungsaufgabe** [engl.: instance of a linear-convex minimization problem]. Die Menge aller linear-konvexen Optimierungsaufgaben heißt **linear-konvexes Optimierungsproblem** [engl.: linear-convex optimization problem] oder kürzer **linear-konvexes Programm** [engl.: linear-convex problem].

2.2.2 Bezeichnung. Seien H eine Hyperebene des \mathbb{R}^n , $a \in \mathbb{R}^n \setminus \{0\}$ und $\beta \in \mathbb{R}$, so dass

$$H = \{x : a^T x = \beta\}.$$

Dann wird die Notation $H_{(a,\beta)}$ für H verwendet und a als **Normalenvektor** oder **Normale** [engl.: normal] von H bezeichnet.

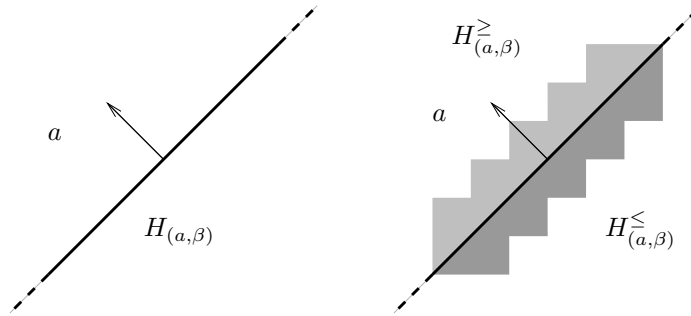
2.2.3 Definition. (a) Eine Teilmenge H^\leq des \mathbb{R}^n heißt (abgeschlossener) **Halbraum** [engl.: (closed) halfspace] des \mathbb{R}^n , wenn es $a \in \mathbb{R}^n \setminus \{0\}$ und $\beta \in \mathbb{R}$ gibt mit $H^\leq = \{x : a^T x \leq \beta\}$. Entsprechend heißt a **äußerer Normalenvektor** oder **äußere Normale** [engl.: outer normal] von H^\leq .

(b) Ist $H = H_{(a,\beta)}$ eine Hyperebene, so heißen $H_{(a,\beta)}^\leq = \{x : a^T x \leq \beta\}$ bzw. $H_{(a,\beta)}^\geq = \{x : a^T x \geq \beta\}$ (abgeschlossene) **Halbräume** [engl.: positive (negative) halfspace] zur Hyperebene H ; $H_{(a,\beta)}^{<} = \{x : a^T x < \beta\}$ bzw. $H_{(a,\beta)}^{>} = \{x : a^T x > \beta\}$ (offene) **Halbräume**. Bisweilen werden abgeschlossene Halbräume auch einfacher mit H^+ und H^- bezeichnet.

(c) Seien $S \in \mathbb{R}^n$ und $k \in \{1, \dots, n\}$. S heißt (relativ offener bzw. abgeschlossener) k -dimensionaler **Halbunterraum**, wenn es einen k -dimensionalen affinen Teilraum F und einen (offenen bzw. abgeschlossenen) Halbraum H^+ des \mathbb{R}^n gibt mit

$$\emptyset \neq F \cap H^+ \neq F \quad \wedge \quad S = F \cap H^+.$$

Ein 1-dimensionaler Halbunterraum wird auch **Halbgerade** oder **Strahl** [engl.: half-line, ray] genannt.



2.3 Abbildung Hyperebene und zugehörige Halbräume im \mathbb{R}^2 .

2.2.4 Definition. Seien $X, Y \subset \mathbb{R}^n$, H eine Hyperebene. Dann **trennt** [engl.: separates] H die Mengen X und Y genau dann, wenn es $a \in \mathbb{R}^n \setminus \{0\}$ und $\beta \in \mathbb{R}$ gibt mit $H = H_{(a,\beta)}$ und

$$X \subset H_{(a,\beta)}^{\leq} \quad \wedge \quad Y \subset H_{(a,\beta)}^{\geq}.$$

Die Hyperebene H **trennt** X und Y **strikt** [engl.: strictly separates], genau dann, wenn es $a \in \mathbb{R}^n \setminus \{0\}$ und $\beta \in \mathbb{R}$ gibt mit

$$X \subset H_{(a,\beta)}^{<} \quad \wedge \quad Y \subset H_{(a,\beta)}^{>}.$$

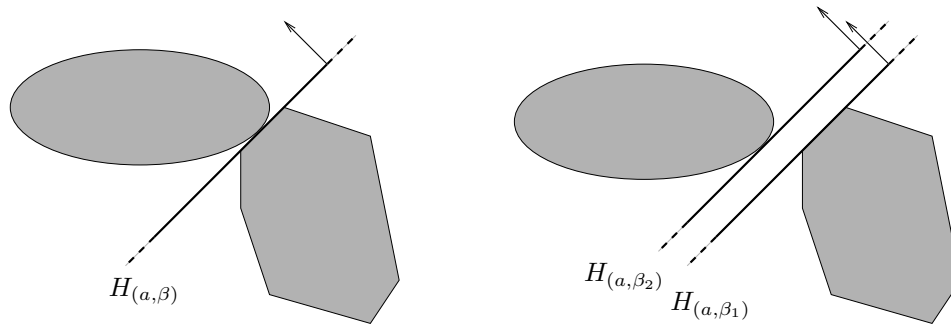
H **trennt** X und Y **streng** [engl.: strongly separates], genau dann, wenn es $a \in \mathbb{R}^n \setminus \{0\}$ und $\beta, \beta_1, \beta_2 \in \mathbb{R}$ gibt mit $\beta_1 < \beta < \beta_2$, $H = H_{(a,\beta)}$ und

$$X \subset H_{(a,\beta_1)}^{\leq} \quad \wedge \quad Y \subset H_{(a,\beta_2)}^{\geq}.$$

Trennt die Hyperebene H die Mengen X und Y (strikt, streng), so wird H als X und Y (**strikt, streng**) **trennende Hyperebene** [engl.: (strictly, strongly) separating hyperplane] bezeichnet.

2.2.5 Satz. Seien $K, C \subset \mathbb{R}^n$ konvex, K beschränkt, und es gelte $\text{cl}(K) \cap \text{cl}(C) = \emptyset$. Dann gibt es eine K und C streng trennende Hyperebene.

Beweis: Wir können o.B.d.A. annehmen, dass K und C nicht leer und abgeschlossen sind. Sei $x \in K$. Da man sich zur Berechnung von $d(x, C)$ auf ein Kompaktum in C beschränken kann, gibt es stets ein y aus C mit $d(x, C) = \|x - y\|$. Da C konvex ist (und \mathbb{S}^{n-1} keine Strecke positiver Länge enthält), ist y eindeutig bestimmt, d.h. y kann als Funktion $y(x)$ von x aufgefasst werden. Als solche ist $d(x, C)$ stetig; also gibt es ein x_0 , so dass mit $y_0 = y(x_0)$ gilt



2.4 Abbildung Trennung (links), strenge Trennung (rechts)

$$\|x_0 - y_0\| = \min\{d(x, C) : x \in K\}.$$

Seien nun

$$a := y_0 - x_0 \quad \wedge \quad \beta_1 := a^T x_0 \quad \wedge \quad \beta_2 := a^T y_0.$$

Dann gilt nach Voraussetzung $0 < \|y_0 - x_0\|^2 = (y_0 - x_0)^T (y_0 - x_0)$ und somit

$$\beta_1 = a^T x_0 < a^T y_0 = \beta_2.$$

Mittels Widerspruchsbeweises zeigen wir $K \subset H_{(a, \beta_1)}^{\leq}$. Der Beweis für $C \subset H_{(a, \beta_2)}^{\geq}$ folgt analog.³ Sei also $x \in K \cap H_{(a, \beta_1)}^{\geq}$. Für $\lambda \in]0, 1[$ gilt

$$\begin{aligned} \left\| y_0 - (\lambda x + (1 - \lambda)x_0) \right\|^2 &= (\lambda(y_0 - x) + (1 - \lambda)a)^T (\lambda(y_0 - x) + (1 - \lambda)a) \\ &= \lambda^2 ((y_0 - x) - a)^T ((y_0 - x) - a) + 2\lambda ((y_0 - x) - a)^T a + a^T a \\ &= \lambda^2 \|x_0 - x\|^2 + 2\lambda (x_0 - x)^T a + \|a\|^2. \end{aligned}$$

Da nach Voraussetzung $a^T (x_0 - x) < 0$ ist, gibt es ein $\delta \in]0, 1[$, so dass für alle $\lambda \in]0, \delta[$

$$\lambda \|x_0 - x\|^2 + 2(x_0 - x)^T a < 0$$

gilt. Somit folgt für jedes solche λ

$$\left\| y_0 - (\lambda x + (1 - \lambda)x_0) \right\| < \|y_0 - x_0\|,$$

im Widerspruch zur Wahl von x_0 und y_0 . □

2.2.6 Bezeichnung. Die im Beweis von Satz 2.2.5 konstruierte Abbildung $x \mapsto y(x)$ heißt **nearest-point-map** von C , und wird mit π_C bezeichnet.

2.2.7 Beispiel. Seien $a_1, \dots, a_k, b \in \mathbb{R}^n$, $A = (a_1, \dots, a_k)$ und $Q = \text{conv}\{a_1, \dots, a_k\}$. Genau dann ist $b \notin Q$, wenn sich b nicht als Konvexkombination der Vektoren a_1, \dots, a_k darstellen lässt, d.h.

³ Alternativ könnte man rein elementargeometrisch argumentieren.

$$\neg \exists x = (\xi_1, \dots, \xi_k)^T \in [0, \infty]^k: \sum_{i=1}^k \xi_i = 1 \wedge Ax = b.$$

Nach Satz 2.2.5 ist hingegen $b \notin Q$ genau dann, wenn es eine b und Q streng trennende Hyperebene gibt, d.h. wenn

$$\exists (y \in \mathbb{R}^n \wedge \eta \in \mathbb{R}) : y^T b < \eta \wedge y^T A > \eta \mathbf{1}.$$

Somit ist

$$\begin{aligned} Ax &= b \\ \mathbf{1}^T x &= 1 \\ x &\geq 0 \end{aligned}$$

genau dann unzulässig, wenn

$$\begin{aligned} y^T b - \eta &< 0 \\ -y^T A + \eta \mathbf{1} &< 0 \end{aligned}$$

zulässig ist. Der Trennungssatz hat hier also die Gestalt eines **Alternativsatzes**. Es gibt viele Varianten solcher Sätze, die im Bereich der linearen Optimierung in der Regel unter dem Namen **Lemma von Farkas** firmieren. Satz 2.2.29 enthält eine Optimierungsversion.

2.2.8 Satz. Seien $K, C \subset \mathbb{R}^n$ nicht leer und konvex. Ist $\text{relint}(K) \cap \text{relint}(C) = \emptyset$, so existiert eine K und C trennende Hyperebene.

Existiert umgekehrt eine K und C trennende Hyperebene, und ist $\text{aff}(K \cup C) = \mathbb{R}^n$, so gilt $\text{relint}(K) \cap \text{relint}(C) = \emptyset$.

2.2.9 Korollar. Seien $K, C \subset \mathbb{R}^n$ nicht leer und konvex, $S := \text{aff}(K \cup C)$, $s \in S$, und es gelte $\text{relint}(K) \cap \text{relint}(C) = \emptyset$. Dann existieren $a \in (-s + S) \setminus \{0\}$ und $\beta \in \mathbb{R}$, so dass $H_{(a, \beta)}$ die Mengen K und C trennt.

2.2.10 Satz. Jede abgeschlossene konvexe Menge des \mathbb{R}^n ist Durchschnitt aller abgeschlossenen (oder aller offenen) Halbräume, die sie enthalten.

2.2.11 Definition. Seien $X \subset \mathbb{R}^n$, $x \in \text{cl}(X)$ und H eine Hyperebene. Dann heißt H **Stützhyperebene** [engl.: supporting hyperplane] **an X im Punkt x** , wenn H die Mengen X und $\{x\}$ trennt. Man sagt dann auch ‘ H stützt X in x ’ und bezeichnet x als **Stützpunkt** [engl.: support point] von X . Eine Hyperebene H heißt **Stützhyperebene an X** , wenn ein Punkt $x \in \text{cl}(X)$ existiert, so dass H die Menge X im Punkt x stützt. Eine Stützhyperebene an X heißt **eigentlich**, wenn $X \not\subset H$ ist; andernfalls **uneigentlich**. Ist H eine Stützhyperebene an X und H^+ der von H berandete abgeschlossene Halbraum, der X enthält, dann heißt H^+ zu H gehöriger **Stützhalbraum** [engl.: supporting halfspace].

2.2.12 Bemerkung. Seien $X \subset \mathbb{R}^n$ beschränkt, H eine Hyperebene und H^+ , H^- die beiden zugehörigen abgeschlossenen Halbräume. H stützt X genau dann, wenn gilt

$$d(X, H) = 0 \quad \wedge \quad (x, y \in X \wedge x \in \text{int}(H^+) \Rightarrow y \in H^+).$$

2.2.13 Satz. Seien $C \subset \mathbb{R}^n$ und $K \subset \text{relbd}(C)$ beide konvex und $K \neq \emptyset$. Dann gibt es eine eigentliche Stützhyperebene H an C mit $K \subset H$.

Beweis: Sei $s \in C$. Nach Korollar 2.2.9 gibt es eine C und K trennende Hyperebene $H := H_{(a,\beta)}$ mit $a + s \in \text{aff}(C)$, und es folgt $C \not\subset H$. Seien H^\pm die zugehörigen abgeschlossenen Halbräume mit $C \subset H^+$ und $K \subset H^-$. Aus $K \subset C \subset H^+$ folgt $K \subset H^+ \cap H^- = H$ und damit die Behauptung. \square

2.2.14 Korollar. Seien $C \subset \mathbb{R}^n$ abgeschlossen und konvex sowie $x^* \in C$.

- (a) Seien $c \in \mathbb{R}^n \setminus \{0\}$ und $\beta \in \mathbb{R}$. Der Vektor x^* ist genau dann ein Optimalpunkt der linear-konvexen Maximierungsaufgabe $\max_{x \in C} c^T x$, wenn $H_{(c,\beta)}^{\leq}$ Stützhalbraum an C in x^* ist. In diesem Fall gilt $\beta = \max_{x \in C} c^T x$.
- (b) Ist $x^* \in \text{bd}(C)$, so existiert ein Vektor $c \in \mathbb{R}^n \setminus \{0\}$, so dass x^* Optimalpunkt der linear-konvexen Maximierungsaufgabe $\max_{x \in C} c^T x$ ist.⁴ Ist $x^* \in \text{relbd}(C)$, so kann noch verlangt werden, dass die durch $x \mapsto c^T x$ definierte Zielfunktion auf C nicht konstant ist.

2.2.15 Bemerkung. Seien $X \subset \mathbb{R}^n$, $x^* \in X$ und

$$K := \{c \in \mathbb{R}^n : \max_{x \in X} c^T x = c^T x^*\}.$$

Dann gilt:

$$\lambda_1, \lambda_2 \in [0, \infty[\wedge c_1, c_2 \in K \Rightarrow \lambda_1 c_1 + \lambda_2 c_2 \in K.$$

2.2.16 Definition. (a) Sei $K \subset \mathbb{R}^n$. K heißt **Kegel** [engl.: cone], wenn gilt

$$[0, \infty[K \subset K \wedge K + K \subset K.$$

- (b) Ist K ein Kegel und ein Polyeder, so wird K auch **polyederischer Kegel** [engl.: polyhedral cone] genannt.
- (c) Sei $X \subset \mathbb{R}^n$. Die Menge

$$\bigcap \{K : X \subset K \wedge K \text{ ist Kegel}\}$$

heißt **positive** (oder **konische**) **Hülle** [engl.: positive hull] von X und wird mit $\text{pos}(X)$ bezeichnet.

- (d) Seien $x, x_1, \dots, x_k \in \mathbb{R}^n$. Dann heißt x **Nichtnegativkombination** oder **Konische Kombination** [engl.: conic combination] von x_1, \dots, x_k , falls $\lambda_1, \dots, \lambda_k$ aus \mathbb{R} existieren mit

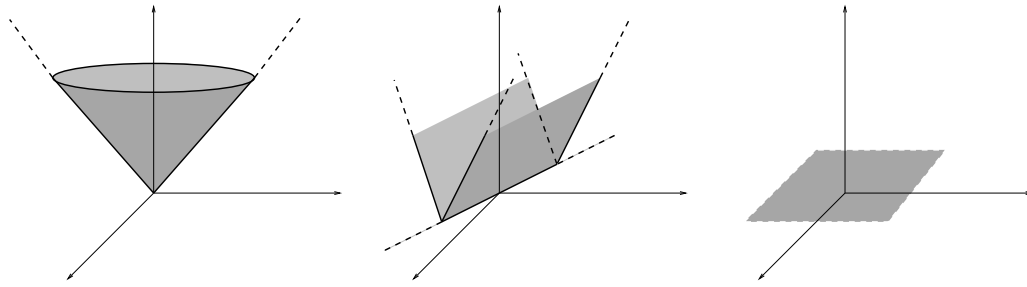
$$\lambda_1, \dots, \lambda_k \geq 0 \wedge \sum_{i=1}^k \lambda_i x_i = x.$$

2.2.17 Beispiel. Seien

$$A \subset \mathbb{R}^n \setminus \{0\} \wedge K := \bigcap_{a \in A} H_{(a,0)}^{\leq}.$$

Für $a \in A$, $x, y \in K$ und $\lambda \in [0, \infty[$ gilt

$$a^T(\lambda x) = \lambda a^T x \leq 0 \wedge a^T(x + y) = a^T x + a^T y \leq 0.$$



2.5 Abbildung Verschiedene Kegel des \mathbb{R}^3 .

Also ist K ein Kegel.

Die drei in Abbildung 2.5 skizzierten Kegel geben verschiedene Beispiele an. Links ist A überabzählbar, in der Mitte und rechts kommt man mit nur zwei Elementen aus.

2.2.18 Bemerkung. (a) Sei K ein Kegel im \mathbb{R}^n . Dann ist K konvex, und es gilt $0 \in K$.

(b) Sei $X \subset \mathbb{R}^n$. Dann ist $\text{pos}(X)$ ein Kegel.

2.2.19 Bemerkung. Seien K ein abgeschlossener Kegel im \mathbb{R}^n und A die Menge der äußeren Normalen aller Stützhyperebenen an K in 0 . Dann gilt $K = \bigcap_{a \in A} H_{(a,0)}^{\leq}$.

Beweis: Nach Satz 2.2.10 ist K Durchschnitt aller abgeschlossenen Halbräume, die K enthalten. Seien nun $a \in \mathbb{R}^n \setminus \{0\}$ und $\beta \in \mathbb{R}$ mit $K \subset H_{(a,\beta)}^{\leq}$. Nach Bemerkung 2.2.18 ist $0 \in K$, und es folgt $\beta \geq 0$. Sei nun $x \in K$. Für jedes $\lambda \in [0, \infty[$ gilt dann $a^T(\lambda x) = \lambda a^T x \leq \beta$, und somit $a^T x \leq 0$. Es folgt $K \subset H_{(a,0)}^{\leq}$ und damit die Behauptung. \square

2.2.20 Lemma. Sei $X \subset \mathbb{R}^n$. Dann gilt

$$\text{pos}(X) = [0, \infty[\text{conv}(X) = \text{conv}([0, \infty[X),$$

d.h. $\text{pos}(X)$ ist die Menge aller Nichtnegativkombinationen von Elementen von X .

2.2.21 Bezeichnung. Seien $C \subset \mathbb{R}^n$ konvex und $x^* \in C$. Ferner seien

$$N_C(x^*) := \left\{ c \in \mathbb{R}^n : \max_{x \in C} c^T x = c^T x^* \right\}$$

sowie

$$S_C(x^*) := \bigcap_{c \in N_C(x^*)} H_{(c,0)}^{\leq}.$$

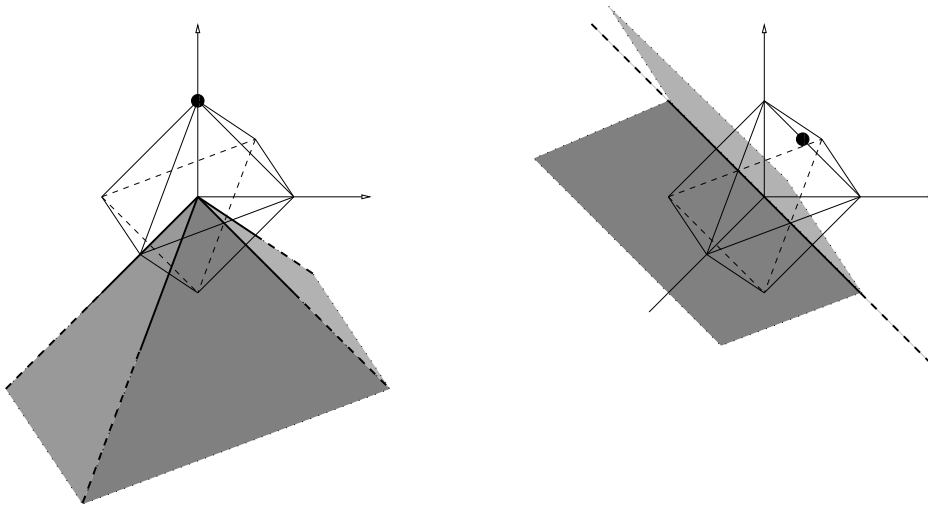
Dann heißen $N_C(x^*)$ **Kegel der äußeren Normalen** [engl.: normal cone] und $S_C(x^*)$ **Stützkegel** [engl.: support cone] an C in x^* . Bisweilen wird $S_C(x^*)$ auch **Tangentialekegel** [engl.: tangential cone] oder (für Polyeder) **Innenkegel** [engl.: inner cone] genannt.

⁴ Korollar 2.2.14 (b) lässt sich in der Sprache der Aktivitätsanalyse linear-konvexer ökonomischer Modelle so formulieren, dass zu jedem effizienten Produktionsprogramm x^* ein Effizienzpreissystem $c \in \mathbb{R}^n \setminus \{0\}$ existiert. Wir werden auf dieses dort 'Koopmans Effizienzpreissystem' genannte Ergebnis in Kapitel ?? noch etwas genauer eingehen.

2.2.22 Bemerkung. Seien $C \subset \mathbb{R}^n$ konvex, $c \in \mathbb{R}^n$ sowie $x^* \in C$. Dann ist x^* genau dann ein Optimalpunkt der linear-konvexen Maximierungsaufgabe $\max_{x \in C} c^T x$, wenn $c \in N_C(x^*)$ gilt.

2.2.23 Bemerkung. Seien $C \subset \mathbb{R}^n$ konvex und $x^* \in C$. Dann sind $N_C(x^*)$ und $S_C(x^*)$ abgeschlossene Kegel.

Beweis: Dass $N_C(x^*)$ und $S_C(x^*)$ tatsächlich Kegel sind, wurde bereits in Bemerkung 2.2.15 bzw. Beispiel 2.2.17 gezeigt. $S_C(x^*)$ ist als Durchschnitt abgeschlossener Halbräume abgeschlossen. Seien nun $(c_i)_{i \in \mathbb{N}}$ eine Folge in $N_C(x^*)$ und $c \in \mathbb{R}^n$ mit $c_i \rightarrow c$ für $i \rightarrow \infty$. Für $x \in C$ und $i \in \mathbb{N}$ gilt $c_i^T x \leq c_i^T x^*$, und aus der Stetigkeit linearer Funktionen folgt dann auch $c^T x \leq c^T x^*$. \square



2.6 Abbildung Stützkegel $S_P(x^*)$ an zwei verschiedenen Randpunkten x^* des Standardkreuzpolytops.

2.2.24 Lemma. Seien $C \subset \mathbb{R}^n$ konvex und $x^* \in C$. Dann gilt

$$S_C(x^*) = \text{cl}([0, \infty[(C - x^*)).$$

Beweis: Sei $Q := \text{cl}([0, \infty[(C - x^*))$. Da $S_C(x^*)$ ein abgeschlossener Kegel ist, der $C - x^*$ enthält, gilt $Q \subset S_C(x^*)$.

Sei nun $y \in \text{relint}(S_C(x^*))$, und es gelte

$$]0, \infty[y \cap (C - x^*) = \emptyset.$$

Nach Korollar 2.2.9 gibt es $a \in \text{lin}(S_C(x^*)) \setminus \{0\}$ und $\beta \in \mathbb{R}$, so dass

$$]0, \infty[y \subset H_{(a, \beta)}^{\geq} \quad \wedge \quad C - x^* \subset H_{(a, \beta)}^{\leq}.$$

Wegen $0 \in]0, \infty[y \cap (C - x^*)$ gilt $\beta = 0$, d.h. $H_{(a, a^T x^*)}^{\leq}$ ist ein Stützhalbraum an C in x^* . Nach Definition von $S_C(x^*)$ gilt somit $y \notin S_C(x^*)$, im Widerspruch zur Wahl von y . Somit ist $\text{relint}(S_C(x^*)) \subset Q$, und da Q abgeschlossen ist, folgt die Behauptung. \square

2.2.25 Lemma. Seien C konvex, $c \in \mathbb{R}^n$ und $x^* \in C$. Der Punkt x^* ist genau dann Optimalpunkt der linear-konvexen Optimierungsaufgabe

$$\begin{aligned} \max c^T x \\ x \in C, \end{aligned}$$

wenn er Optimalpunkt von

$$\begin{aligned} \max c^T x \\ x \in x^* + S_C(x^*) \end{aligned}$$

ist.

Beweis: Wegen $x^* \in C \subset x^* + S_C(x^*)$ gilt natürlich

$$\max_{x \in C} c^T x \leq \max_{x \in x^* + S_C(x^*)} c^T x.$$

Ist also x^* Maximalpunkt bez. $x^* + S_C(x^*)$, so ist x^* erst recht Maximalpunkt bez. C .

Seien nun x^* Maximalpunkt bez. C , aber $y \in x^* + S_C(x^*)$ mit $c^T x^* < c^T y$. Nach Lemma 2.2.24 gilt

$$y - x^* \in \text{cl}([0, \infty[(C - x^*)).$$

Seien entsprechend $(\lambda_i)_{i \in \mathbb{N}}$ und $(x_i)_{i \in \mathbb{N}}$ Folgen mit $\lambda_i \geq 0$ und $x_i \in C - x^*$ für alle $i \in \mathbb{N}$, so dass

$$x^* + \lambda_i x_i \rightarrow y \quad (i \rightarrow \infty).$$

Dann gilt

$$c^T(x^* + \lambda_i x_i) \rightarrow c^T y > c^T x^*,$$

also gibt es ein $i_0 \in \mathbb{N}$ mit $c^T(x^* + \lambda_{i_0} x_{i_0}) > c^T x^*$. Es folgt $c^T x_{i_0} > 0$, also

$$x^* + x_{i_0} \in C \quad \wedge \quad c^T(x^* + x_{i_0}) > c^T x^*,$$

im Widerspruch zur Optimalität von x^* . □

2.2.26 Korollar. Seien C konvex, $c \in \mathbb{R}^n$, $x^* \in C$ und $S := S_C(x^*)$. Dann gilt

$$N_C(x^*) = N_S(0).$$

2.2.27 Bezeichnung. Seien $n, m \in \mathbb{N}$, $M := \{1, \dots, m\}$, $a_1, \dots, a_m \in \mathbb{R}^n \setminus \{0\}$, $A := (a_1, \dots, a_m)^T$, $b := (\beta_1, \dots, \beta_m)^T \in \mathbb{R}^m$, $P := \{x \in \mathbb{R}^n : Ax \leq b\}$ und $x \in P$. Sei $I(x) \subset M$, so dass

$$\begin{aligned} a_i^T x &= \beta_i & (i \in I(x)) \\ a_i^T x &< \beta_i & (i \in M \setminus I(x)). \end{aligned}$$

Dann heißt genau jede Nebenbedingung $a_i^T x \leq \beta_i$ mit $i \in I(x)$ in x **aktiv** [engl.: active constraints]. $I(x)$ wird Menge der Indizes aller aktiven Nebenbedingungen genannt. Ferner sei

$$S^=(x) := \{y \in \mathbb{R}^n : i \in I(x) \Rightarrow a_i^T y = \beta_i\}.$$

2.2.28 Lemma. Seien $n, m \in \mathbb{N}$, $M := \{1, \dots, m\}$, $a_1, \dots, a_m \in \mathbb{R}^n \setminus \{0\}$, $A := (a_1, \dots, a_m)^T$, $b := (\beta_1, \dots, \beta_m)^T \in \mathbb{R}^m$, $P := \{x \in \mathbb{R}^n : Ax \leq b\}$ und $x^* \in P$. Dann gilt

$$N_P(x^*) = \text{pos}(\{a_i : i \in I(x^*)\}) \quad \wedge \quad S_P(x^*) = \bigcap_{i \in I(x^*)} \{x \in \mathbb{R}^n : a_i^T x \leq 0\}.$$

Beweis: Sei $Q := \text{pos}(\{a_i : i \in I(x^*)\})$. Für $i \in I(x^*)$ gilt

$$\max_{x \in P} a_i^T x \leq \beta_i = a_i^T x^*,$$

d.h. $Q \subset N_P(x^*)$.

Sei nun $c \in N_P(x^*) \setminus Q$. Nach Satz 2.2.5 gibt es eine c und Q streng trennende Hyperebene. Seien also $q \in \mathbb{R}^n \setminus \{0\}$ und $\kappa_1, \kappa_2 \in \mathbb{R}$ mit $\kappa_1 < \kappa_2$ und

$$Q \subset H_{(q, \kappa_1)}^{\leq} \quad \wedge \quad c^T q \geq \kappa_2.$$

Wegen $0 \in Q$ gilt $0 \leq \kappa_1 < \kappa_2$. Nach Lemma 2.2.20 ist die linke Inklusion äquivalent zu

$$\lambda_1, \dots, \lambda_m \in [0, \infty[\implies \left(\sum_{i \in I(x^*)} \lambda_i a_i \right)^T q = \sum_{i \in I(x^*)} \lambda_i a_i^T q \leq \kappa_1,$$

und das gilt genau dann, wenn

$$a_1^T q, \dots, a_m^T q \leq 0$$

sind. Es folgt

$$i \in I(x^*) \wedge \mu \in [0, \infty[\implies a_i^T(x^* + \mu q) = \beta_i + \mu a_i^T q \leq \beta_i.$$

Somit gibt es ein $\mu_0 \in]0, \infty[$, so dass

$$x^* + \mu_0 q \in P \quad \wedge \quad c^T(x^* + \mu_0 q) = c^T x^* + \mu_0 c^T q \geq c^T x^* + \mu_0 \kappa_2 > c^T x^*,$$

im Widerspruch zu $c \in N_P(x^*)$. Somit folgt auch $N_P(x^*) \subset Q$, und damit die erste Identität.

Nach Definition gilt dann auch

$$S_P(x^*) = \bigcap_{c \in \text{pos}(\{a_i : i \in I(x^*)\})} H_{(c, 0)}^{\leq} \subset \bigcap_{i \in I(x^*)} \{x \in \mathbb{R}^n : a_i^T x \leq 0\}.$$

Für $i \in I(x^*)$ seien nun $\lambda_i \in [0, \infty[$ und $a := \sum_{i \in I(x^*)} \lambda_i a_i$. Ferner sei $x \in \mathbb{R}^n$ mit $a_i^T x \leq 0$ für alle $i \in I(x^*)$. Dann gilt

$$a^T x = \left(\sum_{i \in I(x^*)} \lambda_i a_i \right)^T x = \sum_{i \in I(x^*)} \lambda_i a_i^T x \leq 0.$$

Hiermit folgt auch die zweite Behauptung. \square

2.2.29 Korollar. (Lemma von Farkas)

Seien $n, m \in \mathbb{N}$, $a_1, \dots, a_m \in \mathbb{R}^n$, $A := (a_1, \dots, a_m)^T \in \mathbb{R}^{m \times n}$ und $c \in \mathbb{R}^n$. Dann gilt

$$(Ax \leq 0 \implies c^T x \leq 0) \iff (\exists y \in \mathbb{R}^m : y \geq 0 \wedge A^T y = c).$$

2.3 Darstellungssätze

2.3.1 Bemerkung. Seien $C \subset \mathbb{R}^n$ abgeschlossen und konvex und $s \in \mathbb{R}^n \setminus \{0\}$. Dann gilt

$$x, y \in C \wedge x + [0, \infty[s \subset C \implies y + [0, \infty[s \subset C.$$

2.3.2 Satz. Sei $C \subset \mathbb{R}^n$ abgeschlossen und konvex. Dann gilt

$$C = \text{conv}(\text{relbd}(C))$$

genau dann, wenn C kein nicht-leerer affiner Unterraum oder Halbunterraum des \mathbb{R}^n ist.

Beweis: Ist C ein affiner Unterraum, so gilt $\text{relbd}(C) = \emptyset$. Ist C hingegen ein Halbunterraum, und H^+ ein zugehöriger Halbraum, so gilt $\text{relbd}(C) = \text{aff}(C) \cap \text{bd}(H^+) \neq C$.

Sei nun $K := \text{conv}(\text{relbd}(C))$. Da C abgeschlossen ist, gilt $\text{relbd}(C) \subset C$; aus der Konvexität folgt daher $K \subset C$. Wir nehmen an, dass die umgekehrte Inklusion nicht gelte. Da $\text{relbd}(C) \subset K$ ist, sei

$$y \in \text{relint}(C) \setminus K.$$

Nach Korollar 2.2.9 gibt es eine $\{y\}$ und K trennende Hyperebene mit Normale in $(-y) + \text{aff}(C)$. Sei also $H_{(a,\beta)}^<$ ein abgeschlossener Halbraum mit

$$y \in H_{(a,\beta)}^< \quad \wedge \quad K \subset H_{(a,\beta)}^>.$$

Ist $z \in H_{(a,\beta)}^< \cap \text{aff}(C)$, so folgt

$$\text{conv}(\{y, z\}) \cap \text{relbd}(C) \subset \text{conv}(\{y, z\}) \cap K = \emptyset,$$

d.h. $z \in \text{relint}(C)$. Somit gilt

$$H_{(a,\beta)}^< \cap \text{aff}(C) \subset C.$$

Da C abgeschlossen und konvex ist, gilt nach Bemerkung 2.3.1 für jeden Punkt $c \in \text{relbd}(C)$

$$c + (H_{(a,0)}^< \cap \text{aff}(C)) \subset C.$$

Somit ist C ein (nicht-leerer) affiner Unterraum oder Halbunterraum des \mathbb{R}^n . \square

2.3.3 Definition. Seien $C \subset \mathbb{R}^n$ konvex, $x \in C$ und

$$F(x) := \bigcup \left\{ \text{conv}(\{y, z\}) : y, z \in C \wedge x \in \text{relint}(\text{conv}(\{y, z\})) \right\}.$$

Dann heißt $F(x)$ **Seite** von x in C [engl.: face]. Ist $k \in \mathbb{N}_0$ und $\dim(\text{aff}(F(x))) = k$, so wird $F(x)$ auch **k -Seite** von C genannt. Zur besonderen Kennzeichnung von C wird bisweilen auch $F_C(x)$ geschrieben.

Gilt $F(x) = \{x\}$, so heißt x **Extremalpunkt** [engl.: extreme point] von C . Die Menge der Extremalpunkte von C wird mit $\text{ext}(C)$ bezeichnet.

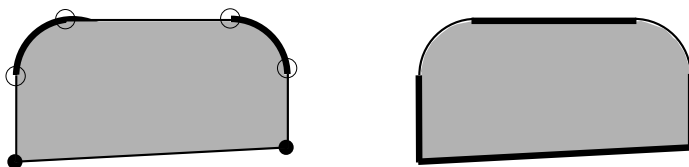
Ist $F(x)$ ein Strahl, so heißt $F(x)$ **Extremalstrahl** [engl.: extreme ray] von C .

Eine Menge $F \subset C$ heißt **Seite** von C , wenn $F = \emptyset$ ist, oder wenn es einen Punkt $x \in C$ gibt mit $F = F(x)$. \emptyset und C werden auch **uneigentliche Seiten** von C genannt, alle übrigen **eigentliche Seiten** von C .

2.3.4 Beispiel. Wir bestimmen die eigentlichen Seiten von \mathbb{B}^n . Seien $x \in \mathbb{S}^{n-1}$ und $y, z \in \mathbb{B}^n$, $\lambda \in]0, 1[$ mit $x = \lambda y + (1 - \lambda)z$. Nach der Cauchy-Schwarzschen Ungleichung gilt

$$1 = x^T x = x^T (\lambda y + (1 - \lambda)z) \leq \lambda \|y\| + (1 - \lambda) \|z\| \leq 1$$

mit Gleichheit genau dann, wenn $y, z \in \mathbb{S}^{n-1} \cap \mathbb{R}x$ gilt. Es folgt $y = z = x$. Jeder Randpunkt von \mathbb{B}^n ist also Extrempunkt; \mathbb{B}^n besitzt daher nur k -Seiten für $k \in \{-1, 0, n\}$.



2.7 Abbildung Fett hervorgehoben sind (links) die Menge der Extrempunkte (zwei Kreisbögen und zwei Eckpunkte) und (rechts) die vier 1-dimensionalen Seiten der abgebildeten 2-dimensionalen Menge C . Die vier eingekreisten Extrempunkte (links) sind nicht Durchschnitt von C mit einer Stützgeraden.

Für jeden Extrempunkt x von \mathbb{B}^n gilt $\{x\} = \mathbb{B}^n \cap H_{(x,1)}$, d.h. jede eigentliche Seite von \mathbb{B}^n ist Durchschnitt mit einer Stützhyperbene. Abbildung 2.9 zeigt eine andere 2-dimensionale kompakte konvexe Menge und ihre eigentlichen Seiten, die Extrempunkte enthält, für die das nicht der Fall ist.⁵

2.3.5 Bemerkung. Seien $C \subset \mathbb{R}^n$ konvex und $x \in \mathbb{R}^n$. x ist Extrempunkt von C genau dann, wenn $x \in C$ ist und

$$x \notin \text{conv}(C \setminus \{x\})$$

gilt, d.h. wenn

$$y_1, y_2 \in C \wedge \lambda \in]0, 1[\wedge x = \lambda y_1 + (1 - \lambda)y_2 \implies x = y_1 = y_2$$

Ist C nicht einpunktig, so liegen insbesondere die Extrempunkte von C in $\text{relbd}(C)$.

2.3.6 Lemma. Sei $C \subset \mathbb{R}^n$ konvex. Dann gelten die folgenden Aussagen:

- Ist H eine Stützhyperbene an C , so ist $C \cap H$ eine Seite von C .
- Seien S eine Seite von C und T eine Seite von S . Dann ist T eine Seite von C .

Beweis: (a) Seien $H = H_{(a,\beta)}$, H^+ der zugehörige Stützhalbraum, $x \in \text{relint}(C \cap H)$ und $y, z \in C$, $\lambda \in]0, 1[$ mit $x = \lambda y + (1 - \lambda)z$. Aus $\beta = a^T x = \lambda a^T y + (1 - \lambda)a^T z \leq \beta$ folgt $y, z \in H$. Es folgt $C \cap H \subset F(x)$. Sei umgekehrt $y \in C \cap H$. Dann gilt $x \in \text{relint}(C \cap \text{aff}(\{y, x\}))$, und es folgt $y \in F(x)$. Insgesamt gilt also $C \cap H = F(x)$.

⁵ Hier zeigt sich der Unterschied zwischen Seiten und so genannten 'exponierten' Seiten.

(b) Seien $s \in C$ und $t \in S$ mit $S = F_C(s)$ und $T = F_S(t)$. Ferner seien $y, z \in C$ mit $t \in \text{relint}(\text{conv}(\{y, z\}))$. Da $t \in S$ ist, gibt es ein $v \in S$ mit $s \in \text{relint}(\text{conv}(\{t, v\}))$. Sei $D := \text{conv}(\{v, y, z\})$. Da $t \in \text{relint}(\text{conv}(\{y, z\}))$ ist, die Gerade $\text{aff}(\{t, v\})$ die Punkte y und z in $\text{aff}(D)$ also streng trennt, gilt $s \in \text{relint}(D)$. Der Punkt s ist daher jeweils relativ innerer Punkt der Intervalle $D \cap \text{aff}\{s, y\}$ und $D \cap \text{aff}\{s, z\}$, und es gibt somit Punkte $\hat{y}, \hat{z} \in C$ mit $s \in \text{relint}(\text{conv}(\{y, \hat{y}\}))$ und $s \in \text{relint}(\text{conv}(\{z, \hat{z}\}))$. Somit gilt $y, z \in S$, und es folgt die Behauptung. \square

2.3.7 Korollar. Seien $C \subset \mathbb{R}^n$ konvex und $p + [0, \infty[q$ ein Extremalstrahl von C . Dann ist p ein Extremalpunkt von C .

2.3.8 Lemma. Seien $C \subset \mathbb{R}^n$ konvex und $x \in C$. Dann ist $F(x)$ konvex, und es gilt $x \in \text{relint}(F(x))$. Ist C abgeschlossen, so ist auch $F(x)$ abgeschlossen.

Beweis: Für Extremalpunkte ist die Aussage trivial. Sei daher x kein Extremalpunkt von C . Sei $S := \text{conv}(F(x))$. Da C konvex ist, gilt $S \subset C$. Angenommen, $x \notin \text{relint}(S)$, d.h. $x \in \text{relbd}(S)$. Nach Satz 2.2.13 gibt es eine eigentliche Stützhyperebene H an S in x . Aus der Definition von $F(x)$ folgt aber dann $F(x) \subset H$, also $S \subset H$, im Widerspruch dazu, dass H eigentlich ist. Somit gilt $x \in \text{relint}(S)$, und insbesondere schneidet jede Gerade G mit $G \subset \text{aff}(F(x))$ und $x \in G$ die Menge S in einem Intervall positiver Länge, und es gilt $x \in \text{relint}(C \cap G)$. Hieraus folgt $F(x) = S$ sowie $x \in \text{relint}(F(x))$.

Zum Beweis der letzten Behauptung sei $y \in \text{cl}(F(x))$. Da C abgeschlossen ist, folgt $y \in C$. Ferner ist x relativ innerer Punkt von $(x + \mathbb{R}(y - x)) \cap C$, d.h. es gibt ein $z \in C$ mit $x \in \text{relint}(\text{conv}(\{y, z\}))$. Insgesamt folgt damit die Behauptung. \square

2.3.9 Korollar. Seien $C \subset \mathbb{R}^n$ konvex und $x \in C$. Dann gelten die folgenden Aussagen:

- (a) $F(x) = C \cap \text{aff}(F(x))$;
- (b) $\hat{x} \in \text{relint}(F(x)) \Rightarrow F(x) = F(\hat{x})$.

Beweis: (a) Natürlich gilt $F(x) \subset C \cap \text{aff}(F(x))$. Sei andererseits $y \in C \cap \text{aff}(F(x))$. Nach Lemma 2.3.8 ist $x \in \text{relint}(F(x))$. Somit ist x relativ innerer Punkt des Intervalls $C \cap \text{aff}\{x, y\}$, und es folgt $y \in F(x)$.

(b) Sei $y \in F(x)$. Dann gilt $\hat{x} \in \text{relint}(C \cap \text{aff}\{\hat{x}, y\})$, und es folgt $y \in F(\hat{x})$. Damit gilt $F(x) \subset F(\hat{x})$.

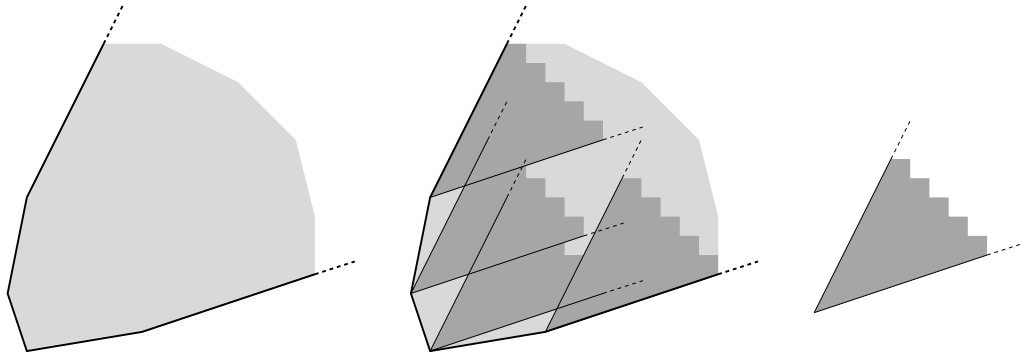
Die umgekehrte Inklusion folgt durch Vertauschung der Rollen von x und \hat{x} , wenn noch $x \in \text{relint}(F(\hat{x}))$ gezeigt wird. Wäre aber $x \in \text{relbd}(F(\hat{x}))$, so gäbe es nach Korollar 2.2.13 eine Hyperebene H mit zugehörigem abgeschlossenem Halbraum H^+ , so dass

$$F(\hat{x}) \subset H^+ \quad \wedge \quad F(\hat{x}) \not\subset H \quad \wedge \quad F(x) \subset H.$$

Das ist aber ein Widerspruch, denn zu keinem Punkt $y \in F(\hat{x}) \cap \text{int}(H^+)$ kann es einen Punkt $z \in F(\hat{x})$ geben mit $\hat{x} \in \text{relint}(C \cap \text{aff}\{y, z\})$. \square

2.3.10 Definition. Seien $C \subset \mathbb{R}^n$ abgeschlossen und konvex.

- (a) Ist $c \in C$, so heißt der (bez. Inklusion) größte in $(-c) + C$ enthaltene lineare Teilraum bzw. Kegel **Linealitätsraum** [engl.: lineality space] bzw. **Rezessionskegel** [engl.: recession cone] von C ; er wird mit $\text{ls}(C)$ bzw. $\text{rec}(C)$ bezeichnet.
- (b) Gilt $C = \emptyset$ oder $(C \neq \emptyset \wedge \text{ls}(C) = \{0\})$, so heißt C **geradenfrei** [engl.: line free].



2.8 Abbildung Rezessionskegel $\text{rec}(P)$ (rechts) eines Polyeders P (links). In der Mitte ist $\text{rec}(P)$ an jedem Extrempunkt von P abgetragen.

2.3.11 Bemerkung. Sei $C \subset \mathbb{R}^n$ abgeschlossen, konvex und nicht leer. Dann gilt

$$\text{ls}(C) = \text{ls}(\text{rec}(C)).$$

2.3.12 Bemerkung. Seien $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ und $P := \{x : Ax \leq b\} \neq \emptyset$. Dann gilt $\text{ls}(P) = \ker(A)$ und $\text{rec}(P) = \{x \in \mathbb{R}^n : Ax \leq 0\}$. Insbesondere ist P genau dann geradenfrei, wenn $\text{rang}(A) = n$ ist.

2.3.13 Lemma. Seien $C \subset \mathbb{R}^n$ nicht leer, abgeschlossen und konvex und $K := C \cap (\text{ls}(C))^\perp$. Dann ist K geradenfrei, abgeschlossen und konvex, und es gilt

$$C = K + \text{ls}(C).$$

Beweis: Als Durchschnitt von zwei abgeschlossenen, konvexen Mengen ist K abgeschlossen und konvex. Enthielte K eine Gerade G , so wäre nach Bemerkung 2.3.1

$$\text{ls}(C) \subsetneq \text{ls}(G) + \text{ls}(C) \subset \text{ls}(C),$$

im Widerspruch zur Definition des Linealitätsraums. K ist also geradenfrei.

Nach Bemerkung 2.3.1 gilt für jeden Punkt $x \in C$ auch $x + \text{ls}(C) \subset C$; es folgt

$$K + \text{ls}(C) \subset C.$$

Sei andererseits $x \in C$. Das orthogonale Komplement $(\text{ls}(C))^\perp$ schneidet $x + \text{ls}(C)$ in genau einem Punkt; sei $y \in K$ mit

$$(x + \text{ls}(C)) \cap (\text{ls}(C))^\perp = \{y\}.$$

Dann gilt $x \in y + \text{ls}(C)$. Somit ist auch

$$C \subset K + \text{ls}(C)$$

und es folgt die Behauptung. \square

2.3.14 Satz. *Seien $C \subset \mathbb{R}^n$ abgeschlossen, konvex und geradenfrei, und R bezeichne die Vereinigung aller Extremalstrahlen von C . Dann gilt*

$$C = \text{conv}(\text{ext}(C) \cup R).$$

Beweis: Natürlich gilt $\text{conv}(\text{ext}(C) \cup R) \subset C$.

Wir beweisen die umgekehrte Inklusion durch vollständige Induktion nach der Dimension k von C . Offenbar sind die Fälle $k \in \{-1, 0, 1\}$ trivial. Wir nehmen also an, die Behauptung sei für alle $k \in \mathbb{N}$ mit $k \leq k_0 \in \mathbb{N}$ bereits bewiesen, und es sei C eine abgeschlossene, konvexe Teilmenge des \mathbb{R}^n der Dimension $k_0 + 1$.

Nach Satz 2.3.2 gilt $C = \text{conv}(\text{relbd}(C))$. Sei also $x \in \text{relbd}(C)$. Nach Satz 2.2.13 existiert eine Stützhyperebene an C in x , die C nicht ganz enthält; H sei eine solche. Seien E_H die Menge der Extrempunkte sowie R_H die Vereinigung aller Extremalstrahlen von $C \cap H$. Aus Lemma 2.3.6 und der Induktionsvoraussetzung folgt somit

$$x \in C \cap H = \text{conv}(E_H \cup R_H) \subset \text{conv}(\text{ext}(C) \cup R),$$

und damit die Behauptung. \square

2.3.15 Korollar. *Sei $C \subset \mathbb{R}^n$ abgeschlossen und konvex. C besitzt genau dann einen Extrempunkt, wenn C geradenfrei und nicht leer ist.*

Beweis: ‘ \Rightarrow ’ Da C einen Extrempunkt besitzt, ist $C \neq \emptyset$. Aus Bemerkung 2.3.1 folgt ferner, dass C geradenfrei ist.

‘ \Leftarrow ’ Nach Satz 2.3.14 ist C die konvexe Hülle seiner Extrempunkte und Extremalstrahlen, und nach Korollar 2.3.7 gilt $\text{ext}(C) \neq \emptyset$. \square

2.3.16 Korollar. *Seien $C \subset \mathbb{R}^n$ nicht leer, abgeschlossen und konvex, $K := C \cap (\text{ls}(C))^\perp$, und $Y \subset \mathbb{R}^n \setminus \{0\}$, so dass für jeden Extremalstrahl S von $\text{rec}(K)$ ein $y \in Y$ existiert mit $S = [0, \infty[y$. (Y enthält also für jeden Extremalstrahl von $\text{rec}(K)$ einen Richtungsvektor.) Ferner sei Z eine Basis von $\text{ls}(C)$. Dann gilt*

$$\begin{aligned} C &= \text{conv}(\text{ext}(K)) + \text{rec}(C) = \text{conv}(\text{ext}(K)) + \text{rec}(K) + \text{ls}(C) \\ &= \text{conv}(\text{ext}(K)) + \text{pos}(Y) + \text{lin}(Z). \end{aligned}$$

Beweis: Sei R die Vereinigung aller Extremalstrahlen von K . Nach Lemma 2.3.13 ist $C = K + \text{ls}(C)$, und nach Satz 2.3.14 gilt $K = \text{conv}(\text{ext}(K) \cup R)$ sowie $\text{rec}(K) = \text{conv}([0, \infty[Y) = \text{pos}(Y)$.

Durch Anwendung von Lemma 2.3.13 auf $\text{rec}(C)$ und Bemerkung 2.3.11 erhalten wir

$$\begin{aligned} \text{rec}(C) &= \left(\text{rec}(C) \cap (\text{ls}(\text{rec}(C)))^\perp \right) + \text{ls}(\text{rec}(C)) = \left(\text{rec}(C) \cap (\text{ls}(C))^\perp \right) + \text{ls}(C) \\ &= \text{rec}(K) + \text{ls}(C). \end{aligned}$$

Es bleibt somit nur noch zu zeigen, dass

$$\text{conv}(\text{ext}(K) \cup R) = \text{conv}(\text{ext}(K)) + \text{rec}(K)$$

gilt. Die Inklusion ‘ \supset ’ folgt mit Satz 2.3.14 und Bemerkung 2.3.1 aus $\text{ext}(K) + \text{rec}(K) \subset K$. Ist umgekehrt $S := x + [0, \infty[y$ ein Extremalstrahl von K , so folgt mit Korollar 2.3.7 $S \subset \text{ext}(K) + \text{rec}(K)$, und damit auch die Inklusion ‘ \subset ’.

Insgesamt ist damit die Behauptung bewiesen. \square

2.3.17 Korollar. (*Minkowski*)

Jede kompakte konvexe Teilmenge des \mathbb{R}^n ist die konvexe Hülle ihrer Extremalpunkte.

2.3.18 Korollar. *Ist der zulässige Bereich einer linear-konvexen Maximierungsaufgabe geradenfrei und existiert ein Maximalpunkt, so gibt es einen Extremalpunkt, der optimal ist.*

Beweis: Seien C der zulässige Bereich, c der Zielfunktionsvektor der gegebenen Aufgabe, x^* Maximalpunkt und $\gamma^* = c^T x^*$. Gilt $c = 0$, so sei $H := \mathbb{R}^n$; andernfalls sei $H := H_{(c, \gamma^*)}$. In jedem Fall ist die Menge $C \cap H$ der Maximalpunkte nicht leer und geradenfrei. Mit Korollar 2.3.15 folgt die Behauptung. \square

2.3.19 Beispiel. *Sei*

$$C := \{x = (\xi_1, \xi_2)^T \in \mathbb{R}^2 : \xi_1 \geq 1 \wedge 0 \leq \xi_1 \xi_2 \leq \xi_1 - 1\}.$$

Ferner sei $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ für $x = (\xi_1, \xi_2)^T \in \mathbb{R}^2$ gegeben durch $\varphi(x) := \xi_2$. Dann gilt

$$\max_{x \in C} \varphi(x) = 1 \quad \wedge \quad \xi_2 < 1.$$

Das Maximum 1 der Zielfunktion wird also nicht angenommen.

2.3.20 Satz. *Seien $n, m \in \mathbb{N}$, $a_1, \dots, a_m \in \mathbb{R}^n \setminus \{0\}$, $A := (a_1, \dots, a_m)^T$, $b := (\beta_1, \dots, \beta_m)^T \in \mathbb{R}^m$ und $P := \{x \in \mathbb{R}^n : Ax \leq b\}$. Sei $x \in P$. Dann gilt*

$$\text{aff}(F(x)) = S^=(x) \quad \wedge \quad F(x) = P \cap S^=(x).$$

Beweis: Es reicht zu zeigen, dass $\text{aff}(F(x)) = S^=(x)$ gilt. Die zweite Behauptung folgt dann aus Korollar 2.3.9 (a).

Nach Lemma 2.3.8 ist x relativ innerer Punkt von $F(x)$. Seien $\dim(F(x)) = k$, $y_1, \dots, y_k \in \mathbb{R}^n$ linear unabhängig und $\mu \in]0, \infty[$ mit

$$j \in \{1, \dots, k\} \quad \Rightarrow \quad x + [-\mu, \mu] y_j \subset F(x).$$

Dann gilt für alle $i \in I(x)$ und $j \in \{1, \dots, k\}$

$$a_i^T(x \pm \mu y_j) = \beta_i \pm \mu a_i^T y_j \leq \beta_i$$

also $a_i^T y_1 = \dots = a_i^T y_k = 0$, und es folgt

$$\text{aff}(F(x)) \subset S^=(x).$$

Seien umgekehrt $y \in S^=(x)$, $z := y - x$ und $\mu \in]0, \infty[$. Dann gilt für jedes $i \in I(x)$

$$\begin{aligned} a_i^T(x \pm \mu z) &= \beta_i & (i \in I(x)) \\ a_i^T(x \pm \mu z) &< \beta_i \pm \mu a_i^T z & (i \notin I(x)). \end{aligned}$$

Für hinreichend kleines positives μ ist somit $x \pm \mu z \in P$, und es folgt $y \in \text{aff}(F(x))$. Insgesamt ist damit $S^=(x) = \text{aff}(F(x))$ bewiesen. \square

2.3.21 Korollar. *Polyeder besitzen höchstens endlich viele Seiten.*

2.3.22 Bemerkung. *Seien $n, m \in \mathbb{N}$, $a_1, \dots, a_m \in \mathbb{R}^n \setminus \{0\}$, $A := (a_1, \dots, a_m)^T$, $b := (\beta_1, \dots, \beta_m)^T \in \mathbb{R}^m$ und $P := \{x \in \mathbb{R}^n : Ax \leq b\}$. Ferner seien $x \in \text{relbd}(P)$ und $a \in \text{relint}(N_P(x))$. Dann gilt*

$$F(x) = P \cap H_{(a, a^T x)}.$$

Beweis: Nach Satz 2.3.20 gilt $F(x) = P \cap S^=(x)$; wir zeigen $P \cap S^=(x) = P \cap H_{(a, a^T x)}$. Gemäß Lemma 2.2.28 seien $\lambda_i \in [0, \infty[$ für $i \in I(x)$ mit $a = \sum_{i \in I(x)} \lambda_i a_i$. Dann gilt für $y \in S^=(x)$

$$a^T y = \sum_{i \in I(x)} \lambda_i a_i^T y = \sum_{i \in I(x)} \lambda_i a_i^T x = a^T x,$$

d.h. $y \in H_{(a, a^T x)}$, und es folgt $P \cap S^=(x) \subset P \cap H_{(a, a^T x)}$.

Sei nun umgekehrt $y \in P \cap H_{(a, a^T x)}$. Angenommen es gäbe ein $i_0 \in I(x)$ mit $a_{i_0}^T y < \beta_{i_0} = a_{i_0}^T x$. Da $a \in \text{relint}(N_P(x))$ ist, gibt es $z \in N_P(x)$ und $\lambda \in]0, 1[$ mit $a = \lambda a_{i_0} + (1 - \lambda)z$, und es folgt

$$a^T y = \lambda a_{i_0}^T y + (1 - \lambda)z^T y < \lambda a_{i_0}^T x + (1 - \lambda)z^T x = a^T x.$$

Somit gilt $y \notin H_{(a, a^T x)}$, im Widerspruch zur Wahl von y . Insgesamt folgt damit die Behauptung. \square

2.3.23 Korollar. *Seien $P \subset \mathbb{R}^n$ ein Polyeder und $F \subset P$. F ist genau dann eine eigentliche Seite von P , wenn es eine eigentliche Stützhyperbene H an P gibt mit $P \cap H = F$.*

2.3.24 Bezeichnung. *Sei $P \subset \mathbb{R}^n$ ein Polyeder. Jede 0-Seite von P heißt **Ecke** [engl.: vertex], jede 1-Seite **Kante** [engl.: edge] und jede $\dim(P) - 1$ -Seite **Facette** [engl.: edge] von P .*

2.3.25 Bemerkung. *Sei P ein Polyeder des \mathbb{R}^n . P besitzt genau dann eine Ecke, wenn $P \neq \emptyset$ und $\text{rang}(A) = n$ gilt.*

Beweis: Die Behauptung folgt aus Bemerkung 2.3.12 und Korollar 2.3.15. \square

2.3.26 Beispiel. *Wir betrachten noch einmal das Polytop (in der \mathcal{H} -Darstellung) aus Beispiel 1.2.2.*

$$\begin{array}{rcll}
\max & \xi_1 + \xi_2 + \xi_3 & & \\
& \xi_1 + 2\xi_2 + \xi_3 & \leq & 3 \quad (1) \\
& -2\xi_1 + \xi_2 & \leq & 0 \quad (2) \\
& \xi_1 & \leq & 1 \quad (3) \\
& & \xi_2 & \leq 1 \quad (4) \\
& & & \xi_3 \leq 1 \quad (5) \\
& -\xi_1 & & \leq 0 \quad (6) \\
& & -\xi_2 & \leq 0 \quad (7) \\
& & & -\xi_3 \leq 0 \quad (8)
\end{array}$$

Sei

$$x^* := \left(1, \frac{3}{4}, \frac{1}{2}\right)^T.$$

Dann gilt $x^* \in P$, $I(x^*) = \{1,3\}$, und $S^=(x^*)$ ist durch die beiden Gleichungen

$$\begin{array}{rcl}
\xi_1 + 2\xi_2 + \xi_3 & = & 3 \\
\xi_1 & = & 1
\end{array}$$

beschrieben. Als Lösung dieses Teilsystems erhält man

$$S^=(x^*) = \text{aff}(F(x^*)) = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \mathbb{R} \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}.$$

Bezeichnen wir den reellen Parameter mit λ , so ergeben sich aus den Bedingungen (2), (4), (5), (7) und (8) die folgenden Restriktionen:

$$\lambda \leq 2 \wedge \lambda \leq 1 \wedge \lambda \geq 1/2 \wedge \lambda \geq 0 \wedge \lambda \leq 1;$$

Ungleichung (6) ist keine Einschränkung. Es folgt

$$F(x^*) = \text{conv} \left\{ \begin{pmatrix} 1 \\ 1/2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

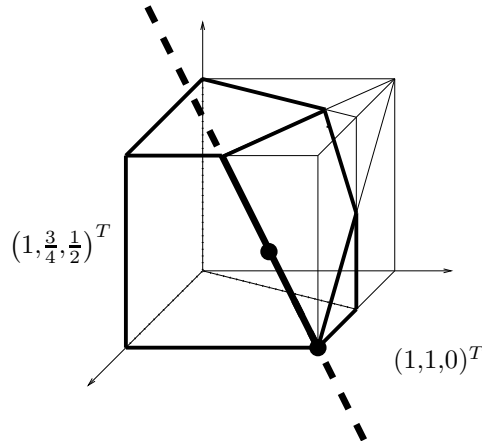
Die Ecke $v := (1,1,0)^T$ ist Durchschnitt der vier Hyperebenen

$$\begin{array}{rcl}
\xi_1 + 2\xi_2 + \xi_3 & = & 3 \\
\xi_1 & = & 1 \\
& \xi_2 & = 1 \\
& & \xi_3 = 0,
\end{array}$$

d.h. es gilt $I(v) = \{1,3,4,8\}$. Jeweils drei dieser Gleichungen bestimmen v aber bereits eindeutig; die gegebene Darstellung von $S^=(v)$ ist also redundant.

2.3.27 Korollar. Sei P ein Polyeder im \mathbb{R}^n . Dann existieren endliche Teilmengen V , Y und Z des \mathbb{R}^n , so dass

$$P = \text{conv}(V) + \text{pos}(Y) + \text{lin}(Z).$$



2.9 Abbildung Seiten, relativ innere Punkte und affinen Hüllen.

Seien $P \neq \emptyset$ und $k := n - \text{rang}(A)$. Ist $k = n$, so folgt $Z = \{0\}$. Andernfalls seien z_1, \dots, z_k eine Basis des Lösungsraums des homogenen Gleichungssystems $Ax = 0$, und Z kann gleich $\{z_1, \dots, z_k\}$ gewählt werden. Ferner seien $\hat{A} := A$ und $\hat{b} := b$ für $k = 0$ und $\hat{A} := (A^T, z_1, \dots, z_k)^T \in \mathbb{R}^{(m+k) \times n}$ und $\hat{b} := (b^T, 0, \dots, 0)^T \in \mathbb{R}^{m+k}$ für $k \geq 1$. Dann können die Elemente von V und Y als Lösungen von $n \times n$ Teilsystemen von $\hat{A}x = \hat{b}$ vom Rang n bzw. $n - 1$ Teilsystemen von $\hat{A}x = 0$ vom Rang $n - 1$ gewählt werden.

Beweis: Die Behauptung ist trivial für $P = \emptyset$. Seien daher $P \neq \emptyset$, $Z := \{0\}$ für $\text{ls}(P) = \{0\}$ bzw. eine Basis von $\text{ls}(P)$ und $Q := P \cap (\text{lin}(Z))^\perp$. Dann sind Q ein geradenfreies Polyeder und $\text{rec}(P)$ ein polyederischer Kegel. Seien V die Menge der Ecken von Q und $Y := \{0\}$ für $\text{rec}(Q) = \{0\}$. Gilt $\text{rec}(Q) \neq \{0\}$, so sei $Y \subset \mathbb{R}^n \setminus \{0\}$ und enthalte für jede Kante von $\text{rec}(Q)$ einen Richtungsvektor. Dann gilt nach Korollar 2.3.16

$$P = \text{conv}(V) + \text{pos}(Y) + \text{lin}(Z).$$

Nach Bemerkung 2.3.12 ist $\text{ls}(P) = \ker(A)$. Mit $k = \dim(\text{ls}(P))$ folgt, dass die Wahl $Z = \{0\}$ bzw. $Z = \{z_1, \dots, z_k\}$ zulässig ist. Nach Satz 2.3.20 ist jede Ecke von Q Lösung eines $n \times n$ Teilsystems von $\hat{A}x = \hat{b}$ mit einer Teilmatrix vom Rang n . Nach Bemerkung 2.3.12 gilt ferner $\text{rec}(Q) = \{x : \hat{A}x \leq \hat{b}\}$. Die affine Hülle jeder Kante von $\text{rec}(Q)$ ist somit, wieder nach Satz 2.3.20, die Lösungsmenge eines $(n - 1) \times n$ Teilsystems von $\hat{A}x = 0$ mit einer Teilmatrix vom Rang $n - 1$. Insgesamt folgt damit die Behauptung. \square

2.3.28 Bezeichnung. Seien $P \subset \mathbb{R}^n$ und V, S endliche Teilmengen des \mathbb{R}^n . Gilt

$$P = \text{conv}(V) + \text{pos}(S),$$

so heißt P \mathcal{V} -**Polyeder** [engl.: \mathcal{V} -polyhedron]. Sind P ein Polyeder und V, S endliche Teilmengen des \mathbb{R}^n mit

$$P = \text{conv}(V) + \text{pos}(S),$$

so heißt $(n; V, S)$ eine \mathcal{V} -**Darstellung** [engl.: \mathcal{V} -presentation] von P .

Ist $(n; V, S)$ eine \mathcal{V} -Darstellung von P und existiert keine \mathcal{V} -Darstellung $(n; V', S')$ von P mit $V' \subsetneq V$ oder $S' \subsetneq S$, so heißt $(n; V, S)$ **irredundant** [engl.: irredundant].

2.3.29 Satz. Sei $P \subset \mathbb{R}^n$. P ist genau dann ein \mathcal{H} -Polyeder, wenn P ein \mathcal{V} -Polyeder ist.

Beweis: Für $P = \emptyset$ ist die Aussage trivial. Sei also im folgenden $P \neq \emptyset$.

‘ \Rightarrow ’ folgt aus Korollar 2.3.27.

‘ \Leftarrow ’ Seien $V := \{v_1, \dots, v_p\}$ und $S := \{s_1, \dots, s_q\}$ endliche Teilmengen des \mathbb{R}^n mit $P = \text{conv}(V) + \text{pos}(S)$. Ferner seien $A_1 := (v_1, \dots, v_p)$, $A_2 := (s_1, \dots, s_q)$ sowie

$$Q = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : x \in \mathbb{R}^n \wedge y \in \mathbb{R}^p \wedge z \in \mathbb{R}^q \wedge x = A_1 y + A_2 z \wedge y, z \geq 0 \wedge \mathbf{1}^T y = 1 \right\}.$$

Da Q durch die linearen Bedingungen

$$\begin{array}{rcl} x & - & A_1 y & - & A_2 z & = & 0 \\ & & \mathbf{1}^T y & & & = & 1 \\ & & y & & & \geq & 0 \\ & & & & z & \geq & 0 \end{array}$$

gegeben ist, ist Q ein \mathcal{H} -Polyeder, und P ist die Orthogonalprojektion von Q auf den \mathbb{R}^n der Koordinaten x . Nach Korollar 1.3.14 ist P ein \mathcal{H} -Polyeder, und es folgt die Behauptung. \square

2.3.30 Korollar. Die Anzahl der in der Fourier-Motzkin-Elimination auftretenden Operationen kann exponentiell mit der Anzahl der Ungleichungen des Ausgangssystems wachsen, selbst dann, wenn die Koeffizientenmatrix nur Komponenten aus $\{-1, 0, 1\}$ besitzt.

Beweis: Wir wenden die Konstruktion des Beweises von Satz 2.3.29 auf $P := \text{conv}(\{\pm u_1, \dots, \pm u_n\})$, d.h. auf das Standard-Kreuzpolytop Q_n des \mathbb{R}^n an. Gemäß Bemerkung 1.3.5 überführt die Fourier-Motzkin-Elimination das lineare System

$$\begin{array}{rcl} x & - & (E_n, -E_n)y & = & 0 \\ & & \mathbf{1}^T y & = & 1 \\ & & y & \geq & 0 \end{array}$$

durch Elimination von y in $2n$ Schritten in eine \mathcal{H} -Darstellung von Q_n . Nach Lemma 2.1.8 gilt

$$Q_n = \{x \in \mathbb{R}^n : y \in \{-1, 1\}^n \Rightarrow y^T x \leq 1\}.$$

Keine dieser Ungleichungen ist redundant, d.h. Q_n hat 2^n Facetten. Das Ausgangssystem besteht aus $n + 1$ Gleichungen und $2n$ Ungleichungen (oder, wenn man möchte, $4n + 2$ Ungleichungen) in $3n$ Variablen. Von den hieraus in $2n$ Schritten entstehenden Ungleichungen sind aber 2^n irredundant. \square

2.3.31 Satz. Durch (n, m, A, b, c) sei eine lineare Optimierungsaufgabe in natürlicher Form spezifiziert; seien $P := \{x \in \mathbb{R}^n : Ax \leq b\} \neq \emptyset$ und φ die durch $\varphi(x) := c^T x$ für $x \in \mathbb{R}^n$ definierte Zielfunktion. Sind P geradenfrei und φ auf P nach oben beschränkt, so enthält P eine Ecke, die maximal ist.

Beweis: Für $c = 0$ ist jeder Punkt von P optimal. Mit Korollar 2.3.15 folgt dann die Behauptung.

Im folgenden sei $c \neq 0$. Ferner sei $\gamma^* = \sup_{x \in P} c^T x$. Nach den Korollaren 2.3.18 und Korollar 2.3.23 reicht es zu zeigen, dass $P \cap H_{(c, \gamma^*)} \neq \emptyset$ ist.

Entsprechend Korollar 2.3.27 seien $V := \{v_1, \dots, v_p\}$ und $S := \{s_1, \dots, s_q\}$ endliche Teilmengen des \mathbb{R}^n mit $P = \text{conv}(V) + \text{pos}(S)$. Ferner seien $\lambda_1, \dots, \lambda_p, \mu_1, \dots, \mu_q \in [0, \infty[$ mit $\sum_{i=1}^p \lambda_i = 1$. Dann gilt⁶

$$c^T \left(\sum_{i=1}^p \lambda_i v_i + \sum_{j=1}^q \mu_j s_j \right) = \sum_{i=1}^p \lambda_i c^T v_i + \sum_{j=1}^q \mu_j c^T s_j \leq \gamma^*,$$

und es folgt

$$x \in \text{pos}(S) \quad \Rightarrow \quad c^T x \leq 0.$$

Nach Korollar 2.3.7 ist $V \neq \emptyset$; somit gilt

$$\sup_{x \in P} c^T x = \sup_{x \in \text{conv}(V)} c^T x.$$

Nach Satz 2.1.13 ist $\text{conv}(V)$ kompakt, d.h. φ nimmt sein Maximum über P an. \square

⁶ mit der üblichen Setzung $\sum_{x \in \emptyset} \alpha(x) = 0$

3 Der Simplex-Algorithmus

3.1 Grundstruktur des Algorithmus

3.1.1 Bemerkung. Durch (n, m, A, b, c) sei eine lineare Optimierungsaufgabe in natürlicher Form spezifiziert; seien

$$P := \{x \in \mathbb{R}^n : Ax \leq b\} \quad \wedge \quad T := \{x \in \mathbb{R}^n : Ax = 0\}^\perp \quad \wedge \quad Q := P \cap T.$$

Dann gelten die folgenden Aussagen:

- (a) Q ist geradenfrei;
- (b) $P = \emptyset \Leftrightarrow Q = \emptyset$;
- (c) $Q \neq \emptyset \Leftrightarrow \text{ext}(Q) \neq \emptyset$;
- (d) $Q \neq \emptyset \wedge c \notin T \Rightarrow \max_{x \in P} c^T x = \infty$;
- (e) $c \in T \Rightarrow \max_{x \in P} c^T x = \max_{x \in Q} c^T x$.

3.1.2 Lemma. Durch (n, m, A, b, c) sei eine lineare Optimierungsaufgabe in natürlicher Form spezifiziert, v sei eine Ecke des zulässigen Bereichs P . v ist genau dann nicht bereits Maximalpunkt, wenn es eine Kante K von P gibt mit $v \in K$ und $c^T x > c^T v$ für alle $x \in K \setminus \{v\}$.

Beweis: Da ' \Leftarrow ' klar ist, bleibt nur ' \Rightarrow ' zu zeigen.

Sei φ die durch $\varphi(x) := c^T x$ für $x \in \mathbb{R}^n$ definierte Zielfunktion.

Ist $P = \{v\}$, so ist v Maximalpunkt von φ bez. P . Wir setzen daher voraus, dass $\dim(P) \geq 1$ gilt. Nach Korollar 2.3.15 bzw. Bemerkung 2.3.25 gilt $\text{ls}(P) = \{0\}$. Es folgt $\text{ext}(S_P(v)) = \{0\}$, und nach Lemma 2.2.28 und Korollar 2.3.27 ist $S_P(v)$ die konvexe Hülle seiner Kanten. Seien $s_1, \dots, s_k \in \mathbb{R}^n \setminus \{0\}$ mit $S_P(v) = \text{pos}\{s_1, \dots, s_k\}$. Gilt $c^T s_i \leq 0$ für alle $i = 1, \dots, k$, so ist v Maximalpunkt bez. $v + S_P(v)$, nach Lemma 2.2.25 also auch bez. P . Ist v nicht Maximalpunkt bez. P , so gibt es daher mindestens ein $i_0 \in \{1, \dots, k\}$ mit $c^T s_{i_0} > 0$. Sei nun $K := (v + [0, \infty[s_{i_0}) \cap P$. Dann ist K eine Kante von P , und es folgt die Behauptung. \square

3.1.3 Bezeichnung. Jede Kante K gemäß Lemma 3.1.2 heißt **Verbesserungskante** [engl.: *improving edge*] von P in v .

3.1.4 Definition. Seien P ein nicht leeres, geradenfreies Polyeder, $V := \text{ext}(P)$ und E die Menge der Kanten¹ von P . Dann heißt das Paar (V, E) **1-Skelett** [engl.: *1-skeleton*] von P . Ferner sei $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$.

¹ Bislang wurden 1-Seiten mit großen lateinischen Buchstaben wie F oder K bezeichnet. Im folgenden werden wir in Analogie zu den üblichen Bezeichnungen der Graphentheorie die Kanten von P oft mit e_i bezeichnen.

Ein bezüglich φ **aufsteigender (Kanten-) Pfad** [engl.: increasing edge path] im 1-Skelett von P ist eine Sequenz von einem der folgenden beiden Typen:²

- (a) $(v_0, e_1, v_1, e_2, \dots, e_p, v_p)$ mit $v_0, \dots, v_p \in V$ und $e_1, \dots, e_p \in E$ mit $e_i = \text{conv}\{v_{i-1}, v_i\}$ für $i = 1, \dots, p$ und $\varphi(v_0) < \varphi(v_1) < \dots < \varphi(v_p)$;
- (b) $(v_0, e_1, v_1, e_2, \dots, e_p)$ mit $v_0, \dots, v_{p-1} \in V$ und $e_1, \dots, e_p \in E$ mit $e_i = \text{conv}\{v_{i-1}, v_i\}$ für $i = 1, \dots, p-1$ sowie $v_{p-1} \in e_p$; ferner gilt $\varphi(v_0) < \varphi(v_1) < \dots < \varphi(v_{p-1})$, und φ ist auf e_p nach oben unbeschränkt.

Im ersten Fall sprechen wir von einem aufsteigenden Pfad **von** v_0 **nach** v_p , im zweiten von einem **von** v_0 **nach** ∞ .

3.1.5 Bemerkung. Durch (n, m, A, b, c) sei eine lineare Optimierungsaufgabe in natürlicher Form spezifiziert, φ sei die durch $\varphi(x) := c^T x$ für $x \in \mathbb{R}^n$ definierte Zielfunktion, und v sei eine (nicht optimale) Ecke des zulässigen Bereichs P . Ist φ auf P nach oben beschränkt, so existiert eine Ecke v^* von P und ein bez. φ aufsteigender Pfad von v nach v^* . Ist φ auf P nach oben unbeschränkt, so existiert ein bez. φ aufsteigender Pfad von v nach ∞ .

3.1.6 Prozedur: Simplex-Algorithmus (geometrische Form)

```

INPUT:    $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $c \in \mathbb{R}^n$ 
          $P \leftarrow \{x : Ax \leq b\}$ ,  $T \leftarrow \{x : Ax \leq 0\}^\perp$ ,  $Q \leftarrow P \cap T$ 
OUTPUT:  Ecke  $v^*$  von  $Q$  mit  $\max_{x \in P} c^T x = c^T v^*$ , falls eine solche existiert;
         OR Meldung 'unzulässig', falls  $P = \emptyset$ ;
         OR Meldung 'unbeschränkt', falls  $\max_{x \in P} c^T x = \infty$ .
BEGIN    IF  $Q = \emptyset$  THEN Meldung 'unzulässig'
         ELSE IF  $c \notin T$  THEN Meldung 'unbeschränkt'
         ELSE sei  $v$  eine Ecke von  $Q$ 
           WHILE  $v$  ist nicht optimal DO
             BEGIN
               Finde eine Verbesserungskante  $e$  von  $Q$  in  $v$ 
               IF  $e$  ist unbeschränkt THEN Meldung 'unbeschränkt'
               ELSE sei  $w$  Ecke von  $Q$  mit  $e = \text{conv}\{v, w\}$ 
                  $v \leftarrow w$ 
             END
           END
         OUTPUT  $v$ 
END
```

3.1.7 Bemerkung. Durch (n, m, A, b, c) sei eine lineare Optimierungsaufgabe in natürlicher Form spezifiziert. Der Simplex-Algorithmus in seiner geometrischen Form löst die Aufgabe korrekt in endlich vielen Schritten.

Beweis: Die Korrektheit folgt aus den Bemerkungen 3.1.1 und 3.1.5, die Endlichkeit aus Korollar 2.3.21. \square

² Im Sinne von Definition ?? liegt ein Weg vor, wenn die letzte Kante beschränkt ist. Andernfalls müsste $\text{ext}(P)$ um je eine Ecke bei $\pm\infty$ ergänzt werden, mit denen alle unbeschränkten Kanten verbunden sind, je nachdem, wie sich die Zielfunktion auf ihnen verhält.

3.1.8 Lemma. Durch (n, m, A, b, c) sei eine lineare Optimierungsaufgabe in natürlicher Form spezifiziert. Seien $k := \text{rang}(A)$,

$$A_{1,1} \in \mathbb{R}^{k \times k} \quad \wedge \quad A_{1,2} \in \mathbb{R}^{k \times (n-k)} \quad \wedge \quad A_{2,1} \in \mathbb{R}^{(m-k) \times k} \quad \wedge \quad A_{2,2} \in \mathbb{R}^{(m-k) \times (n-k)}$$

mit $\text{rang}(A_{1,1}) = k$ und

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} \quad \wedge \quad S_1 := \begin{pmatrix} A_{1,1} \\ A_{2,1} \end{pmatrix} \quad \wedge \quad S_2 := \begin{pmatrix} A_{1,2} \\ A_{2,2} \end{pmatrix}.$$

Ferner seien

$$P := \{x \in \mathbb{R}^n : Ax \leq b\} \quad \wedge \quad T := \{x \in \mathbb{R}^n : Ax = 0\}^\perp \quad \wedge \quad Q := P \cap T$$

und $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^k$ die orthogonale Projektion auf die ersten k Koordinaten. Außerdem seien für $y \in \mathbb{R}^n$

$$y_1 \in \mathbb{R}^k \quad \wedge \quad y_2 \in \mathbb{R}^{n-k} \quad \wedge \quad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Dann gelten die folgenden Aussagen.

- (a) $c \in T \Leftrightarrow c_2 = A_{1,2}^T (A_{1,1}^T)^{-1} c_1$.
- (b) $x \in T \Rightarrow c^T x = (c_1 + A_{1,1}^{-1} A_{1,2} c_2)^T x_1$.
- (c) $Q = \left\{ x : Ax \leq b \wedge x_2 = A_{1,2}^T (A_{1,1}^T)^{-1} x_1 \right\}$.
- (d) $\pi|_T$ ist injektiv, und es gilt

$$\pi(Q) = \left\{ x_1 \in \mathbb{R}^k : \left(S_1 + S_2 A_{1,2}^T (A_{1,1}^T)^{-1} \right) x_1 \leq b \right\}.$$

Beweis: Es gilt

$$T = (\ker(A))^\perp = \text{lin} \{a_1, \dots, a_k\}.$$

Somit ist $y \in T$ genau dann, wenn ein $z \in \mathbb{R}^k$ existiert mit $y = (A_{1,1}, A_{1,2})^T z$, d.h.

$$y_1 = A_{1,1}^T z \quad \wedge \quad y_2 = A_{1,2}^T z.$$

Da $A_{1,1}$ regulär ist, folgt

$$y_2 = A_{1,2}^T z = A_{1,2}^T (A_{1,1}^T)^{-1} y_1.$$

Durch Anwendung dieser Beziehung auf $y := c$ und $y := x$ folgen (a), (b), (c) und die Injektivität von π auf T . Die zweite Behauptung von (d) ergibt sich mit (c) aus

$$\begin{aligned} Q &= \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : S_1 x_1 + S_2 x_2 \leq b \wedge x_2 = A_{1,2}^T (A_{1,1}^T)^{-1} x_1 \right\} \\ &= \left\{ \begin{pmatrix} x_1 \\ A_{1,2}^T (A_{1,1}^T)^{-1} x_1 \end{pmatrix} : \left(S_1 + S_2 A_{1,2}^T (A_{1,1}^T)^{-1} \right) x_1 \leq b \right\}. \end{aligned}$$

□

3.1.9 Beispiel. Führen wir den Übergang gemäß Lemma 3.1.8 zu der dimensionsreduzierten Aufgabe mit geradenfreiem zulässigen Bereich an dem Beispiel

$$\begin{aligned} \max \quad & -\xi_1 + 2\xi_2 \\ -2\xi_1 \quad & + 4\xi_2 \leq 4 \\ \xi_1 \quad & - 2\xi_2 \leq 2 \end{aligned}$$

explizit durch, so erhalten wir die transformierte Aufgabe

$$\begin{aligned} \max \quad & -5\xi_1 \\ -10\xi_1 \quad & \leq 4 \\ 5\xi_1 \quad & \leq 2. \end{aligned}$$

Der zulässige Bereich dieser Aufgabe ist das Intervall $[-\frac{2}{5}, \frac{2}{5}]$. Es entsteht durch Projektion von $P \cap T$ auf die ξ_1 -Achse.

3.1.10 Lemma. Seien $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $P := \{x \in \mathbb{R}^n : Ax \leq b\}$ Es sei $\text{rang}(A, -b) = n + 1$ und

$$Q := \left\{ z := \begin{pmatrix} x \\ \xi_{n+1} \end{pmatrix} \in \mathbb{R}^{n+1} : Ax - b\xi_{n+1} \leq 0 \wedge \xi_{n+1} \leq 1 \right\}$$

oder³ $\text{rang}(A) = n$ und

$$Q := \left\{ z := \begin{pmatrix} x \\ \xi_{n+1} \end{pmatrix} \in \mathbb{R}^{n+1} : Ax - b\xi_{n+1} \leq 0 \wedge 0 \leq \xi_{n+1} \leq 1 \right\}.$$

Dann ist 0 Ecke von Q , und es gilt

$$\max_{z \in Q} u_{n+1}^T z \in \{0, 1\} \quad \wedge \quad (P \neq \emptyset \Leftrightarrow \max_{z \in Q} u_{n+1}^T z = 1).$$

*Beweis:*⁴ Es gilt $0 \in Q$, also auch $\max_{z \in Q} u_{n+1}^T z \geq 0$. Alle Bedingungen $Ax - b\xi_{n+1} \leq 0$ sowie $\xi_{n+1} \leq 0$ im zweiten Fall sind in 0 aktiv. Entweder wegen $\text{rang}(A, -b) = n + 1$ oder wegen

$$\text{rang} \begin{pmatrix} A & -b \\ 0 & 1 \end{pmatrix} = n + 1$$

ist 0 nach Satz ?? somit Ecke von Q .

Für $0 < \xi_{n+1} \leq 1$ gilt mit $\hat{x} := \xi_{n+1} \cdot x$

$$A\hat{x} - b\xi_{n+1} \leq 0 \quad \Leftrightarrow \quad Ax \leq b.$$

Ferner ist $P \times \{1\} = Q \cap H_{(u_{n+1}, 1)}^{\geq}$. Hieraus folgen auch die zweite und dritte Aussage. \square

3.1.11 Bezeichnung. Seien (n, m, A, b) eine \mathcal{H} -Darstellung eines Polyeders P und v eine Ecke von P . Gilt $|I(v)| = n$, so heißt v **regulär**, andernfalls **überbestimmt**⁵ [engl.: *degenerated*]

³ Der Fall $\text{rang}(A) = \text{rang}(A, -b) = n$ bedeutet, dass bereits $Ax = b$ (eindeutig) lösbar. Der Lösungsvektor v ist Ecke von P ; tatsächlich ist $-v + P$ ein Kegel. Wir brauchen also eigentlich kein weiteres Hilfsproblem zu lösen.

⁴ Vgl. Bemerkung ??.

⁵ In der Literatur findet man hierfür auch den Begriff **entartet** oder **degeneriert**.

3.1.12 Bemerkung. Seien $A := (a_1, \dots, a_m)^T \in \mathbb{R}^{m \times n}$, $b := (\beta_1, \dots, \beta_m)^T \in \mathbb{R}^m$,

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

$$Q := \left\{ z := \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n+m} : Ax + y \leq b \wedge x \geq 0 \wedge y \leq 0 \right\}$$

und für $i = 1, \dots, m$

$$\omega_i := \begin{cases} 0 & \text{für } \beta_i \geq 0; \\ \beta_i & \text{für } \beta_i < 0, \end{cases}$$

sowie

$$w = (\omega_1, \dots, \omega_m)^T \quad \wedge \quad z_0 = \begin{pmatrix} 0 \\ w \end{pmatrix}.$$

Dann ist z_0 Ecke von Q , und es gilt

$$\max_{(x^T, y^T)^T \in Q} \mathbf{1}^T y \leq 0 \quad \wedge \quad \left(P \neq \emptyset \Leftrightarrow \max_{(x^T, y^T)^T \in Q} \mathbf{1}^T y = 0 \right).$$

Beweis: Für die Koeffizientenmatrizen von P bzw. Q gilt

$$\text{rang} \begin{pmatrix} A \\ -E_n \end{pmatrix} = n \quad \wedge \quad \text{rang} \begin{pmatrix} A & E_m \\ -E_n & 0 \\ 0 & E_m \end{pmatrix} = n + m.$$

Nach Konstruktion gilt $z_0 \in Q$. Ferner sind alle Bedingungen $x \geq 0$ sowie mit $y = (\eta_1, \dots, \eta_m)^T$ für jedes $i \in \{1, \dots, m\}$ mindestens eine der Bedingungen $a_i^T x + \eta_i \leq \beta_i$ oder $\eta_i \leq 0$ aktiv. Nach Satz ?? ist z_0 somit Ecke von Q . Natürlich folgt aus $y \leq 0$ auch $\max_{z \in Q} \mathbf{1}^T y \leq 0$ mit Gleichheit genau dann, wenn es einen Punkt $x_0 \in P$ gibt, so dass $(x_0^T, 0)^T \in Q$ gilt. Insgesamt folgt damit die Behauptung. \square

3.1.13 Beispiel. Gegeben sei die LP-Zulässigkeitsaufgabe

$$\begin{aligned} \xi &\leq 1 \\ -\xi &\leq 0. \end{aligned}$$

Die in Bemerkung 3.1.12 konstruierte Hilfsaufgabe ist dann

$$\begin{aligned} \max \quad & \eta_1 + \eta_2 \\ \xi + \quad & \eta_1 \leq 1 \\ -\xi + \quad & \eta_2 \leq 0 \\ & \eta_1 \leq 0 \\ & \eta_2 \leq 0. \end{aligned}$$

Der Punkt $(\xi, \eta_1, \eta_2)^T = 0$ ist (wie nach Bemerkung 3.1.12 zu erwarten) Ecke des zulässigen Bereichs der Hilfsaufgabe.

Betrachten wir andererseits die gleiche Zulässigkeitsaussage, aber ohne die Nichtnegativitätsbedingung, d.h.

$$\xi \leq 1,$$

so ist die zugehörige Hilfsaufgabe

$$\begin{array}{r} \max \quad \eta \\ \xi + \eta \leq 1 \\ \eta \leq 0. \end{array}$$

Der Punkt $(\xi, \eta)^T = 0$ ist zulässig, aber keine Ecke mehr.

3.1.14 Bezeichnung. Seien $A := (a_1, \dots, a_m)^T \in \mathbb{R}^{m \times n}$ und $I \subset \{1, \dots, m\}$. Dann sei

$$A_I := (a_i : i \in I)^T,$$

d.h. A_I entsteht aus A durch Streichen aller Zeilen, deren Indizes nicht in I liegen.

3.1.15 Lemma. Seien $A := (a_1, \dots, a_m)^T$ mit $\text{rang}(A) = n$, $b := (\beta_1, \dots, \beta_m)^T$ und $P = \{x \in \mathbb{R}^n : Ax \leq b\}$. Ferner seien $M := \{1, \dots, m\}$ und $I \subset M$ mit $|I| = n$, so dass $\text{rang}(A_I) = n$ ist,

$$v := (\nu_1, \dots, \nu_n)^T := A_I^{-1}b_I \quad \wedge \quad J := \{i \in M : a_i^T v > \beta_i\} \quad \wedge \quad k := |J|.$$

Die Komponenten der Vektoren $y \in \mathbb{R}^k$ werden mit den Elementen von J indiziert, d.h. wir schreiben $y := (\eta_i : i \in J)^T$ und ferner $z := (x^T, y^T)^T$. Seien nun

$$Q := \left\{ z \in \mathbb{R}^{n+k} : (i \in M \setminus J \Rightarrow a_i^T x \leq \beta_i) \wedge (i \in J \Rightarrow (a_i^T x + \eta_i \leq \beta_i \wedge \eta_j \leq 0)) \right\}.$$

sowie

$$(i \in J \Rightarrow \omega_i := \beta_i - a_i^T v) \quad \wedge \quad w := (\omega_i : i \in J)^T \quad \wedge \quad z_0 := (v^T, w^T)^T.$$

Dann ist z_0 Ecke von Q , und es gilt

$$\max_{z \in Q} \mathbf{1}^T y \leq 0 \quad \wedge \quad (P \neq \emptyset \Leftrightarrow \max_{z \in Q} \mathbf{1}^T y = 0).$$

Beweis: Nach Konstruktion gilt $z_0 \in Q$, und z_0 ist Lösung des Gleichungssystems

$$\begin{array}{rcl} a_i^T x & = & \beta_i \quad (i \in I) \\ a_i^T x + \eta_i & = & \beta_i \quad (i \in J). \end{array}$$

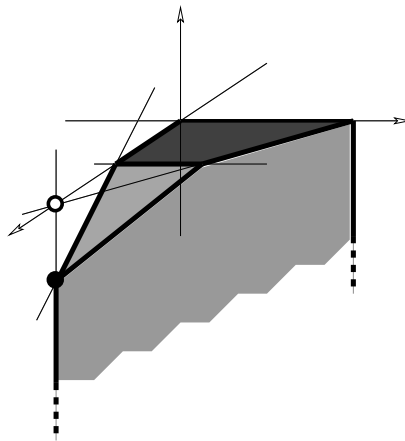
Da die zugehörige $(n+k) \times (n+k)$ Koeffizientenmatrix nach Konstruktion regulär ist, ist z_0 nach Satz ?? Ecke von Q .

Natürlich folgt aus $y \leq 0$ auch $\max_{z \in Q} \mathbf{1}^T y \leq 0$ mit Gleichheit genau dann, wenn es einen Punkt $x_0 \in P$ gibt, so dass $(x_0^T, 0)^T \in Q$ gilt. Insgesamt folgt damit die Behauptung. \square

3.1.16 Beispiel. Sei P gegeben durch die Bedingungen

$$\begin{array}{rcl} \xi_1 + \xi_2 & \leq & 2 \\ -\xi_1 & \leq & 0 \\ & - & \xi_2 \leq 0 \\ \xi_1 & \leq & 1 \end{array}$$

Die Koeffizientenmatrix hat Rang 2, und wir können etwa $I := \{1, 3\}$ wählen. Dann gilt



3.1 Abbildung Die zulässigen Bereiche P (dunkle Fläche in der (ξ_1, ξ_2) -Ebene) und Q . Der Ausgangspunkt v und die Ecke z_0 (schwarz) sind hervorgehoben.

$$A_I = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} = A_I^{-1} \quad \wedge \quad v = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Da v die ersten drei Bedingungen erfüllt, nicht aber die letzte, erhalten wir die Hilfsaufgabe

$$\begin{array}{rcll} \max & \eta & & \\ \xi_1 & + & \xi_2 & \leq 2 \\ -\xi_1 & & & \leq 0 \\ & - & \xi_2 & \leq 0 \\ \xi_1 & & + & \eta \leq 1 \\ & & & \eta \leq 0. \end{array}$$

Der Vektor $z_0 = (2, 0, -1)^T$ ist Ecke des zulässigen Bereichs Q . Die von z_0 ausgehenden Kanten haben die Richtungen

$$\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

Die beiden ersten sind beschränkt, und man erreicht über jede von ihnen eine Ecke von $P \times \{0\}$, die dann als Startecke für den Simplex-Algorithmus jeder LP-Aufgabe über P verwenden kann.

3.2 Finden einer Verbesserungskante

3.2.1 Notation. Durch (n, m, A, b, c) sei eine LP-Aufgabe in natürlicher Form spezifiziert. Es seien $M := \{1, \dots, m\}$, $a_1, \dots, a_m \in \mathbb{R}^n \setminus \{0\}$ mit $A = (a_1, \dots, a_m)^T$, $\beta_1, \dots, \beta_m \in \mathbb{R}$ mit $b = (\beta_1, \dots, \beta_m)^T$, $P := \{x \in \mathbb{R}^n : Ax \leq b\}$, $\gamma_1, \dots, \gamma_n \in \mathbb{R}$ mit $c = (\gamma_1, \dots, \gamma_n)^T$ und $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ bezeichne die durch $\varphi(x) := c^T x$ für alle $x \in \mathbb{R}^n$ definierte Zielfunktion. Sei ferner v eine Ecke von P .

3.2.2 Beispiel. Seien $n \geq 2$ und

$$C := \bigcap_{i=1}^{n-1} \{x = (\xi_1, \dots, \xi_n)^T : \pm \xi_i - \xi_n \leq 0\}.$$

Für jeden Punkt $x \in C$ gilt $u_n^T x \geq 0$ mit Gleichheit genau für $x = 0$. Genauer ist C der vom $(n-1)$ -dimensionalen Würfel $[-1,1]^{n-1} \times \{1\}$ erzeugte Kegel, d.h.

$$C = \text{pos}([-1,1]^{n-1} \times \{1\}).$$

C besitzt 2^{n-1} Kanten, die von seiner Ecke 0 ausgehen.

3.2.3 Bemerkung. Seien $d := (c^T, 0)^T$ und

$$Q := \left\{ z := \begin{pmatrix} x \\ \xi_{n+1} \end{pmatrix} \in \mathbb{R}^{n+1} : Ax - (b - Av)\xi_{n+1} \leq 0 \wedge 0 \leq \xi_{n+1} \leq 1 \right\}.$$

Dann ist 0 Ecke von Q und genau dann Optimalpunkt in Q bez. der durch $z \mapsto d^T z$ gegebenen Zielfunktion, wenn v Optimalpunkt von P bez. φ ist. Andernfalls gibt es eine Verbesserungskante K von Q in 0 mit

$$\max_{x \in P} c^T x = c^T v + \max_{z \in K} d^T z = c^T v + \max_{z \in Q} d^T z.$$

Beweis: Dass 0 Ecke von Q ist, folgt aus Lemma 3.1.10. Ist w eine Ecke von P , so ist $[0,1](w^T, 1)^T$ eine Kante von Q . Umgekehrt führt jede Kante von Q mit Endpunkt 0 , die nicht in $H_{(u_{n+1}, 0)}$ enthalten ist, zu einer Ecke von $P \times \{1\}$. Seien $s \in \mathbb{R}^n \setminus \{0\}$, und $z := (s^T, 0)^T$ liege in einer Kante von Q liegt. Dann gilt $[0, \infty[z \subset Q$ und $s \in \text{ls}(P)$.

Andererseits ist jeder Punkt $z \in Q \cap H_{(u_{n+1}, 0)}^>$ von der Form $\lambda(x^T, 1)^T$ mit $x \in -v + P$ und $0 < \lambda \leq 1$, und es folgt für $c^T x \geq 0$

$$d^T z = \lambda c^T x \leq c^T x.$$

Aus

$$\max_{x \in -v + P} c^T x = \max_{z \in (-v + P) \times \{1\}} d^T z \leq \max_{z \in Q} d^T z$$

folgt nun die Behauptung. \square

3.2.4 Definition. Sei C ein polyedrischer Kegel im \mathbb{R}^n . C heißt **simplizial** [engl.: simplicial], wenn C eine \mathcal{H} -Darstellung $(n, n, T, 0)$ besitzt mit $\text{rang}(T) = n$.

3.2.5 Bemerkung. Sei C ein simplizialer Kegel im \mathbb{R}^n . Dann ist 0 eine reguläre Ecke von C und C besitzt genau n Kanten.

3.2.6 Definition. Seien $B := B(v) \subset I(v)$ und $N := M \setminus B$. B heißt genau dann **Basis** [engl.: basis] von v , wenn $|B| = n$ gilt und $\{a_i : i \in B\}$ linear unabhängig ist.

Sei B eine Basis von v . Dann heißt A_B (zu B gehörige) **Basisteilmatrix** [engl.: basis submatrix] von A .⁶ Die Ecke v wird auch **zulässige Basislösung** [engl.: basic feasible solution] des linearen Ungleichungssystems $Ax \leq b$ genannt. Die zu B gehörigen Restriktionen $a_i^T x \leq \beta_i$ ($i \in B$) heißen **Basisrestriktionen** [engl.: basis constraints], die zu N gehörigen **Nichtbasisrestriktionen** [engl.: non basis constraints]. Das Gleichungssystem $A_B x = b_B$ heißt **Basisgleichung** [engl.: basis equation].

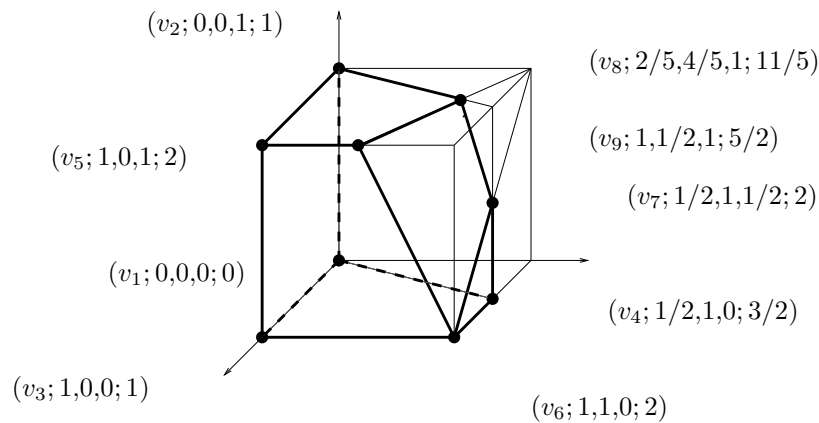
⁶ Bisweilen werden B , $\{a_i : i \in B\}$ und A_B identifiziert, und man spricht etwa auch von $\{a_i : i \in B\}$ als Basis von v .

3.2.7 Bemerkung. Sei B eine Basis von v . Dann gilt $v = A_B^{-1}b_B$.

3.2.8 Beispiel. Wir betrachten erneut das Produktionsproblem von Beispiel 1.2.2.

$$\begin{aligned} \max \quad & \xi_1 + \xi_2 + \xi_3 \\ \xi_1 + 2\xi_2 + \xi_3 & \leq 3 & (1) \\ -2\xi_1 + \xi_2 & \leq 0 & (2) \\ \xi_1 & \leq 1 & (3) \\ \xi_2 & \leq 1 & (4) \\ \xi_3 & \leq 1 & (5) \\ -\xi_1 & \leq 0 & (6) \\ -\xi_2 & \leq 0 & (7) \\ -\xi_3 & \leq 0 & (8) \end{aligned}$$

Der zulässige Bereich ist ein Polytop mit 9 Ecken v_1, \dots, v_9 und 7 Facetten. Die Ungleichung (6) ist redundant.



3.2 Abbildung Zulässiger Bereich der Produktionsaufgabe.

In Abbildung 3.2 werden für $k = 1, \dots, 9$ die Koordinaten der Ecke v_k und die zugehörigen Zielfunktionswerte in der Form $(v_k; \nu_{k1}, \nu_{k2}, \nu_{k3}; \gamma_k)$ angegeben, d.h.

$$v_k := (\nu_{k1}, \nu_{k2}, \nu_{k3})^T \quad \wedge \quad \gamma_k := c^T v_k.$$

Die Ecke v_9 etwa ist durch die Bedingungen (1), (3) und (5) gegeben. Wir haben also

$$B = \{1, 3, 5\} \quad \wedge \quad A_B = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \wedge \quad b_B = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}$$

und

$$A_B^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & -1/2 & -1/2 \\ 0 & 0 & 1 \end{pmatrix} \quad \wedge \quad v_9 = A_B^{-1}b_B = \begin{pmatrix} 1 \\ 1/2 \\ 1 \end{pmatrix}.$$

Zu v_9 gehört eine eindeutig bestimmte Basis; die Ecke v_6 hingegen kann man (wie wir bereits in Beispiel 2.3.26 erkannt hatten) auf verschiedene Weisen als Durchschnitt von drei Restriktionsgleichungen darstellen, da für sie mehr als drei der Restriktionen aktiv sind. Genauer gilt $I(v_6) = \{1,3,4,8\}$. Da alle entsprechenden Teilmatrizen regulär sind, können der Ecke v_6 vier verschiedene Basisgleichungen zugeordnet werden.

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} x = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}; \quad \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} x = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix};$$

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} x = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}; \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

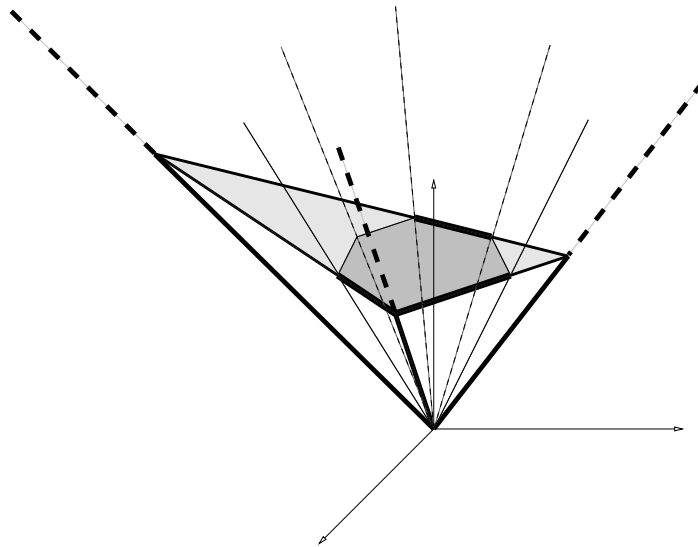
Auch die Ecke v_1 ist überbestimmt; die Bedingungen (2), (6), (7) und (8) sind mit Gleichheit erfüllt. Während Ecke v_6 in vier Facetten von P liegt, ist v_1 aber nur in drei Facetten enthalten. Die Überbestimmtheit einer Ecke ist also eine Eigenschaft der gegebenen \mathcal{H} -Darstellung, nicht notwendigerweise der zugrunde liegenden Geometrie.

3.2.9 Bezeichnung. Seien B eine Basis von v und

$$S_P(B) := S_P(B(v)) := \{x \in \mathbb{R}^n : A_B x \leq 0\}$$

$$N_P(B) := N_P(B(v)) := \text{pos}(\{a_i : i \in B\}).$$

$S_P(B)$ heißt **Basisstützkegel** und $N_P(B)$ **Basisnormalenkegel** [engl.: *basic support (normal) cone*] zur Basis B von v . Zur Abkürzung wird oft auch nur $S(B)$ und $N(B)$ geschrieben.



3.3 Abbildung Approximation von $S_P(v)$ durch einen (simplicialen) Basisstützkegel.

3.2.10 Bemerkung. Sei B eine Basis von v . Dann gilt

$$S_P(v) \subset S_P(B) \quad \wedge \quad N_P(B) \subset N_P(v)$$

mit Gleichheit für $|I(v)| = n$.

3.2.11 Bemerkung. Seien $i_1, \dots, i_n \in M$ mit $i_1 < \dots < i_n$ und $B = \{i_1, \dots, i_n\}$ sowie $s_1, \dots, s_n \in \mathbb{R}^n$ mit $-A_B^{-1} = (s_1, \dots, s_n)$.

(a) Die Kanten von $S(B)$ sind $[0, \infty[s_1, \dots, [0, \infty[s_n$.

(b) Für $i = 1, \dots, n$ ist $\mathbb{R}s_i$ der Lösungsraum von $A_{B \setminus \{i\}}x = 0$.

Beweis: (a) Wir benutzen die Tatsache, dass der Kegel $S(B)$ durch Anwendung der Koordinatentransformation $x \mapsto A_B x$ in den negativen Orthanten $] -\infty, 0]^n$ übergeht, dessen Kanten die nichtpositiven Koordinatenachsen sind. Genauer gilt

$$\begin{aligned} S(B) &= A_B^{-1} \{A_B x \in \mathbb{R}^n : A_B x \leq 0\} \\ &= A_B^{-1} \{y \in \mathbb{R}^n : y \leq 0\} = A_B^{-1} \text{pos}(\{-u_1, \dots, -u_n\}) \\ &= \text{pos}(\{-A_B^{-1}u_1, \dots, -A_B^{-1}u_n\}). \end{aligned}$$

Die Kanten von $S(B)$ werden also von den Spaltenvektoren von $-A_B^{-1}$ erzeugt.

(b) Für $i = 1, \dots, n$ folgt aus $A_B A_B^{-1} = E_n$

$$A_B s_i = -u_i$$

und damit die Behauptung. □

3.2.12 Bemerkung. (a) Gilt $(-A_B^{-1})^T c \leq 0$, so ist v Optimalpunkt von φ über P .

(b) Gilt $(-A_B^{-1})^T c < 0$, so ist v einziger Optimalpunkt von φ über P .

3.2.13 Beispiel. (Fortsetzung von Beispiel 3.2.8). Wie in Beispiel 3.2.8 berechnet wurde, gilt

$$B = I(v_9) = \{1, 3, 5\} \quad \wedge \quad A_B^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & -1/2 & -1/2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Die von v_9 ausgehenden Kanten haben die Richtungen

$$\begin{pmatrix} 0 \\ -1/2 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 1/2 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1/2 \\ -1 \end{pmatrix}.$$

Ferner gilt

$$(-A_B^{-1})^T c = \begin{pmatrix} 0 & -1/2 & 0 \\ -1 & 1/2 & 0 \\ 0 & 1/2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/2 \\ -1/2 \\ -1/2 \end{pmatrix}.$$

Wegen $A_B^{-1}c < 0$ ist v_9 nach Bemerkung 3.2.12 der eindeutig bestimmte Optimalpunkt der gegebenen Aufgabe.

3.2.14 Lemma. Seien S ein Simplex $[0, \infty[s$ eine Kante von $S(B)$ und

$$c^T s > 0 \quad \wedge \quad S := v + [0, \infty[s \quad \wedge \quad R := \{i \in N : a_i^T s > 0\}.$$

(a) Ist $R = \emptyset$, so gilt $S \subset P$, d.h. φ ist über P unbeschränkt.

(b) Seien

$$R \neq \emptyset \quad \wedge \quad \lambda := \min \left\{ \frac{\beta_i - a_i^T v}{a_i^T s} : i \in R \right\} \quad \wedge \quad w := v + \lambda s.$$

Dann ist w eine Ecke von P .

Beweis: Gilt $a_i^T s \leq 0$, so folgt für $\mu \in [0, \infty[$

$$a_i^T (v + \mu s) \leq a_i^T v.$$

(a) Aus $R = \emptyset$ folgt $a_i^T s \leq 0$ für alle $i \in \{1, \dots, n\}$, und damit

$$S \subset H_{(a_i, \beta_i)}^{\leq}.$$

Aus $c^T s > 0$ folgt somit (a).

(b) Für $i \in R$ gilt

$$a_i^T (v + \lambda s) \leq a_i^T v + \frac{\beta_i - a_i^T v}{a_i^T s} a_i^T s = \beta_i.$$

Also gilt $w \in P$. Für $w = v$ ist nichts zu zeigen. Seien also $w \neq v$ und $K := \text{conv}(\{v, w\})$. Sei i_0 ein Index, für den bei der Bestimmung von λ das Minimum angenommen wird. Dann ist $v + [0, \infty[s$ nicht in $H_{(a_{i_0}, \beta_{i_0})}$ enthalten, d.h.

$$\{w\} = S \cap H_{(a_{i_0}, \beta_{i_0})}.$$

$H_{(a_{i_0}, \beta_{i_0})}$ ist somit eigentliche Stützhyperebene an K in w , und die Behauptung folgt aus Lemma 2.3.6. \square

3.2.15 Prozedur: *Simplex-Algorithmus (Verbesserungsphase (Grundform)).*

```

INPUT:    $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $c \in \mathbb{R}^n$ 
         Ecke  $v$  von  $P = \{x : Ax \leq b\}$ 
         Basis  $B = \{i_1, \dots, i_m\}$  von  $v$  mit  $i_1 < \dots < i_m$ 
OUTPUT:  Meldung 'φ ist auf  $P$  nach oben unbeschränkt!'
         OR optimale Ecke  $v^*$ 
BEGIN    fertig ← 'nein'
         WHILE fertig = 'nein' DO
           BEGIN
              $N \leftarrow \{1, \dots, m\} \setminus B$ ,
             Berechne die Spaltenvektoren  $s_1, \dots, s_n$  von  $-A_B^{-1}$ 
             IF  $c^T s_j \leq 0$  für alle  $j = 1, \dots, n$ 
               THEN Meldung ' $v$  ist optimal!' AND fertig ← 'ja'
             ELSE
               BEGIN
                 Wähle  $j_0 \in \{1, \dots, n\}$  mit  $c^T s_{j_0} > 0$ 
                 AND  $R \leftarrow \{i \in N : a_i^T s_{j_0} > 0\}$ 
                 IF  $R = \emptyset$ 
                   THEN Meldung ' $\varphi$  nach oben unbeschränkt!'
                   AND fertig ← 'ja'
                 ELSE
                   BEGIN
                      $\lambda \leftarrow \min \left\{ \frac{\beta_i - a_i^T v}{a_i^T s_{j_0}} : i \in R \right\}$ 
                     AND wähle  $i_{\text{rein}} \in R$  mit  $\lambda = \frac{\beta_{i_{\text{rein}}} - a_{i_{\text{rein}}}^T v}{a_{i_{\text{rein}}}^T s_{j_0}}$ 
                     AND  $i_{\text{raus}} \leftarrow i_{j_0}$ 
                     AND  $B \leftarrow (B \setminus \{i_{\text{raus}}\}) \cup \{i_{\text{rein}}\}$ 
                     AND  $v \leftarrow v + \lambda s_{j_0}$ 
                   END
                 END
               END
           END
         END
END

```

3.2.16 Bezeichnung. *Prozedur 3.2.15 wird **Phase II** [engl.: phase II] des Simplex-Algorithmus genannt. Die Anwendung von Prozedur 3.2.15 zur Entscheidung, ob $P \neq \emptyset$ ist, bzw. der Bestimmung einer Startecke und -basis heißt dann **Phase I** [engl.: phase I] des Simplex-Algorithmus.*

*Jede Regel zur Wahl von $(i_{\text{raus}}, i_{\text{rein}})$ (unter den in Prozedur 3.2.15 angegebenen Möglichkeiten) heißt **Pivotregel** [engl.: pivot rule]; das ausgewählte Paar $(i_{\text{raus}}, i_{\text{rein}})$ heißt **Pivotpaar** [engl.: pivot pair].*

3.2.17 Beispiel. *(Fortsetzung von Beispiel 3.2.8 und 3.2.13) Wir wählen $v_3 = (1, 0, 0)^T$ als Startecke. Die hierzu eindeutig bestimmte Basis, Basisteilmatrix und zugehörige rechte Seite sind*

$$B = \{3, 7, 8\} \quad \wedge \quad A_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \wedge \quad b_B = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

und es gilt $S(B) = S_P(v_3)$. Von v_3 gehen in $v_3 + S(B)$ die drei Kanten

$$v_3 + [0, \infty[(-u_1) \quad \wedge \quad v_3 + [0, \infty[u_2 \quad \wedge \quad v_3 + [0, \infty[u_3$$

aus. Nur die letzten beiden Kanten führen potentiell zu Verbesserungen des Zielfunktionswertes. Wir wählen etwa

$$v_3 + [0, \infty[u_2.$$

Man beachte, dass die von der zweiten Spalte von $-A_B^{-1}$ erzeugte Gerade Lösung des durch Streichen der zweiten Restriktion von $A_B x = b_B$ entstehenden Gleichungssystems ist, und diese ist Restriktion (7), d.h. $i_{\text{raus}} = 7$. Weiter gilt $R = \{i \in N : a_i^T u_2 > 0\} = \{1, 2, 4\}$, und es folgt

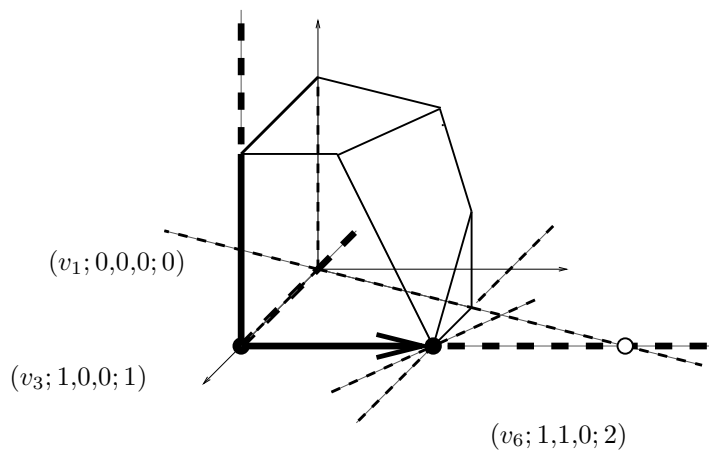
$$\lambda = \min \left\{ \frac{\beta_i - a_i^T v_3}{a_i^T u_2} : i = 1, 2, 4 \right\} = \min \left\{ \frac{3-1}{2}, \frac{0+2}{1}, \frac{1-0}{1} \right\} = \min\{1, 2, 1\} = 1.$$

Das Minimum wird durch die Restriktionen (1) und (4), d.h. durch

$$\xi_1 + 2\xi_2 + \xi_3 \leq 3 \quad \wedge \quad \xi_2 \leq 1$$

bestimmt. Hieran sieht man bereits, dass die neue Ecke $w = v_3 + \lambda u_2 = v_6$ überbestimmt ist. Es gilt $I(v_6) = \{1, 3, 4, 8\}$, wobei die neue Basis 3 und 8 enthält, und die Wahl zwischen 1 und 4 als drittem Element besteht. Man kann sich somit zwischen den folgenden Basisgleichungen für v_6 entscheiden:

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} x = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} \quad \wedge \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$



3.4 Abbildung Der Kegel $S(B(v_3))$ und der Übergang von v_3 zu v_6 .

Wählen wir etwa $i_{\text{rein}} = 4$, so erhalten wir die neue Basis

$$B = B^1(v_6) = \{3, 4, 8\}.$$

Der Kegel $S(B)$ hat die drei Kanten

$$[0, \infty[(-u_1) \quad \wedge \quad [0, \infty[(-u_2) \quad \wedge \quad [0, \infty[u_3,$$

von denen nur die letzte zur Verbesserung der Zielfunktion führt. Wir wählen also $i_{\text{raus}} = 8$, d.h. $-\xi_3 \leq 0$ wird zu einer Nichtbasisrestriktion. Es gilt $R = \{1, 5\}$, also

$$\lambda = \min\{0,1\} = 0,$$

d.h. $i_{\text{rein}} = 1$. Hier tritt der Fall ein, dass wir bei v_6 bleiben, aber zu einer neuen Basis übergehen:

$$B = B^2(v_6) = \{1,3,4\} \quad \wedge \quad A_B = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \wedge \quad b_B = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}.$$

Es gilt

$$A_B^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & -2 \end{pmatrix},$$

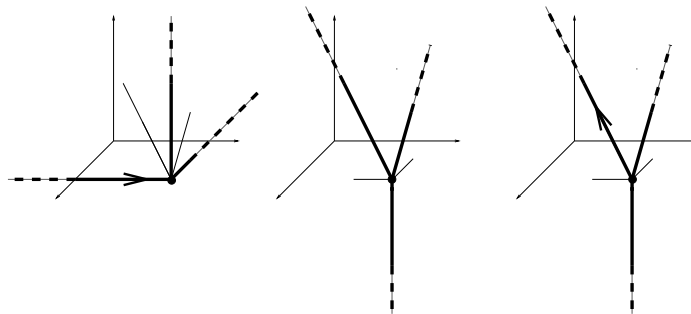
d.h. die Kanten sind

$$v_6 + [0, \infty[(-u_3) \quad \wedge \quad v_6 + [0, \infty[(-1, 0, 1)^T \quad \wedge \quad v_6 + [0, \infty[(0, -1, 2)^T.$$

Die letzte dieser Kante ist zu wählen, d.h. $i_{\text{raus}} = 4$; die Bedingung $\xi_2 \leq 1$ wird inaktiv. Es gilt $R = \{5,7\}$, es folgt

$$\lambda = \min\left\{\frac{1}{2}, 1\right\} = \frac{1}{2},$$

also $i_{\text{rein}} = 5$, und wir erhalten die neue Ecke v_9 zusammen mit der Basis $B = B(v_9) = \{1,3,5\}$. Wie wir bereits in Beispiel 3.2.13 gesehen hatten, ist keine weitere Verbesserung



3.5 Abbildung Übergang zu v_6 , anschließender Basiswechsel und Verlassen von v_6 in Richtung v_9 .

der Zielfunktion mehr möglich; v_9 ist optimal.

3.3 Über Zykel und ihre Vermeidung

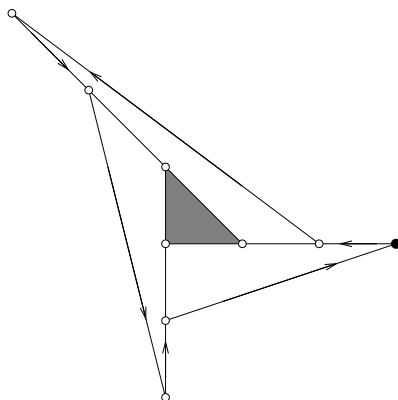
3.3.1 Bezeichnung. Tritt bei der Durchführung des Simplex-Algorithmus für eine gegebene LP-Aufgabe eine vorher berechnete Basis einer Ecke ein zweites Mal auf, so spricht man von einem **Zyklus** [engl.: cycle] (oder davon, dass der Simplex-Algorithmus **zykelt**).

3.3.2 Satz. Bei entsprechender Wahl der Pivotpaare kann der Simplex-Algorithmus **zykeln**.

Beweis: Wir geben eine LP-Aufgabe im \mathbb{R}^3 mit 6 bzw. (im beschränkten Fall) 7 Restriktionen an, für die der Simplex-Algorithmus bei ungünstiger Wahl der Pivotpaare zyckelt:

$$\begin{array}{rcccccl} \max & \xi_3 & & & & \\ -\xi_1 & & & & \leq & 0 & (1) \\ & -\xi_2 & & & \leq & 0 & (2) \\ \xi_1 & + & \xi_2 & - & \xi_3 & \leq & 0 & (3) \\ -4\xi_1 & - & \xi_2 & - & 2\xi_3 & \leq & 0 & (4) \\ \xi_1 & - & 3\xi_2 & - & 3\xi_3 & \leq & 0 & (5) \\ 3\xi_1 & + & 4\xi_2 & - & 6\xi_3 & \leq & 0 & (6) \\ & & & & \xi_3 & \leq & 1 & (7). \end{array}$$

Wir betrachten $v_1 = 0$, d.h. relevant sind nur die 6 Restriktionen (1)-(6). Tatsächlich kann (7) für die gesamte Argumentation weggelassen werden. Ohne Bedingung (7) ist der zulässige Bereich ein Kegel und die Zielfunktion unbeschränkt, ohne dass der Simplex-Algorithmus das merkt. Mit Restriktion (7) zeigt die Konstruktion, dass der Simplex-Algorithmus das Optimum nicht findet. In beiden Fällen verlässt er die Ecke v_1 nicht.



3.6 Abbildung Schnitt der Restriktionshyperebenen mit $H_{(u_3,1)}$; Startkante (schwarzer Punkt), in den Simplex-Schritten ausgewählte Kantenfolge.

Startbasis und zugehörige inverse Basismatrix sind

$$B^1 := \{1,2,5\} \quad \wedge \quad A_{B^1}^{-1} = \frac{1}{3} \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ -1 & 3 & -1 \end{pmatrix}.$$

Nach Prozedur 3.2.15 ist eine Verbesserungskante s_j zu wählen mit $c^T s_j > 0$, und nach Bemerkung 3.2.11 kommen hierfür genau die Kantenrichtungen

$$\begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

in Frage. Wir nehmen die erste und erhalten bzw. wählen

$$s_1 := \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \wedge i_{\text{raus}} := 1 \wedge R_1 := \{3,6\} \wedge i_{\text{rein}} := 6.$$

Man beachte, dass die Schrittweite λ hier und im folgenden stets 0 ist. Die weiteren Schritte sind wie folgt.

$$B^2 := \{2,5,6\} \wedge A_{B^2}^{-1} = \frac{1}{3} \begin{pmatrix} 30 & -6 & 3 \\ -3 & 0 & 0 \\ 13 & -3 & 1 \end{pmatrix} \wedge s_2 := \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \\ \wedge i_{\text{raus}} := 5 \wedge R_2 := \{3\} \wedge i_{\text{rein}} := 3;$$

$$B^3 := \{2,3,6\} \wedge A_{B^3}^{-1} = \frac{1}{3} \begin{pmatrix} 2 & 6 & -1 \\ -3 & 0 & 0 \\ -1 & 3 & -1 \end{pmatrix} \wedge s_3 := \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix} \\ \wedge i_{\text{raus}} := 2 \wedge R_3 := \{1,4\} \wedge i_{\text{rein}} := 4;$$

$$B^4 := \{3,4,6\} \wedge A_{B^4}^{-1} = \frac{1}{3} \begin{pmatrix} -14 & -2 & 3 \\ 30 & 3 & -6 \\ 13 & 1 & -3 \end{pmatrix} \wedge s_4 := \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \\ \wedge i_{\text{raus}} := 6 \wedge R_4 := \{1\} \wedge i_{\text{rein}} := 1;$$

$$B^5 := \{1,3,4\} \wedge A_{B^5}^{-1} = \frac{1}{3} \begin{pmatrix} -3 & 0 & 0 \\ 6 & 2 & -1 \\ 3 & -1 & -1 \end{pmatrix} \wedge s_5 := \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \\ \wedge i_{\text{raus}} := 3 \wedge R_5 := \{2,5\} \wedge i_{\text{rein}} := 5;$$

$$B^6 := \{1,4,5\} \wedge A_{B^6}^{-1} = \frac{1}{3} \begin{pmatrix} -3 & 0 & 0 \\ -14 & 3 & -2 \\ 13 & -3 & 1 \end{pmatrix} \wedge s_6 := \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \\ \wedge i_{\text{raus}} := 4 \wedge R_6 := \{2\} \wedge i_{\text{rein}} := 2.$$

Danach wiederholt sich B^1 ; der Zyklus ist also komplett. \square

3.3.3 Definition. Seien $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$. Das Paar (oder System) (A, b) heißt **generisch** [engl.: generic] oder **in allgemeiner Lage** [engl.: in general position], wenn es keine Teilmenge I von $\{1, \dots, m\}$ gibt mit $|I| = n + 1$, so dass $A_I x = b_I$ lösbar ist.

3.3.4 Lemma. Seien $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, und für $\varepsilon \in]0, \infty[$ sei $e(\varepsilon) := (\varepsilon, \varepsilon^2, \dots, \varepsilon^m)^T$. Dann ist $(A, b + e(\varepsilon))$ generisch für alle bis auf endliche viele positive ε .

Beweis: Sei $I \subset \{1, \dots, m\}$ mit $|I| = n + 1$. Das Gleichungssystem

$$A_I x = (b + e(\varepsilon))_I$$

ist genau dann lösbar, wenn

$$\text{rang}(A_I) = \text{rang}\left(A_I, (b + e(\varepsilon))_I\right)$$

gilt. Seien $k := \text{rang}(A_I)$, $J \subset I$ mit $|J| = k + 1$, und \bar{A}_J eine $(k + 1) \times k$ Teilmatrix von A_J vom Rang k . Dann ist $(\bar{A}_J, b_J + e(\varepsilon)_J)$ eine $(k + 1) \times (k + 1)$ Matrix, und wir zeigen, dass $\det(\bar{A}_J, b_J + e(\varepsilon)_J)$ für höchstens endlich viele Werte von ε verschwindet.

Durch Entwicklung der Determinante nach der letzten Spalte gemäß dem Entwicklungssatz von Laplace erhält man eine Darstellung

$$\det(\bar{A}_J, b_J + e(\varepsilon)_J) = \det(\bar{A}_J, b_J) + \det(\bar{A}_J, e(\varepsilon)_J) = \det(\bar{A}_J, b_J) + \sum_{i \in J} \delta_i \varepsilon^i;$$

die Koeffizienten δ_i sind bis auf das Vorzeichen die durch Streichen einer Zeile entstehenden $k \times k$ Subdeterminanten von \bar{A}_J . Da $\text{rang}(\bar{A}_J) = k$ ist, muss mindestens ein δ_i von 0 verschieden sein. Das durch

$$\varepsilon \mapsto \det(\bar{A}_J, b_J + e(\varepsilon)_J)$$

definierte Polynom in ε ist also nicht identisch 0. Sein Grad ist höchstens m ; somit hat es höchstens m Nullstellen. Da andererseits die Anzahl der verschiedenen $(n + 1) \times n$ Teilmatrizen von A durch $\binom{m}{n+1}$ beschränkt ist, folgt die Behauptung. \square

3.3.5 Bezeichnung. Seien $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ und $\varepsilon \in]0, \infty[$. Ferner seien

$$e(\varepsilon) := (\varepsilon, \varepsilon^2, \dots, \varepsilon^m)^T \in \mathbb{R}^m \quad \wedge \quad P(\varepsilon) = \{x \in \mathbb{R}^n : Ax \leq b + e(\varepsilon)\}.$$

$e(\varepsilon)$ heißt **Perturbationsvektor** [engl.: *perturbation vector*], und $P(\varepsilon)$ wird **perturbiertes Polyeder** [engl.: *perturbed polyhedra*] genannt.

3.3.6 Korollar. Sei $P = \{x \in \mathbb{R}^n : Ax \leq b\}$. Dann gibt es ein $\varepsilon_0 \in]0, \infty[$, so dass für jedes $\varepsilon \in]0, \varepsilon_0]$ kein Punkt des \mathbb{R}^n in mehr als n der Restriktionshyperebenen von $Ax \leq b + e(\varepsilon)$ enthalten ist, d.h. insbesondere, dass jede Ecke des perturbierten Polyeders $P(\varepsilon)$ regulär ist.

3.3.7 Bemerkung. Durch (n, m, A, b, c) sei eine LP-Aufgabe (in natürlicher Form) spezifiziert. Dann gibt es ein $\varepsilon_0 \in]0, \infty[$, so dass die folgende Aussage gilt. Für jedes $\varepsilon \in]0, \varepsilon_0]$ und jede Startecke $v(\varepsilon)$ löst der Simplex-Algorithmus die durch $(m, n, A, b + e(\varepsilon), c)$ gegebene LP-Aufgabe in höchstens

$$\binom{m}{n} - 1$$

vielen Schritten.

3.3.8 Notation. Die gegebene LP-Aufgabe in natürlicher Form sei spezifiziert durch die Daten (n, m, A, b, c) . Es seien ferner $M := \{1, \dots, m\}$, $a_1, \dots, a_m \in \mathbb{R}^n \setminus \{0\}$ mit $A = (a_1, \dots, a_m)^T$, $\beta_1, \dots, \beta_m \in \mathbb{R}$ mit $b = (\beta_1, \dots, \beta_m)^T$, $P := \{x \in \mathbb{R}^n : Ax \leq b\}$, und $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ bezeichne die durch $\varphi(x) := c^T x$ definierte Zielfunktion.

Ferner seien $\varepsilon_0 \in]0, \infty[$ gemäß Korollar 3.3.6 gewählt, und alle im folgenden auftretenden Störungsparameter ε sind auf das Intervall $]0, \varepsilon_0]$ beschränkt. Seien $e := e(\varepsilon)$, $v(\varepsilon)$ eine (im Laufe des Algorithmus erreichte) Ecke von $P(\varepsilon)$, $B := \{i_1, \dots, i_n\}$ mit $i_1 < \dots < i_n$ ihre (eindeutig bestimmte) Basis und $N := M \setminus B$.

3.3.9 Lemma. Sei $v := A_B^{-1}b_B$. Dann ist v eine Ecke von P , und die Kanten von $S_{P(\varepsilon)}(v(\varepsilon))$ und $S_P(B)$ stimmen überein.

Beweis: Sei zunächst $i \in B$. Dann gilt

$$a_i^T v(\varepsilon) = a_i^T A_B^{-1}(b_B + e(\varepsilon)_B) = u_i^T(b_B + e(\varepsilon)_B) = \beta_i + \varepsilon^i$$

unabhängig von der speziellen Wahl von $\varepsilon \in]0, \varepsilon_0]$. Nach Konstruktion (oder, wenn man möchte, aus Stetigkeitsgründen) gilt auch $a_i^T v = \beta_i$.

Wir zeigen nun mittels Widerspruchsbeweis, dass $a_i^T v \leq \beta_i$ auch für $i \in N$ gilt. Sei daher $i_0 \in N$ mit $a_{i_0}^T v > \beta_{i_0}$. Da $v(\varepsilon)$ eine Ecke von $P(\varepsilon)$ und regulär ist, gilt

$$a_{i_0}^T v(\varepsilon) < \beta_{i_0} + \varepsilon^{i_0} \quad \wedge \quad v(\varepsilon) = A_B^{-1}b_B + A_B^{-1}e_B = v + A_B^{-1}e_B.$$

Es folgt

$$a_{i_0}^T v < \beta_{i_0} + \varepsilon^{i_0} - a_{i_0}^T A_B^{-1}e_B.$$

Als Funktional in ε ist die rechte Seite stetig und strebt für $\varepsilon \rightarrow 0$ gegen β_{i_0} . Nach dem Zwischenwertsatz gibt es ein $\varepsilon_1 \in]0, \varepsilon_0[$ mit

$$a_{i_0}^T v = \beta_{i_0} + \varepsilon^{i_0} - a_{i_0}^T A_B^{-1}e_B.$$

Somit gilt

$$a_{i_0}^T v(\varepsilon) = \beta_{i_0} + \varepsilon^{i_0}$$

und damit

$$n + 1 \leq |B \cup \{i_0\}| \leq |I(v(\varepsilon_1))|,$$

im Widerspruch zur Generizität, d.h. zur Wahl von ε_0 . Also gilt auch $A_N v \leq b_N$, d.h. v ist eine zulässige Basislösung von $Ax \leq b$.

Da $v(\varepsilon)$ regulär ist, gilt $S_{P(\varepsilon)}(B) = S_{P(\varepsilon)}(v(\varepsilon))$. Nach Bemerkung 3.2.11 sind die Kanten von $S_{P(\varepsilon)}(v(\varepsilon))$ daher die positiven Hüllen der Spaltenvektoren von $-A_B^{-1}$, ebenso wie die von $S_P(B)$. \square

3.3.10 Bezeichnung. Seien v eine Ecke von P und B eine zugehörige Basis. B heißt $e(\varepsilon)$ -kompatibel, falls $v + A_B^{-1}e_B(\varepsilon_0) \in P(\varepsilon_0)$ gilt.

3.3.11 Bemerkung. Seien v eine Ecke von P , B eine Basis von v und $\varepsilon \in]0, \varepsilon_0]$.

- (a) B ist genau dann $e(\varepsilon)$ -kompatibel, wenn $v + A_B^{-1}e_B(\varepsilon) \in P(\varepsilon)$ gilt.
- (b) B ist genau dann $e(\varepsilon)$ -kompatibel, wenn $\varepsilon^i - a_i^T A_B^{-1}e_B > 0$ für alle $i \in I_P(v) \setminus B$ gilt.

Beweis: Man beachte zunächst, dass für $i \in B$ für alle $\varepsilon \in]0, \varepsilon_0]$

$$\varepsilon^i - a_i^T A_B^{-1}e_B = 0.$$

gilt.

- (a) folgt nun aus der Generizität von $(A, b + e_B)$.
- (b) '⇒' Seien B $e(\varepsilon)$ -kompatibel und $i \in I_P(v) \setminus B$. Dann gilt

$$a_i^T v(\varepsilon) = a_i^T v + a_i^T A_B^{-1}e_B = \beta_i + a_i^T A_B^{-1}e_B < \beta_i + \varepsilon^i$$

und damit $\varepsilon^i - a_i^T A_B^{-1} e_B > 0$

‘ \Leftarrow ’ Für $i \in N \setminus I_P(v)$ gilt $a_i^T v < \beta_i$, also auch

$$a_i^T v(\varepsilon) = a_i^T v + a_i^T A_B^{-1} e_B \leq \beta_i + \varepsilon^i$$

für hinreichend kleines positives ε . Aus der Generizität von $(A, b + e(\varepsilon))$ folgt daher

$$a_i^T v(\varepsilon_0) < \beta_i + \varepsilon_0^i.$$

Für $i \in I_P(v) \setminus B$ gilt nach Voraussetzung $a_i^T A_B^{-1} e_B < \varepsilon^i$ und damit

$$a_i^T v(\varepsilon) = a_i^T v + a_i^T A_B^{-1} e_B < \beta_i + \varepsilon^i.$$

Insgesamt gilt somit $v(\varepsilon) \in P(\varepsilon)$, und aus (a) folgt die Behauptung. \square

3.3.12 Lemma. *Seien v eine Ecke von P und B eine Basis von v . Ferner seien die Ungleichungen so angeordnet, dass $B = \{m-n+1, \dots, m\}$ gilt. Dann ist B $e(\varepsilon)$ -kompatibel.*

Beweis: Sei $i \in I_P(v) \setminus B$. Dann gilt

$$i \leq m-n \quad \wedge \quad \varepsilon^i - a_i^T A_B^{-1} e_B = \varepsilon^i \left(1 - \varepsilon^{m-n-i} a_i^T A_B^{-1} (\varepsilon, \varepsilon^2, \dots, \varepsilon^n)^T\right).$$

Für hinreichend kleine positive ε gilt somit

$$\varepsilon^i - a_i^T A_B^{-1} e_B > 0.$$

Nach Bemerkung 3.3.11 (b) folgt die Behauptung. \square

3.3.13 Beispiel. *(Fortsetzung von Beispiel 3.2.8) Wir vollziehen den Schritt von v_3 zu v_6 aus Beispiel 3.2.8, jetzt aber durch Rechnung im perturbierten Polyeder. Da v_3 regulär ist, ist die Basis von v_3 $e(\varepsilon)$ -kompatibel, so dass wir starten können wie vorher.⁷ Es gilt $B = \{3, 7, 8\}$ und*

$$A_B = A_{B(v_3)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad b_B = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Als Verbesserungskante des zugehörigen Kegels $S(B)$ hatten wir $v_3 + [0, \infty[u_2$ gewählt. Es gilt

$$R = \{1, 2, 4\} \quad \wedge \quad e_B = (\varepsilon^3, \varepsilon^7, \varepsilon^8)^T \quad \wedge \quad v_3(\varepsilon) = v_3 + A_B^{-1} e_B = \begin{pmatrix} 1 + \varepsilon^3 \\ -\varepsilon^7 \\ -\varepsilon^8 \end{pmatrix}.$$

Für $i \in R$ haben wir somit

⁷ Es ist natürlich klar, dass der Simplex-Algorithmus auf der unperturbierten Aufgabe nicht nur nicht zyckelt, sondern auch nach nur einem Schritt v_6 wieder verlassen muss. Schließlich gibt es unter den sechs Kanten aller vier möglichen Basistützel nur eine Verbesserungskante, nämlich $v_6 + [0, \infty[u_3$, die aber höchstens einmal auftreten kann.

$$\frac{\beta_i + \varepsilon^i - a_i^T v_3(\varepsilon)}{a_i^T u_2} = \begin{cases} 1 + \frac{1}{2}\varepsilon^1 - \frac{1}{2}\varepsilon^3 + \varepsilon^7 + \frac{1}{2}\varepsilon^8 & \text{für } i = 1; \\ 2 + \varepsilon^2 + 2\varepsilon^3 + \varepsilon^7 & \text{für } i = 2; \\ 1 + \varepsilon^4 + \varepsilon^7 & \text{für } i = 4. \end{cases}$$

Für hinreichend kleines positives ε ist der entsprechende Quotient für $i = 4$ am kleinsten; wir nehmen also den Index 4 in die Basis auf. Das hatten wir 'vorausschauend' vorher auch bereits getan; mit der auf Perturbation beruhenden Regel ist die Setzung $i_{\text{rein}} = 4$ jetzt aber zwingend.

3.3.14 Notation. (a) Für $i \in M$ seien $\eta_i :]0, \varepsilon_0] \rightarrow \mathbb{R}$ das durch

$$\eta_i(\varepsilon) := \sum_{j=1}^m \eta_{i,j} \varepsilon^j := \varepsilon^i - a_i^T A_B^{-1} e_B$$

definierte Polynom und $y_i := (\eta_{i,1}, \dots, \eta_{i,m})^T \in \mathbb{R}^m$ sein Koeffizientenvektor.

(b) Seien s ein Spaltenvektor von $-A_B^{-1}$ mit $c^T s > 0$ und $R := \{i \in N : a_i^T s > 0\}$. Für $i \in R$ seien $\zeta_i :]0, \varepsilon_0] \rightarrow \mathbb{R}$ das durch

$$\zeta_i(\varepsilon) := \sum_{j=0}^m \zeta_{i,j} \varepsilon^j := \frac{\beta_i + \varepsilon^i - a_i^T v(\varepsilon)}{a_i^T s}$$

definierte Polynom in ε und $z_i := (\zeta_{i,0}, \dots, \zeta_{i,m})^T \in \mathbb{R}^{m+1}$ sein Koeffizientenvektor.

3.3.15 Bemerkung. (a) Sei $i \in M$. Dann gilt

$$\eta_{i,j} = \begin{cases} 0 & \text{für } i \in B; \\ 1 & \text{für } i \in N \wedge j = i; \\ -a_i^T A_B^{-1} u_k & \text{für } i \in N \wedge j = i_k \in B; \\ 0 & \text{für } i \in N \wedge j \in M \setminus (B \cup \{i\}). \end{cases}$$

(b) Sei $i \in R$. Dann gilt

$$\zeta_{i,j} = \frac{1}{a_i^T s} \begin{cases} \beta_i - a_i^T v & \text{für } j = 0; \\ 1 & \text{für } j = i; \\ -a_i^T A_B^{-1} u_k & \text{für } j = i_k \in B; \\ 0 & \text{für } j \in M \setminus (B \cup \{i\}). \end{cases}$$

Beweis: Es gilt

$$\eta_i(\varepsilon) = \sum_{j=1}^m \eta_{i,j} \varepsilon^j = \varepsilon^i - a_i^T A_B^{-1} (\varepsilon^{i_1}, \dots, \varepsilon^{i_n})^T = \varepsilon^i - \sum_{k=1}^n (a_i^T A_B^{-1} u_k) \varepsilon^{i_k}$$

sowie

$$a_i^T s \cdot \zeta_i(\varepsilon) = a_i^T s \cdot \sum_{j=0}^m \zeta_{i,j} \varepsilon^j = \beta_i - a_i^T v + \eta_i(\varepsilon) = \beta_i - a_i^T v + \varepsilon^i - \sum_{k=1}^n (a_i^T A_B^{-1} u_k) \varepsilon^{i_k}.$$

Durch Koeffizientenvergleich folgt hieraus die Behauptung. \square

3.3.16 Definition. (a) Seien $x = (\xi_1, \dots, \xi_m)^T \in \mathbb{R}^m \setminus \{0\}$ und $j_0 \in \{1, \dots, m\}$ mit

$$(j \in \{1, \dots, j_0 - 1\} \Rightarrow \xi_j = 0) \quad \wedge \quad \xi_{j_0} \neq 0.$$

Dann heißt x **lexikographisch positiv (negativ)** [engl.: lexicographically positive (negative)], genau dann, wenn gilt

$$\xi_{j_0} > 0 \quad (\xi_{j_0} < 0);$$

In diesem Fall schreiben wir $x \succ 0$ ($x \prec 0$).

(b) Seien $x_1, x_2 \in \mathbb{R}^m$. x_1 heißt **lexikographisch größer (kleiner)** [engl.: lexicographically greater (smaller)] als x_2 genau dann, wenn $x_1 - x_2 \succ 0$ ($x_1 - x_2 \prec 0$) ist; Schreibweise: $x_1 \succ x_2$ ($x_1 \prec x_2$).

(c) Gilt

$$x_1 \succ x_2 \vee x_1 = x_2 \quad (x_1 \prec x_2 \vee x_1 = x_2),$$

so heißt x_1 **lexikographisch größer (kleiner) gleich** [engl.: lexicographically greater (smaller) than or equal] x_2 . Schreibweise: $x_1 \succeq x_2$ ($x_1 \preceq x_2$).

(d) Sei $X \subset \mathbb{R}^n$ und $x^* \in X$ mit

$$x \in X \quad \Rightarrow \quad x^* \succeq x.$$

Dann heißt x^* **lexikographisches Maximum** von X . Man schreibt auch $x^* = \text{lexmax}(X)$. Ist $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$, so heißt x^* **lexikographisches Maximum** der Optimierungsaufgabe $\max_{x \in X} \psi(x)$, genau dann, wenn

$$x \in X \quad \Rightarrow \quad \begin{pmatrix} \psi(x^*) \\ x^* \end{pmatrix} \succeq \begin{pmatrix} \psi(x) \\ x \end{pmatrix}.$$

Die Aufgabe, ein lexikographisches Maximum⁸ zu finden, wird oft mit $\text{lexmax}_{x \in X} \psi(x)$ bezeichnet.

Wird ‘ \succeq ’ durch ‘ \preceq ’ und ‘max’ durch ‘min’ ersetzt, so spricht man analog von einem **lexikographischen Minimum**.

3.3.17 Bezeichnung. Die **lexikographischen Pivotregel** [engl.: lexicographic pivot rule] bestimmt in jedem Schritt des Simplex Algorithmus i_{rein} gemäß

$$z_{i_{\text{rein}}} = \text{lexmin} \{z_i : i \in R\}.$$

3.3.18 Bemerkung. Seien $\alpha_0, \dots, \alpha_m \in \mathbb{R}$, $\beta \in]0, \infty[$, $\pi : \mathbb{R} \rightarrow \mathbb{R}$ für $\xi \in \mathbb{R}$ definiert durch

$$\pi(\xi) := \sum_{i=0}^m \alpha_i \xi^i,$$

und es gelte

$$0 < \xi \leq \beta \quad \Rightarrow \quad \pi(\xi) \neq 0.$$

Dann gilt

$$\pi(\beta) > 0 \quad \Leftrightarrow \quad (\alpha_0, \dots, \alpha_m)^T \succ 0.$$

⁸ Das ist eine spezielle Form der ‘Mehrzieloptimierung’.

Beweis: Sei $a := (\alpha_0, \dots, \alpha_m)^T$. Nach Voraussetzung sind π nicht das Nullpolynom und $a \neq 0$. Sei k der kleinste Index mit $\alpha_k \neq 0$. Dann gilt

$$\pi(\xi) = \xi^k \left(\alpha_k + \xi \sum_{l=k+1}^m \alpha_l \xi^{l-k-1} \right).$$

Für hinreichend kleines positives ξ hat $\pi(\xi)$ dasselbe Vorzeichen wie α_k . Da π stetig ist, aber keine Nullstelle in $]0, \beta]$ besitzt, folgt hieraus die Behauptung. \square

3.3.19 Korollar. *Seien $i, j \in R$. Dann gilt*

$$z_i \succ 0 \quad \wedge \quad \left(\zeta_i(\varepsilon_0) < \zeta_j(\varepsilon_0) \Leftrightarrow z_i \prec z_j \right).$$

3.3.20 Bezeichnung. *Seien v eine Ecke von P und B eine zugehörige Basis. B heißt **lexikographisch zulässig** [engl.: *lexicographically feasible*], genau dann, wenn gilt*

$$i \in I_P(v) \setminus B \Rightarrow y_i \succ 0.$$

3.3.21 Lemma. *Seien v eine Ecke von P und B eine zugehörige Basis.*

(a) *B ist genau dann lexikographisch zulässig, wenn B $e(\varepsilon)$ -kompatibel ist.*

(b) *Seien $i_w \in R$, $z_{i_w} = \text{lexmin} \{z_i : i \in R\}$, $B_w := (B \setminus \{i_{\text{raus}}\}) \cup \{i_w\}$ sowie $w := A_{B_w}^{-1} b_{B_w}$. Dann gilt*

$$\begin{pmatrix} c^T(w - v) \\ y_{i_w} \end{pmatrix} \succ 0.$$

Beweis: (a) folgt mit den Bemerkungen 3.3.11, 3.3.18 und 3.3.15 aus der Definition der y_i .

(b) Mit $v(\varepsilon_0) = A_B^{-1}(b_B + e(\varepsilon_0))$ und $w(\varepsilon_0) = A_{B_w}^{-1}(b_{B_w} + e(\varepsilon_0))$ gilt

$$c^T w(\varepsilon_0) - c^T v(\varepsilon_0) = \lambda(\varepsilon) c^T s = \beta_{i_w} - a_{i_w}^T v + \eta_{i_w}(\varepsilon_0) > 0.$$

Mit Bemerkungen 3.3.18 folgt die Behauptung. \square

3.3.22 Satz. *Ausgehend von einer Startecke mit lexikographisch zulässiger Basis löst der Simplex-Algorithmus mit lexikographischer Pivotregel das lineare Optimierungsproblem (in natürlicher Form) in endlich vielen Schritten.*

Beweis: Die Aussage folgt mit Bemerkung 3.3.7 aus der Tatsache, dass ausgehend von einer lexikographisch zulässigen Startecke der Simplex-Algorithmus mit lexikographischer Pivotregel dieselbe Folge von Basen produziert wie der Simplex-Algorithmus für $P(\varepsilon)$. \square

3.3.23 Beispiel. *Gegeben sei die LP-Aufgabe aus dem Beweis von Satz 3.3.2 (in der Variante ohne die Restriktion (7), die für das Zykeln ohnehin irrelevant war). Wir müssen zunächst überprüfen, ob die gewählte Startbasis B^1 lexikographisch zulässig ist, d.h. ob $y_3, y_4, y_6 \succ 0$ ist. Mit*

$$B^1 := \{1,2,5\} \quad \wedge \quad A_{B^1}^{-1} = \frac{1}{3} \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ -1 & 3 & -1 \end{pmatrix}.$$

gilt aber gemäß Bemerkung 3.3.15

$$-a_4^T A_{B^1}^{-1} = -\frac{1}{3}(-4, -1, -2) \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ -1 & 3 & -1 \end{pmatrix} = \frac{1}{3}(-14, 3, -2),$$

und es folgt

$$3y_4 = (-14, 3, 0, 3, -2, 0)^T \neq 0.$$

Nach den Lemmata 3.3.21 (a) und 3.3.12 reicht es, die Restriktionen so umzunummern, dass die Basisindizes die Nummern 4,5 und 6 tragen. Man kann also etwa die durch

$$\sigma(1) := 4 \wedge \sigma(2) := 5 \wedge \sigma(3) := 1 \wedge \sigma(4) := 2 \wedge \sigma(5) := 6 \wedge \sigma(6) := 3$$

definierte Permutation σ der Restriktionsindizes von $\{1,2,3,4,5,6\}$ benutzen. Alternativ und völlig äquivalent können wir auch alle auftretenden Vektoren gemäß der neuen 'lexikographischen Ordnung'

$$(3) \succ (4) \succ (6) \succ (1) \succ (2) \succ (5)$$

bewerten. Zur einfachen Vergleichbarkeit mit der ursprünglichen Rechnung behalten wir die Nummerierung der Ungleichungen bei, und permutieren nur die Komponenten der Vektoren z_i .

Wir starten wieder mit

$$s_1 := \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \quad \wedge \quad i_{\text{raus}} := 1 \quad \wedge \quad R_1 := \{3,6\}.$$

Die Wahl von i_{rein} erfolgt jetzt aber nach der lexikographischen Regel. Zu berechnen sind also die Vektoren z_3 und z_6 . Für die Vektoren \hat{z}_3 und \hat{z}_6 mit der ursprünglichen Komponentenindizierung gilt nach Bemerkung 3.3.15

$$\begin{aligned} \hat{z}_3 &= \frac{1}{a_3^T s_1} (\beta_3 - a_3^T v_1, -a_3^T A_{B^1}^{-1} u_1, -a_3^T A_{B^1}^{-1} u_2, 1, 0, -a_3^T A_{B^1}^{-1} u_3, 0)^T \\ \hat{z}_6 &= \frac{1}{a_6^T s_1} (\beta_6 - a_6^T v_1, -a_6^T A_{B^1}^{-1} u_1, -a_6^T A_{B^1}^{-1} u_2, 0, 0, -a_6^T A_{B^1}^{-1} u_3, 1)^T. \end{aligned}$$

Mit $\beta_3 - a_3^T v_1 = \beta_6 - a_6^T v_1 = 0$ sowie $a_3^T s_1 = 2$ und $a_6^T s_1 = 3$ folgt

$$\begin{aligned} z_3 &= \frac{1}{2} (0, 1, 0, 0, -a_3^T A_{B^1}^{-1} u_1, -a_3^T A_{B^1}^{-1} u_2, -a_3^T A_{B^1}^{-1} u_3)^T \\ z_6 &= \frac{1}{3} (0, 0, 0, 1, -a_6^T A_{B^1}^{-1} u_1, -a_6^T A_{B^1}^{-1} u_2, -a_6^T A_{B^1}^{-1} u_3)^T. \end{aligned}$$

Bereits ohne Berechnung der hinteren Komponenten folgt $z_6 \prec z_3$. Auch die lexikographische Pivotregel liefert also $i_{\text{rein}} := 6$, und wir erhalten wieder

$$B^2 := \{2,5,6\} \quad \wedge \quad A_{B^2}^{-1} = \frac{1}{3} \begin{pmatrix} 30 & -6 & 3 \\ -3 & 0 & 0 \\ 13 & -3 & 1 \end{pmatrix} \quad \wedge \quad s_2 := \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \\ \wedge \quad i_{\text{raus}} := 5 \quad \wedge \quad R_2 := \{3\}.$$

Da R_2 einpunktig ist, hat auch die lexikographische Pivotregel keine andere Wahl, als $i_{\text{rein}} := 3$ zu setzen. Wir erhalten somit

$$B^3 := \{2,3,6\} \quad \wedge \quad A_{B^3}^{-1} = \frac{1}{3} \begin{pmatrix} 2 & 6 & -1 \\ -3 & 0 & 0 \\ -1 & 3 & -1 \end{pmatrix} \quad \wedge \quad s_3 := \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix} \\ \wedge \quad i_{\text{raus}} := 2 \quad \wedge \quad R_3 := \{1,4\}.$$

Zur Bestimmung der aufzunehmenden Restriktion sind nun z_1 und z_4 zu berechnen. Es gilt

$$z_1 = \frac{1}{2} (0, -a_1^T A_{B^1}^{-1} u_2, 0, -a_1^T A_{B^1}^{-1} u_3, 1, -a_1^T A_{B^1}^{-1} u_1, 0)^T \\ z_4 = \frac{1}{3} (0, -a_4^T A_{B^1}^{-1} u_2, 1, -a_4^T A_{B^1}^{-1} u_3, 0, -a_4^T A_{B^1}^{-1} u_1, 0)^T.$$

Mit

$$-\frac{1}{3} (-1, 0, 0) \begin{pmatrix} 2 & 6 & -1 \\ -3 & 0 & 0 \\ -1 & 3 & -1 \end{pmatrix} = \frac{1}{3} (2, 6, -1) \\ -\frac{1}{3} (-4, -1, -2) \begin{pmatrix} 2 & 6 & -1 \\ -3 & 0 & 0 \\ -1 & 3 & -1 \end{pmatrix} = \frac{1}{3} (3, 30, -6),$$

erhalten wir

$$z_1 = \frac{1}{6} (0, 6, 0, -1, 3, 2, 0)^T \quad \wedge \quad z_4 = \frac{1}{9} (0, 30, 3, -6, 0, 3, 0)^T.$$

Es gilt $z_1 \prec z_4$ und damit $i_{\text{rein}} := 1$. Hier weicht also die lexikographische Pivotregel von unserer ursprünglichen Wahl ab. Wir erhalten

$$\bar{B}^4 := \{1,3,6\} \quad \wedge \quad A_{\bar{B}^4}^{-1} = \frac{1}{2} \begin{pmatrix} -2 & 0 & 0 \\ 3 & 6 & -1 \\ 1 & 4 & -1 \end{pmatrix} \quad \wedge \quad \bar{s}_4 := \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \wedge \quad i_{\text{raus}} := 6.$$

Es gilt $R = \{i : a_i^T \bar{s}_4 > 0\} = \emptyset$, d.h. wir verlassen jetzt die Ecke $v_1 = 0$ längs einer unbeschränkten Verbesserungskante. (Hätten wir die Restriktion (7) berücksichtigt, so wäre $R = \{7\}$, und wir gingen jetzt direkt zu v_4 , d.h. zu einer optimalen Ecke über.)

Zum Abschluss kann man für die einzelnen Schritte noch den 'lexikographisch Fortschrittsvektor' $(c^T(w-v), y_{i_{\text{rein}}}^T)^T$ bestimmen, um auch Lemma 3.3.21 an diesem konkreten Beispiel zu überprüfen. Mit Bemerkung 3.3.15 erhält man analog zu den obigen Berechnungen die Fortschrittsvektoren der ersten drei Schritte

$$(0, y_6^T) = \frac{1}{3} (0, 0, 0, 3, 3, 30, -6) \\ (0, y_3^T) = \frac{1}{3} (0, 3, 0, -2, 0, -14, 3) \\ (0, y_1^T) = \frac{1}{3} (0, 6, 0, -1, 3, 2, 0);$$

sie sind tatsächlich streng lexikographisch wachsend.

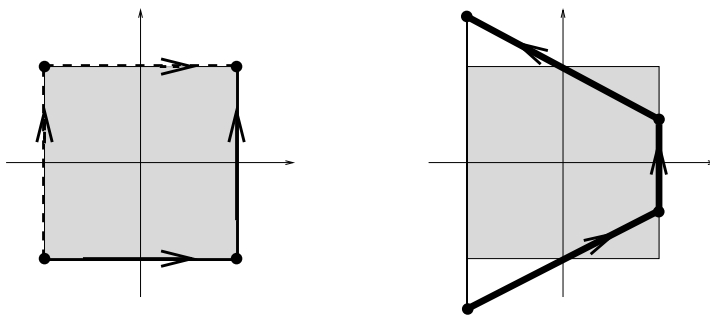
3.4 Über die Laufzeit des Simplex-Algorithmus

3.4.1 Beispiel. Seien P das Einheitsquadrat $[-1,1]^2$, $c := (1,3)^T$ und φ die durch $\varphi(x) := c^T x$ für $x \in \mathbb{R}^2$ definierte Zielfunktion. Dann ist $(1,1)^T$ Optimalpunkt von P bez. φ . Startend von der Ecke $(-1, -1)^T$ können wir als Zwischenecke $(-1,1)^T$ oder $(1, -1)^T$ durchlaufen, nicht aber beide. Der Simplex-Algorithmus benötigt somit zwei Verbesserungsschritte.

Der zulässige Bereich Q des Ungleichungssystems

$$\begin{array}{rcl} \xi_1 & \leq & 1 \\ -\xi_1 & \leq & 1 \\ \xi_1 + 2\xi_2 & \leq & 2 \\ \xi_1 - 2\xi_2 & \leq & 2 \end{array}$$

ist hingegen ein Trapez, durch dessen vier Ecken ein bez. der gleichen Zielfunktion φ aufsteigender Kantenzug verläuft.



3.7 Abbildung Quadrat P (links) und Trapez Q (rechts) mit aufsteigenden Pfaden von der Startecke $(-1, -1)^T$ zur optimalen Ecke $(1,1)^T$ bzw. $(-1,1)^T$.

Sei $\mu \in [0, 1/2]$. Gegeben sei die lineare Optimierungsaufgabe

$$\begin{array}{rcl} \max & \xi_n & \\ \xi_1 & \leq & 1 \\ -\xi_1 & \leq & 1 \\ \mu\xi_1 + \xi_2 & \leq & 1 \\ \mu\xi_1 - \xi_2 & \leq & 1 \\ & \mu\xi_2 + \xi_3 & \leq 1 \\ & \mu\xi_2 - \xi_3 & \leq 1 \\ & \ddots & \vdots \\ & \mu\xi_{n-1} + \xi_n & \leq 1 \\ & \mu\xi_{n-1} - \xi_n & \leq 1. \end{array}$$

3.4.2 Bemerkung. Seien $x := (\xi_1, \dots, \xi_n)^T \in P$ und $j \in \{1, \dots, n\}$. Dann gilt

$$-\frac{1-\mu^j}{1-\mu} \leq \xi_j \leq \frac{1-\mu^j}{1-\mu}.$$

Beweis: Wir zeigen die Behauptung mittels vollständiger Induktion nach j . Die Aussage ist korrekt für $j = 1$. Sei nun $j \geq 2$, und wir nehmen an, dass die Ungleichungen bereits für kleinere Indizes bewiesen sind. Dann gilt

$$-\frac{1-\mu^j}{1-\mu} = -1 - \mu \frac{1-\mu^{j-1}}{1-\mu} \leq -1 + \mu \xi_{j-1} \leq \xi_j \leq 1 - \mu \xi_{j-1} \leq 1 + \mu \frac{1-\mu^{j-1}}{1-\mu} = \frac{1-\mu^j}{1-\mu},$$

und es folgt die Behauptung. \square

3.4.3 Bemerkung. Für $s := (\sigma_1, \dots, \sigma_n)^T \in \{-1, 1\}^n$ sei

$$B_s := \begin{pmatrix} \sigma_1 & 0 & 0 & \dots & 0 \\ \mu & \sigma_2 & 0 & \dots & 0 \\ 0 & \mu & \sigma_3 & & 0 \\ \vdots & & \ddots & \ddots & \\ 0 & 0 & 0 & \mu & \sigma_n \end{pmatrix}$$

Sei v eine Ecke von P . Dann existiert ein $s \in \{-1, 1\}^n$ mit $v = B_s^{-1} \mathbf{1}$.

Beweis: Man beachte zunächst, dass jede der Matrizen B_s regulär ist.

Wir nehmen an, dass v nicht gleich einem der Punkte $B_s^{-1} \mathbf{1}$ ist. Sei $Bx = \mathbf{1}$ eine zu v gehörige Basisgleichung. Da $\text{rang}(B) = n$ gilt, kann B nicht beide Zeilen u_1^T und $-u_1^T$ enthalten. Es muss daher einen Index $j \in \{2, \dots, n\}$ geben, so dass B die beiden Zeilen $(\mu u_{j-1}^T \pm u_j^T)$ enthält. Durch Addition der entsprechenden Gleichungen des Systems $Bx = \mathbf{1}$ folgt mit $1 - 2\mu + \mu^j \geq \mu^j > 0$

$$\xi_{j-1} = \frac{1}{\mu} > \frac{1-\mu^{j-1}}{1-\mu},$$

im Widerspruch zu Bemerkung 3.4.2. Somit enthält B für jedes $k \in \{1, \dots, n\}$ jeweils genau eine der Zeilen mit Indizes aus $\{2k-1, 2k\}$, ist also von der behaupteten Form. \square

3.4.4 Bemerkung. Seien $s := (\sigma_1, \dots, \sigma_n)^T \in \{-1, 1\}^n$, die Matrix B_s wie in Bemerkung 3.4.3 und

$$v(s) := (\nu_1(s), \dots, \nu_n(s))^T := B_s^{-1} \mathbf{1}.$$

Dann gilt für $j = 1, \dots, n$

$$\nu_j(s) = \sum_{i=0}^{j-1} (-1)^i \left(\prod_{l=0}^i \sigma_{j-l} \right) \mu^i.$$

Beweis: Vollständige Induktion nach j . \square

3.4.5 Lemma. (a) Für jedes $s \in \{-1, 1\}^n$ ist $v(s)$ eine reguläre Ecke von P .

(b) Seien $s_1, s_2 \in \{-1, 1\}^n$. Dann ist $\text{conv}(\{v(s_1), v(s_2)\})$ genau dann eine Kante von P , wenn es ein $j \in \{1, \dots, n\}$ gibt mit $s_1 - s_2 \in \{-2u_j, 2u_j\}$, d.h. wenn sich s_1 und s_2 in genau einer Komponente unterscheiden.

Beweis: Seien $s := (\sigma_1, \dots, \sigma_n)^T$ und $t := (\tau_1, \dots, \tau_n)^T \in \{-1, 1\}^n$. Wir bestimmen $B_t v(s)$. Natürlich gilt

$$u_1^T B_t v(s) = \tau_1 \nu_1(s) \in \{-1, 1\}.$$

Sei also nun $j \in \{2, \dots, n\}$. Dann gilt nach Bemerkung 3.4.4

$$\begin{aligned} u_j^T B_t v(s) &= \mu \nu_{j-1}(s) + \tau_j \nu_j(s) \\ &= \sum_{i=0}^{j-2} (-1)^i \left(\prod_{l=0}^i \sigma_{j-1-l} \right) \mu^{i+1} + \tau_j \sum_{i=0}^{j-1} (-1)^i \left(\prod_{l=0}^i \sigma_{j-l} \right) \mu^i \\ &= \tau_j \sigma_j + \sum_{i=0}^{j-2} (-1)^i \left(\prod_{l=1}^{i+1} \sigma_{j-l} - \tau_j \prod_{l=0}^{i+1} \sigma_{j-l} \right) \mu^{i+1} \\ &= \tau_j \sigma_j + (1 - \tau_j \sigma_j) \cdot \sum_{i=0}^{j-2} (-1)^i \left(\prod_{l=1}^{i+1} \sigma_{j-l} \right) \mu^{i+1}. \end{aligned}$$

Für $\tau_j = \sigma_j$ gilt somit $u_j^T B_t v(s) = 1$; für $\tau_j \neq \sigma_j$ folgt $\tau_j = -\sigma_j$ und damit

$$\begin{aligned} u_j^T B_t v(s) &= -1 + 2 \cdot \sum_{i=0}^{j-2} (-1)^i \left(\prod_{l=1}^{i+1} \sigma_{j-l} \right) \mu^{i+1} \leq -1 + 2 \cdot \sum_{i=0}^{j-2} \mu^{i+1} \\ &= -1 + 2\mu \frac{1 - \mu^{j-1}}{1 - \mu} \leq -1 + 2 \left(1 - \frac{1}{2^{j-1}} \right) = 1 - \frac{1}{2^{j-2}} < 1. \end{aligned}$$

Daher erfüllt $v(s)$ genau die zu B_s gehörigen Restriktionen mit Gleichheit, alle übrigen sogar strikt. Insgesamt folgt (a).

Zum Beweis von (b) sei $s_1 \neq s_2$. Die beiden Ecken $v(s_1)$ und $v(s_2)$ sind genau dann durch eine Kante verbunden, wenn $v(s_2)$ alle bis auf genau eine der Basisrestriktionen von $v(s_1)$ mit Gleichheit erfüllt. Anwendung der bereits bewiesenen Ungleichung mit $s = s_1$ und $t = s_2$ zeigt, dass das genau dann der Fall ist, wenn sich s_1 und s_2 in genau einer Komponente unterscheiden.

Insgesamt ist das Lemma hiermit bewiesen. \square

3.4.6 Bezeichnung. Das Polytop P wird *Klee-Minty Würfel* genannt.

Wir zeigen nun, dass es tatsächlich einen bez. des Wertes der letzten Komponente aufsteigenden Pfad in P gibt.

3.4.7 Lemma. Für $s := (\sigma_1, \dots, \sigma_n)^T \in \{-1, 1\}^n$ seien

$$\kappa_i(s) := (-1)^i \prod_{l=0}^i \sigma_{n-l} \quad (i = 0, \dots, n-1) \quad \wedge \quad q(s) := (\kappa_0(s), \dots, \kappa_{n-1}(s))^T.$$

Seien nun $s := (\sigma_1, \dots, \sigma_n)^T$ und $t := (\tau_1, \dots, \tau_n)^T \in \{-1, 1\}^n$ mit $s \neq t$.

(a) Dann gilt

$$\varphi(v(s)) \neq \varphi(v(t)) \quad \wedge \quad \left(\varphi(v(s)) < \varphi(v(t)) \Leftrightarrow q(s) \prec q(t) \right).$$

(b) Sei $q(s) \prec q(t)$, sei k der Index der ersten von 0 verschiedenen Komponente von $q(t) - q(s)$, und es gebe kein $r \in \{-1, 1\}^n$ mit $q(s) \prec q(r) \prec q(t)$. Dann gilt $t - s = 2u_k$, d.h. t entsteht aus s durch Wechsel der k -ten Komponente von -1 zu 1 .

Beweis: (a) Nach Bemerkung 3.4.4 gilt

$$\varphi(v(t)) - \varphi(v(s)) = \sum_{i=0}^{n-1} (-1)^i \left(\prod_{l=0}^i \tau_{n-l} - \prod_{l=0}^i \sigma_{n-l} \right) \mu^i = \sum_{i=0}^{n-1} (\kappa_i(t) - \kappa_i(s)) \mu^i.$$

Für $k \in \{0, 1, \dots, n-1\}$ folgt mit $0 < \mu \leq 1/2$

$$\begin{aligned} \left| \sum_{i=k+1}^{n-1} (\kappa_i(t) - \kappa_i(s)) \mu^i \right| &\leq \sum_{i=k+1}^{n-1} |\kappa_i(t) - \kappa_i(s)| \mu^i \leq 2 \sum_{i=k+1}^{n-1} \mu^i \\ &= 2\mu^{k+1} \frac{1 - \mu^{n-k-1}}{1 - \mu} \leq 2\mu^k (1 - \mu^{n-k-1}). \end{aligned}$$

Sei nun k der Index der kleinsten Komponente, in der sich $q(t)$ und $q(s)$ unterscheiden. Dann gilt

$$\varphi(v(t)) - \varphi(v(s)) = (\kappa_k(t) - \kappa_k(s))\mu^k + \sum_{i=k+1}^{n-1} (\kappa_i(t) - \kappa_i(s)) \mu^i.$$

Es folgt

$$\begin{aligned} \left| \varphi(v(t)) - \varphi(v(s)) \right| &\geq \left| |(\kappa_k(t) - \kappa_k(s))\mu^k| - |2\mu^k(1 - \mu^{n-k-1})| \right| \\ &= |2\mu^k - 2\mu^k(1 - \mu^{n-k-1})| = 2\mu^{n-1} > 0, \end{aligned}$$

und damit die erste Behauptung von (a).

Ferner ist für $\kappa_k(t) > \kappa_k(s)$

$$\begin{aligned} \varphi(v(t)) - \varphi(v(s)) &\geq (\kappa_k(t) - \kappa_k(s))\mu^k - 2\mu^k(1 - \mu^{n-k-1}) \\ &= 2\mu^k - 2\mu^k(1 - \mu^{n-k-1}) = 2\mu^{n-1} > 0 \end{aligned}$$

und für $\kappa_k(t) < \kappa_k(s)$

$$\begin{aligned} \varphi(v(t)) - \varphi(v(s)) &\leq (\kappa_k(t) - \kappa_k(s))\mu^k + 2\mu^k(1 - \mu^{n-k-1}) \\ &= -2\mu^k + 2\mu^k(1 - \mu^{n-k-1}) = -2\mu^{n-1} < 0. \end{aligned}$$

Insgesamt ist damit (a) nachgewiesen.

(b) Natürlich gilt

$$\kappa_0(t) = \kappa_0(s) \quad \wedge \quad \dots \quad \wedge \quad \kappa_{k-1}(t) = \kappa_{k-1}(s) \quad \wedge \quad \kappa_k(t) = 1 > -1 = \kappa_{k-1}(s).$$

Angenommen, es gäbe einen Index $l \in \{k+1, \dots, n-1\}$ mit $\kappa_l(t) = 1$ oder $\kappa_l(t) = -1$. Im ersten Fall wäre

$$q(t) \succ (\kappa_0(t), \dots, \kappa_{l-1}(t), -1, \kappa_{l+1}(t), \dots, \kappa_{n-1}(t))^T \succ q(s),$$

im zweiten

$$q(t) \succ (\kappa_0(s), \dots, \kappa_{l-1}(s), 1, \kappa_{l+1}(s), \dots, \kappa_{n-1}(s))^T \succ q(s),$$

jeweils im Widerspruch zur Voraussetzung, dass $q(t)$ der lexikographische Nachfolger von $q(s)$ ist. Somit gilt mit $\kappa_i(t) = \kappa_i(s) = \kappa_i$ für $i = 0, \dots, k-1$

$$q(t) = (\kappa_0, \dots, \kappa_{k-1}, +1, -1, \dots, -1)^T \quad \wedge \quad q(s) = (\kappa_0, \dots, \kappa_{k-1}, -1, +1, \dots, +1)^T.$$

Nun ist für $i = 0, \dots, n-1$

$$\kappa_i(t) - \kappa_i(s) = (-1)^i \left(\prod_{l=0}^i \tau_{n-l} - \prod_{l=0}^i \sigma_{n-l} \right).$$

Somit folgt für $i = 0, \dots, k-1$

$$\prod_{l=0}^i \tau_{n-l} = \prod_{l=0}^i \sigma_{n-l}$$

und damit sukzessive (in umgekehrter Reihenfolge)

$$\tau_{n-k+1} = \sigma_{n-k+1} \quad \wedge \quad \dots \quad \wedge \quad \tau_n = \sigma_n.$$

Ferner gilt

$$2 = \kappa_k(t) - \kappa_k(s) = (-1)^k (\tau_{n-k} - \sigma_{n-k}) \prod_{l=0}^{k-1} \tau_{n-l},$$

und das impliziert

$$\tau_{n-k} = -\sigma_{n-k}.$$

Für $i = k+1, \dots, n-1$ erhalten wir abschließend

$$-2 = \kappa_i(t) - \kappa_i(s) = (-1)^i \prod_{l=0}^k \tau_{n-l} \left(\prod_{l=k+1}^i \tau_{n-l} + \prod_{l=0}^i \sigma_{n-l} \right),$$

und es folgt sukzessive (und wieder in umkehrter Reihenfolge)

$$\tau_1 = \sigma_1 \quad \wedge \quad \dots \quad \wedge \quad \tau_{n-k-1} = \sigma_{n-k-1}.$$

Insgesamt ist damit die Behauptung bewiesen. □

3.4.8 Beispiel. Für den Klee-Minty Würfel im \mathbb{R}^2 gilt für $\mu = 1/2$

$$\begin{aligned} v((-1, -1)^T) &= \begin{pmatrix} -1 \\ -3/2 \end{pmatrix} \quad \wedge \quad v((+1, -1)^T) = \begin{pmatrix} +1 \\ -1/2 \end{pmatrix} \\ v((+1, +1)^T) &= \begin{pmatrix} +1 \\ +1/2 \end{pmatrix} \quad \wedge \quad v((-1, +1)^T) = \begin{pmatrix} -1 \\ +3/2 \end{pmatrix} \end{aligned}$$

sowie

$$\begin{aligned} q((-1, -1)^T) &= \begin{pmatrix} -1 \\ -1 \end{pmatrix} \quad \wedge \quad q((+1, -1)^T) = \begin{pmatrix} -1 \\ +1 \end{pmatrix} \\ q((+1, +1)^T) &= \begin{pmatrix} +1 \\ -1 \end{pmatrix} \quad \wedge \quad q((-1, +1)^T) = \begin{pmatrix} +1 \\ +1 \end{pmatrix}. \end{aligned}$$

3.4.9 Satz. *(Beispiel von Klee und Minty)*

Bei entsprechender Wahl der Pivotpaare kann der Simplex-Algorithmus für eine Aufgabe mit $2n$ Restriktionen im \mathbb{R}^n genau $2^n - 1$ Verbesserungsschritte durchlaufen.

3.4.10 Forschungsproblem. *Gibt es eine Pivotregel, mit der der Simplex-Algorithmus jede durch (m, n, A, b, c) spezifizierte Aufgabe in einer Anzahl von Schritten löst, die durch ein Polynom in m und n beschränkt ist.*

4 LP-Dualität

4.1 Dualität linearer Programme

4.1.1 Notation. Durch (n, m, A, b, c) sei eine LP-Aufgabe in natürlicher Form spezifiziert. Es seien $M := \{1, \dots, m\}$, $a_1, \dots, a_m \in \mathbb{R}^n \setminus \{0\}$ und $\bar{a}_1, \dots, \bar{a}_n \in \mathbb{R}^m$ mit $A = (a_1, \dots, a_m)^T = (\bar{a}_1, \dots, \bar{a}_n)$, $\beta_1, \dots, \beta_m \in \mathbb{R}$ mit $b = (\beta_1, \dots, \beta_m)^T$, $P := \{x \in \mathbb{R}^n : Ax \leq b\}$ sowie $\gamma_1, \dots, \gamma_n \in \mathbb{R}$ mit $c = (\gamma_1, \dots, \gamma_n)^T$. Die Variablenvektoren seien $x := (\xi_1, \dots, \xi_n)^T \in \mathbb{R}^n$ und $y := (\eta_1, \dots, \eta_m)^T \in \mathbb{R}^m$.

4.1.2 Bemerkung. Sei $x^* \in \mathbb{R}^n$ mit $Ax^* \leq b$. x^* ist genau dann ein Optimalpunkt der linearen Optimierungsaufgabe

$$\begin{aligned} \max c^T x \\ Ax \leq b \end{aligned}$$

wenn es einen Vektor $y^* := (\eta_1^*, \dots, \eta_m^*)^T \in \mathbb{R}^m$ gibt mit

$$A^T y^* = c \quad \wedge \quad y^* \geq 0 \quad \wedge \quad (i \in M \wedge a_i^T x^* < \beta_i \Rightarrow \eta_i^* = 0).$$

4.1.3 Beispiel. Gegeben sei die in Abbildung 4.1 (links) dargestellte lineare Optimierungsaufgabe

$$\begin{aligned} \max c^T x \\ a_i^T x \leq b \quad (i = 1, \dots, 5). \end{aligned}$$

v_1 und v_2 seien die Schnittpunkte der Restriktionsgeraden

$$a_1^T x = \beta_1 \quad \wedge \quad a_4^T x = \beta_4$$

bzw.

$$a_2^T x = \beta_2 \quad \wedge \quad a_4^T x = \beta_4.$$

Es gilt

$$c \in \text{pos}(\{a_1, a_4\}) \quad \wedge \quad c \in \text{pos}(\{a_2, a_4\}).$$

Im Sinne des Kegelkriteriums sind v_1 und v_2 daher 'optimal', nicht aber zulässig.

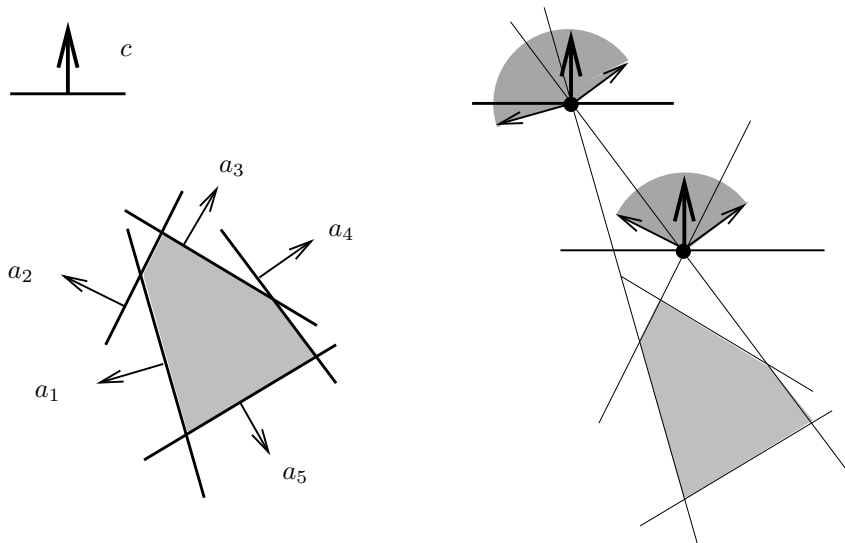
Seien $I \subset \{1, \dots, 5\}$ mit $|I| = 2$, $v_I := A_I^{-1} b_I$ und $y_I \in [0, \infty[$ mit $c = (A_I)^T y_I$. Dann gilt

$$c^T v_I = ((A_I)^T y_I)^T A_I^{-1} b_I = y_I^T b_I.$$

Das Minimum der rechten Seite wird gerade für $I^* = \{2, 3\}$ angenommen; v_{I^*} ist optimal.

4.1.4 Notation. Die beiden LP-Aufgaben

$$\begin{aligned} \max c^T x \\ \text{(I) } Ax \leq b \end{aligned} \quad \begin{aligned} \min b^T y \\ \text{(II) } A^T y = c \\ y \geq 0, \end{aligned}$$



4.1 Abbildung Links: Zielfunktionsvektor c und zulässiges Polyeder einer LP-Aufgabe; Rechts: Zwei verschiedene Darstellungen von c als Nichtnegativkombinationen von zwei Restriktionsnormalen.

werden im folgenden mit (I) bzw. (II) bezeichnet. Es seien

$$Q := \{y \in \mathbb{R}^m : A^T y = c \wedge y \geq 0\},$$

$\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ und $\psi : \mathbb{R}^m \rightarrow \mathbb{R}$ bezeichnen die durch $\varphi(x) := c^T x$ bzw. $\psi(y) := b^T y$ definierten Zielfunktionen, und $\zeta_{(I)}$ bzw. $\zeta_{(II)}$ seien die (möglicherweise uneigentlichen) Optimalwerte von (I) bzw. (II).

4.1.5 Bezeichnung. Die linearen Optimierungsaufgaben (I) und (II) heißen zueinander **dual** [engl.: dual]. Eine der LP-Aufgaben wird oft als **primale** [engl.: primal], die andere dann als zugehörige **duale** Aufgabe bezeichnet.¹

4.1.6 Lemma. Sei $y^* := (\eta_1^*, \dots, \eta_m^*)^T \in Q$. y^* ist genau dann ein Optimalpunkt von (II), wenn es einen Punkt $x^* \in \mathbb{R}^n$ gibt mit

$$Ax^* \leq b \quad \wedge \quad (i \in M \wedge \eta_i^* > 0 \Rightarrow a_i^T x^* = \beta_i).$$

Beweis: Um Bemerkung 4.1.2 bzw. Bemerkung 2.2.22 und Lemma 2.2.28 anzuwenden, überführen wir (II) zunächst in natürliche Form. Offenbar gilt für jedes $y \in Q$

$$b^T y^* \leq b^T y \quad \Leftrightarrow \quad (-b)^T y^* \geq (-b)^T y$$

sowie

$$Q = \{y : A^T y \leq c \wedge -A^T y \leq -c \wedge -y \leq 0\}.$$

Die Restriktionen $A^T y \leq c$ und $-A^T y \leq -c$ sind natürlich stets aktiv. Sei $I^* := \{i \in M : \eta_i^* = 0\}$. Der Vektor y^* ist genau dann Optimalpunkt von (II), wenn

¹ Für uns ist meistens (I) primal und (II) dual, aber das ist willkürlich, da Dualität in gewissem Sinn involutorisch ist; vgl. Lemma 4.1.6 sowie die Dualitätstabelle 4.1.11.

$$-b \in N_Q(y^*) = \text{pos}(\{\pm \bar{a}_1, \dots, \pm \bar{a}_n\} \cup \{-u_i : i \in I^*\})$$

gilt, d.h. wenn es $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in [0, \infty[$ und für $i \in I^*$ ferner $\sigma_i \in [0, \infty[$ gibt mit

$$\begin{aligned} -\beta_i &= \sum_{j=1}^n u_i^T \bar{a}_j \lambda_j - \sum_{j=1}^n u_i^T \bar{a}_j \mu_j - \sigma_i & (i \in I^*) \\ -\beta_i &= \sum_{j=1}^n u_i^T \bar{a}_j \lambda_j - \sum_{j=1}^n u_i^T \bar{a}_j \mu_j & (i \notin I^*). \end{aligned}$$

Mit der Setzung $\xi_j^* := \mu_j - \lambda_j$ für $j \in \{1, \dots, n\}$ und $x^* := (\xi_1^*, \dots, \xi_n^*)^T$ ist diese Bedingung äquivalent zu

$$\begin{aligned} \beta_i &\geq u_i^T A x^* & (i \in I^*) \\ \beta_i &= u_i^T A x^* & (i \notin I^*). \end{aligned}$$

Hieraus folgt die Behauptung. \square

4.1.7 Satz. (Dualitätssatz der linearen Optimierung)

Es gelten die folgenden Aussagen:

- (a) $\zeta_{(I)} \leq \zeta_{(II)}$.
- (b) $P \neq \emptyset \vee Q \neq \emptyset \Rightarrow \zeta_{(I)} = \zeta_{(II)}$.
- (c) Seien $x^* \in P$ und $y^* \in Q$. Dann gilt

$$c^T x^* = b^T y^* \Leftrightarrow \zeta_{(I)} = c^T x^* \wedge \zeta_{(II)} = b^T y^* \Leftrightarrow (y^*)^T (b - A x^*) = 0.$$

Beweis: (a) Für $x \in P$ und $y \in Q$ gilt $c^T x = y^T A x \leq y^T b$. Mit den üblichen Konventionen $\max_{x \in \emptyset} c^T x = -\infty$ und $\min_{y \in \emptyset} b^T y := \infty$ folgt $\zeta_{(I)} \leq \zeta_{(II)}$.

(b) Seien $P \neq \emptyset$, aber φ über P nach oben unbeschränkt. Dann folgt aus (a) $\zeta_{(I)} = \zeta_{(II)} = \infty$. Gilt $Q \neq \emptyset$, und ist ψ auf Q nach unten unbeschränkt, so folgt analog $\zeta_{(I)} = \zeta_{(II)} = -\infty$.

Sei nun x^* ein Optimalpunkt von (I). Nach Bemerkung 4.1.2 gibt es einen Punkt $y^* \in \mathbb{R}^m$ mit

$$A^T y^* = c \quad \wedge \quad y^* \geq 0 \quad \wedge \quad (i \in M \wedge a_i^T x^* < \beta_i \Rightarrow \eta_i^* = 0).$$

Also gilt $Q \neq \emptyset$, und es folgt mit (a)

$$\zeta_{(II)} \geq \zeta_{(I)} = c^T x^* = (y^*)^T A x^* = (y^*)^T b;$$

y^* ist also Minimalpunkt von (II).

Sei nun andererseits y^* ein Optimalpunkt von (II). Mit Lemma 4.1.6 folgt die Existenz eines Punktes $x^* := (\xi_1^*, \dots, \xi_n^*)^T$ mit

$$A x^* \leq b \quad \wedge \quad (i \in M \wedge \eta_i^* > 0 \Rightarrow a_i^T x^* = \beta_i).$$

Also gilt $x^* \in P$ und

$$\zeta_{(II)} = b^T y^* = (x^*)^T A^T y^* = (x^*)^T c \leq \zeta_{(I)},$$

d.h. x^* ist Maximalpunkt von (I).

(c) Die erste der beiden Äquivalenzen ist eine direkte Folge von (a) und (b). Zum Beweis der zweiten beachte man, dass

$$0 \leq b^T y^* - c^T x^* = (y^*)^T (b - Ax^*)$$

gilt. Somit ist $(y^*)^T (b - Ax^*) = 0$ äquivalent zu $b^T y^* = c^T x^*$.

Insgesamt folgt damit die Behauptung. \square

4.1.8 Bezeichnung. (a) Die Ungleichung $\zeta_{(I)} \leq \zeta_{(II)}$ heißt **schwache Dualitätsaussage** [engl.: weak duality]; die quadratische Gleichung $y^T (b - Ax) = 0$ **Komplementaritätsbedingung** [engl.: complementarity condition, complementary slackness].

(b) Jedes Paar (x^*, y^*) eines für (I) zulässigen Punktes x^* und eines für (II) zulässigen Punktes y^* mit $(y^*)^T (b - Ax^*) = 0$ heißt **primal-duales Paar** [engl.: primal-dual pair].

4.1.9 Korollar. Es gilt eine der folgenden Aussagen:

(a) Beide Aufgaben sind zulässig, besitzen endliche Optimalpunkte und die Optima der Zielfunktionen stimmen überein.

Der Punkt $x^* \in P$ ist genau dann optimal in (I), wenn x^* zu einem primal-dualen Paar ergänzt werden kann.

Der Punkt $y^* \in Q$ ist genau dann optimal in (II), wenn y^* zu einem primal-dualen Paar ergänzt werden kann.

(b) Eine der Aufgaben ist unzulässig, die andere (in der Optimierungsrichtung) unbeschränkt.

(c) Beide Aufgaben sind unzulässig.

Beweis: Der Beweis ergibt sich direkt aus dem Dualitätssatz 4.1.7. \square

4.1.10 Beispiel. Seien $\beta_1, \beta_2, \gamma_1, \gamma_2 \in \mathbb{R}$. Die folgenden beiden Aufgaben sind dual zueinander.

$$\begin{array}{ll} \max & \gamma_1 \xi_1 + \gamma_2 \xi_2 \\ \xi_1 + \xi_2 & \leq \beta_1 \\ -\xi_1 - \xi_2 & \leq -\beta_2 \end{array} \quad \begin{array}{ll} \min & \beta_1 \eta_1 - \beta_2 \eta_2 \\ \eta_1 - \eta_2 & = \gamma_1 \\ \eta_1 - \eta_2 & = \gamma_2 \\ \eta_1 & \geq 0 \\ \eta_2 & \geq 0 \end{array}$$

Offenbar ist die links stehende primale Aufgabe genau dann zulässig, wenn

$$\beta_1 \geq \beta_2$$

gilt. Ist der zulässige Bereich P nicht leer, so gilt $\text{ls}(P) = \text{lin}\{(-1, 1)^T\}$. Also hat die primale Aufgabe im Falle $P \neq \emptyset$ genau dann ein endliches Optimum, wenn $c \in (\text{ls}(P))^\perp$ gilt, d.h. falls

$$\gamma_1 = \gamma_2$$

ist.

Die duale Aufgabe ist hingegen genau dann zulässig, wenn die beiden Restriktionsgeraden übereinstimmen, d.h. wenn

$$\gamma_1 = \gamma_2$$

gilt. Im Falle der Zulässigkeit liegt genau dann ein endliches Optimum vor, wenn $(\beta_1, -\beta_2)^T$ im Halbraum $H_{((1,1)^T, 0)}^{\geq}$ liegt, d.h. wenn

$$\beta_1 \geq \beta_2$$

gilt.

4.1.11 Dualitätstabelle. Seien $\{M_1, M_2, M_3\}$ eine Partition von $\{1, \dots, m\}$ und $\{N_1, N_2, N_3\}$ eine Partition von $\{1, \dots, n\}$. Dann entsprechen sich Restriktionen und Variablen in primalen und dazu dualen LP-Aufgaben gemäß der folgenden Tabelle, die für Maximierungsaufgaben von links nach rechts und für Minimierungsaufgaben von rechts nach links zu lesen.

$\max c^T x$		$\min b^T y$
$a_i^T x = \beta_i$	$i \in M_1$	η_i
$a_i^T x \leq \beta_i$	$i \in M_2$	$\eta_i \geq 0$
$a_i^T x \geq \beta_i$	$i \in M_3$	$\eta_i \leq 0$
$\xi_j \geq 0$	$j \in N_1$	$\bar{a}_j^T y \geq \gamma_j$
$\xi_j \leq 0$	$j \in N_2$	$\bar{a}_j^T y \leq \gamma_j$
ξ_j	$j \in N_3$	$\bar{a}_j^T y = \gamma_j$

Beweis: Es reicht, die links stehende LP-Aufgabe zu betrachten, wenn man lediglich noch unterscheidet, ob $\max c^T x$ oder $\min c^T x$ zugrunde liegt. Hierdurch können beide Teile der Behauptung ('Lesen der Tabelle von rechts nach links und von links nach rechts') gleichzeitig bewiesen werden. Wir transformieren daher die links stehende LP-Aufgabe für beide Optimierungsvarianten in die natürliche Form:

$$\begin{array}{lll} \max c^T x & / & -\max(-c)^T x \\ a_i^T x & \leq & \beta_i \quad (i \in M_1) \\ -a_i^T x & \geq & -\beta_i \quad (i \in M_1) \\ a_i^T x & \leq & \beta_i \quad (i \in M_2) \\ -a_i^T x & \leq & -\beta_i \quad (i \in M_3) \\ -\xi_j & \leq & 0 \quad (j \in N_1) \\ \xi_j & \leq & 0 \quad (j \in N_2). \end{array}$$

Bezieht man die beiden Vorzeichen \pm auf die beiden verschiedenen Zielfunktionen, so lautet die hierzu duale Aufgabe:

$$\begin{aligned}
& \pm \min \sum_{i \in M_1} \beta_i (\mu_i - \nu_i) + \sum_{i \in M_2} \beta_i \eta_i - \sum_{i \in M_3} \beta_i \eta_i \\
\sum_{i \in M_1} a_i (\mu_i - \nu_i) + \sum_{i \in M_2} a_i \eta_i - \sum_{i \in M_3} a_i \eta_i - \sum_{j \in N_1} u_j \sigma_j + \sum_{j \in N_2} u_j \sigma_j &= \pm c \\
\mu_i, \nu_i &\geq 0 & (i \in M_1) \\
\eta_i &\geq 0 & (i \in M_2) \\
\eta_i &\geq 0 & (i \in M_3) \\
\sigma_j &\geq 0 & (j \in N_1) \\
\sigma_j &\geq 0 & (j \in N_2).
\end{aligned}$$

Mit der Setzung $\eta_i := \mu_i - \nu_i$ für $i \in M_1$ ist diese Aufgabe äquivalent zu

$$\begin{aligned}
& \pm \min \sum_{i=1}^m \beta_i \eta_i \\
\sum_{i=1}^m a_i \eta_i + \sum_{j \in N_1 \cup N_2} u_j \sigma_j &= \pm c \\
\eta_i &\geq 0 & (i \in M_2) \\
\eta_i &\leq 0 & (i \in M_3) \\
\sigma_j &\leq 0 & (j \in N_1) \\
\sigma_j &\geq 0 & (j \in N_2).
\end{aligned}$$

Die Variablen σ_j für $j \in N_1 \cup N_2$ können ebenfalls entfernt werden, wenn die entsprechenden Gleichungen durch Ungleichungen ersetzt werden; wir erhalten die äquivalente Aufgabe:

$$\begin{aligned}
& \pm \min \sum_{i=1}^m \beta_i \eta_i \\
\sum_{i=1}^m a_i \eta_i &\geq \pm \gamma_j & (j \in N_1) \\
\sum_{i=1}^m a_i \eta_i &\leq \pm \gamma_j & (j \in N_2) \\
\sum_{i=1}^m a_i \eta_i &= \pm \gamma_j & (j \in N_3) \\
\eta_i &\geq 0 & (i \in M_2) \\
\eta_i &\leq 0 & (i \in M_3).
\end{aligned}$$

Unter Verwendung der Spaltenvektoren $\bar{a}_1, \dots, \bar{a}_1$ erhalten wir für die Maximierungsaufgabe (links) und nach anschließendem Ersetzen von y durch $-y$ für die Minimierungsaufgabe (rechts):

$$\begin{array}{lll}
\min b^T y & & \max b^T y \\
\bar{a}_j^T y \geq \gamma_j & (j \in N_1) & \bar{a}_j^T y \leq \gamma_j \\
\bar{a}_j^T y \leq \gamma_j & (j \in N_2) & \bar{a}_j^T y \geq \gamma_j \\
\bar{a}_j^T y = \gamma_j & (j \in N_3) & \bar{a}_j^T y = \gamma_j \\
\eta_i \geq 0 & (i \in M_2) & \eta_i \leq 0. \\
\eta_i \leq 0 & (i \in M_3) & \eta_i \geq 0.
\end{array}$$

Insgesamt folgt damit (bei entsprechender Übersetzung in die Bezeichnungen der Tabellenleserichtung ‘von rechts nach links’) die Behauptung. \square

4.1.12 Korollar. Die lineare Optimierungsaufgabe

$$\begin{aligned} \max \quad & c^T x \\ Ax \leq & b \\ x \geq & 0 \end{aligned}$$

besitzt genau dann einen Optimalpunkt, wenn die lineare Ungleichungsaufgabe

$$\begin{aligned} c^T x - b^T y & \geq 0 \\ Ax & \leq b \\ A^T y & \geq c \\ x & \geq 0 \\ y & \geq 0 \end{aligned}$$

zulässig ist. $\begin{pmatrix} x^* \\ y^* \end{pmatrix}$ ist genau dann Lösung des Ungleichungssystems, wenn (x^*, y^*) ein primal-duales Paar der LP-Aufgabe ist.

Beweis: ‘ \Rightarrow ’ Ist x^* Optimalpunkt der primalen LP-Aufgabe, so existiert nach Korollar 4.1.9 ein y^* , so dass (x^*, y^*) ein primal-duales Paar bilden. Nach Tabelle 4.1.11 und Satz 4.1.7 erfüllt dann $(x^*)^T, (y^*)^T$ das Ungleichungssystem.

‘ \Leftarrow ’ Erfüllt umgekehrt $(x^*)^T, (y^*)^T$ das Ungleichungssystem, so sind x^* primal und y^* dual zulässig. Nach Satz 4.1.7 (a) gilt daher $c^T x^* \leq b^T y^*$, also insgesamt $c^T x^* = b^T y^*$, und die Behauptung folgt aus Satz 4.1.7 (c). \square

Jeder Algorithmus, der lineare Ungleichungssysteme löst, kann somit verwendet werden, um LP-Aufgaben zu lösen. Man beachte aber, daß der Übergang vom primalen Maximierungsproblem zum primal-dualen Zulässigkeitsproblem gemäß Korollar 4.1.12 im allgemeinen neue Instabilitäten generiert. Das liegt an der ‘Zielfunktionsungleichung’ $c^T x^* \geq b^T y^*$, die nach Satz 4.1.7 (a) mit für jeden zulässigen Punkt (x^T, y^T) mit Gleichheit erfüllt sein muss. Auch wenn also sowohl der zulässige Bereich der primalen als auch der der dualen Aufgabe innere Punkte besitzen, ist die Lösungsmenge des zugehörigen primal-dualen Ungleichungssystems im allgemeinen (d.h. wenn nicht gerade $b = 0$ und $c = 0$ gilt) unterdimensional, und für beliebig kleine positive ‘Störung’ ϵ ist

$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix} : Ax \leq b \wedge A^T y \geq c \wedge x \geq 0 \wedge y \geq 0 \wedge c^T x - b^T y \geq \epsilon \right\}$$

leer.

4.1.13 Beispiel. Die folgenden beiden Aufgaben sind dual zueinander.

$$\begin{array}{ll} \max \quad & \xi \\ \xi \leq & 2 \\ \xi \geq & 0 \end{array} \qquad \begin{array}{ll} \min \quad & 2\eta \\ \eta \geq & 1 \\ \eta \geq & 0 \end{array}$$

Zur Bestimmung eines primal-dualen Paares gemäß Korollar 4.1.12 ist den primalen und den dualen Zulässigkeitsbedingungen noch die ‘Zielfunktionsrestriktion’

$$\xi - 2\eta \geq 0$$

hinzuzufügen.

4.2 Dualität in ökonomischen Modellen

4.2.1 Bezeichnung. Für lineare Produktionsaufgabe werden die Ressourcen R_1, \dots, R_m in der Ökonomie oft auch **Faktoren** genannt. Ist (x^*, y^*) ein primal-duales Paar, so heißen die Komponenten $\eta_1^*, \dots, \eta_m^*$ von y^* (zu x^* gehörige) **Schattenpreise**.

4.2.2 Bemerkung. Seien $q := (\kappa_1, \dots, \kappa_m)^T \in \mathbb{R}^m$ und $s := (\sigma_1, \dots, \sigma_m)^T \in \mathbb{R}^m$. Dann gilt

$$\max\{c^T x : Ax \leq b \wedge x \geq 0\} = \max\{c^T x - s^T q : Ax \leq b + q \wedge x \geq 0\}$$

genau dann, wenn s Optimalpunkt der dualen Aufgabe $\min\{b^T y : A^T y \geq c \wedge y \geq 0\}$ ist.

Beweis: Nach dem Dualitätssatz 4.1.7 sind die beiden Maxima genau dann gleich, wenn die zugehörigen dualen Aufgaben gleiche Minima besitzen. Nach der Dualitätstabelle 4.1.11 ist das Duale der um den möglichen Zukauf erweiterten Aufgabe

$$\begin{aligned} \min b^T y \\ A^T y &\geq c \\ y &= s \\ y &\geq 0. \end{aligned}$$

Die beiden Maxima stimmen also genau dann überein, wenn es einen für das Duale der ursprünglichen Produktionsaufgabe optimalen Vektor y^* gibt, der auch $y^* = s$ erfüllt. Insgesamt folgt damit die Behauptung. \square

Literaturverzeichnis

- [Aig04] AIGNER, M.: *Diskrete Mathematik*. Vieweg Studium: Aufbaukurs Mathematik. Vieweg, 5. Auflage, 2004.
- [Alt02] ALT, W.: *Nichtlineare Optimierung*. Vieweg Studium: Aufbaukurs Mathematik. Vieweg, 2002.
- [Alt04] ALT, W.: *Numerische Verfahren der konvexen, nichtglatten Optimierung*. Teubner, 2004.
- [Bar02] BARVINOK, A.: *A course in convexity*. American Mathematical Society, Providence, RI, 2002.
- [Bea96] BEASLEY, J.E. (Herausgeber): *Advances in linear and integer programming*. Oxford lecture series in mathematics and its applications. Oxford University Press, 1996.
- [Bor01] BORWARDT, K.H.: *Optimierung, Operations Research, Spieltheorie: Mathematische Grundlagen*. Birkhäuser, 2001.
- [BTN01] BEN-TAL, A. und A. NEMIROVSKI: *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MPS-SIAM Series on Optimization, 2001.
- [BW05] BERTSIMAS, D. und R. WEISMANTEL: *Optimization over Integers*. Dynamic Ideas, Belmont, Mass., 2005.
- [Cas00] CASSELS, J.W.S.: *An Introduction to the Geometry of Numbers*. Springer, 2nd Printing Auflage, 2000.
- [Chv83] CHVÁTAL, V.: *Linear programming*. Freeman, 1983.
- [Die00] DIESTEL, R.: *Graphentheorie*. Springer, 2nd Auflage, 2000.
- [DS06] DEMPE, S. und H. SCHREIER: *Operations Research: Deterministische Modelle und Methoden*. Teubner, 2006.
- [GJ79] GAREY, M.R. und D.S. JOHNSON: *Computers and intractability: A guide to the theory of NP-completeness*. Freeman, 1979.
- [GK99] GEIGER, C. und C. KANZOW: *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben*. Springer, 1999.
- [GK02] GEIGER, C. und C. KANZOW: *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer, 2002.
- [GL87] GRUBER, P.M. und C.G. LEKKERKERKER: *Geometry of Numbers*. North-Holland, 1987.

-
- [GLS93] GRÖTSCHEL, M., L. LOVÁSZ und A. SCHRIJVER: *Geometric algorithms and combinatorial optimization*, Band 2 der Reihe *Algorithms and Combinatorics*. Springer, 2. Auflage, 1993.
- [Gr"93] GRÜNBAUM, B.: *Convex Polytopes*. Springer, 2. Auflage, 1993.
- [GT97] GROSSMANN, CH. und J. TERNO: *Numerik der Optimierung*. Teubner, 2. Auflage, 1997.
- [HPT95] HORST, R., P.M. PARDALOS und N. THOAI: *Introduction to Global Optimization*. Nonconvex Optimization and Applications. Kluwer, 1995.
- [JS04] JARRE, F. und J. STOER: *Optimierung*. Springer, 2004.
- [Kel99] KELLEY, C.T.: *Iterative Methods for Optimization*. SIAM, 1999.
- [KN05] KRUMKE, S. O. und H. NOLTEMEIER: *Graphentheoretische Konzepte und Algorithmen*. Teubner, 2005.
- [KV06] KORTE, B. und J. VYGEN: *Combinatorial Optimization*, Band 21 der Reihe *Algorithms and Combinatorics*. Springer, 3. Auflage, 2006.
- [LLKe] LAWLER, E.L., J.K. LENSTRA, A.H.G. RINNOOY KAN und D. SMOYS (EDS.).
- [LPV05] LOVÁSZ, L., J. PEKIKÁN und K. VESZTERGOMBI: *Diskrete Mathematik*. Springer, 2005.
- [Min86] MINOUX, M.: *Mathematical programming: Theory and Algorithms*. Wiley–interscience publications. Wiley, 1986.
- [NN93] NESTEROV, Y. und A. NEMIROVSKI: *Interior-point polynomial algorithms in convex programming*. SIAM Studies in Applied Mathematics, 1993.
- [NW88] NEMHAUSER, G. L. und L. A. WOLSEY: *Integer and combinatorial optimization*. Wiley–interscience series in discrete mathematics and optimization. Wiley, 1988.
- [PGW81] P.E. GILL, W. MURRAY und M.H. WRIGHT: *Practical Optimization*. Academic Press, 1981.
- [PS82] PAPADIMITRIOU, C.H. und K. STEIGLITZ: *Combinatorial optimization: Algorithms and complexity*. Prentice–Hall, 1982.
- [PS02] PRÖMEL, H.J. und A. STEGER: *The Steiner Tree Problem*. Vieweg, 2002.
- [Rei94] REINELT, G.: *The Traveling Salesman: Computational Solutions for TSP Applications*. Springer, Lecture Notes in Computer Science 840, 1994.
- [Ren01] RENEGAR, J.: *A Mathematical View of Interior-Point Methods in Convex Optimization*. MPS-SIAM Series on Optimization, 2001.
- [Roc72] ROCHAFELLAR, R.T.: *Convex Analysis*. Princeton University Press, 2. print. Auflage, 1972.

- [Sch86] SCHRIJVER, A.: *Theory of Linear and Integer Programming*. Wiley–Interscience series in discrete mathematics and optimization. Wiley, first edition Auflage, 1986.
- [Sch93] SCHNEIDER, R.: *Convex Bodies: The Brunn-Minkowski Theory*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1993.
- [Sch03] SCHRIJVER, A.: *Combinatorial Optimization A – C*. Algorithms and Combinatorics 24. Springer, first edition Auflage, 2003.
- [Van97] VANDERBEI, R.J.: *Linear programming: Foundations and extensions*. Int. Ser. Oper. Res. Manag. Sci. Kluwer, 2. print. Auflage, 1997.
- [Wan06] WANKA, R.: *Approximationsalgorithmen*. Teubner, 2006.
- [Ye97] YE, Y.: *Interior point algorithms: Theory and analysis*. Wiley–Interscience Series in Discrete Mathematics and Optimization. Wiley, 1997.
- [Zie95] ZIEGLER, G.M.: *Lectures on polytopes*, Band 152 der Reihe *Graduate texts in mathematics*. Springer, 2nd Auflage, 1995.

Index

- $H_{(a,\beta)}^{<}$, *siehe* Halbraum
 $H_{(a,\beta)}^{>}$, *siehe* Halbraum
 $H_{(a,\beta)}^{\geq}$, *siehe* Halbraum
 $H_{(a,\beta)}^{\leq}$, *siehe* Halbraum
 $I(x)$, *siehe* aktive Nebenbedingungen
 $N_C(x^*)$, *siehe* Kegel der äusseren Normalen,
siehe Stutzkegel
 \mathcal{H} -Darstellung, 4
 \mathcal{V} -Darstellung, 39
— irredundant, 39
 \mathcal{V} -Polytop, 16
ext, *siehe* Extrempunkt
ls, *siehe* Linearitätsraum
rec, *siehe* Rezessionskegel
 $d(\cdot, \cdot)$, *siehe* Abstand
k-Seite, 31
- Abbildung
— affin, 16
abhängig
— linear, 15
Abschluß, 19
Abstand, 19
aff, *siehe* affine Hülle
affin abhängig, 15
affin unabhängig, 15
affine Hülle, 16
affiner Unterraum, 15
Affinkombination, 15
aktive Nebenbedingungen, 29
Alternativsatz, 24
Aufgabe
— Produktionsplanung, 5
— Wire Spacing, 8
- bd, *siehe* Rand
- Carathéodory, 19
cl, *siehe* Abschluss
conv, *siehe* konvexe Hülle
- Dilatation, 15
dim, *siehe* Dimension, *siehe* Dimension
Dimension, 15, 20
- Ebene, 15
Ecke, 36
Einheitskugel, *siehe* Kugel
Einheitssphäre, *siehe* Sphäre
Einheitsvektoren
— Standard, 14
- ext, *siehe* Extrempunkt
Extrempunkt, 31
Extremalstrahl, 31
- Facette, 36
Farkas, 30
— Lemma von, 24, 30
Fourier-Motzkin Eliminationsverfahren, 12
Fourier-Motzkin-Elimination, 12
- Gerade, 15
— Halb-, 22
- Hülle
— konische, 25
— konvex, 16
— positive, 25
Halbgerade, 22
Halbraum, 21
— abgeschlossener, 21, 22
— offen, 22
— Stützhilbraum, 24
Halbunterraum, 22
Hülle
— affin, 16
— linear, 16
Hyperebene, 15
— eigentliche Stützhyperebene, 24
— Stützhyperebene, 24
— streng trennende, 22
— strikt trennende, 22
— trennende, 22
— uneigentliche Stützhyperebene, 24
- ILP-Problem, 4
Innenkegel, 27
Inneres, 19
— relatives, 20
int, *siehe* Inneres
irredundant, *siehe* \mathcal{V} -Darstellung oder \mathcal{H} -Darstellung
- Kante, 36
Kegel, 25
— der äusseren Normalen, 27
— Innen-, 27
— polyederisch, 25
— Stütz-, 27
— Tangential-, 27
konische Kombination, 26
konvex, 16
konvexe Hülle, 16
Konvexkombination, 16

- Kreupolytop, 18
- Kreuzpolytop
 - Einheits-, 18
 - regular, 18
 - Standard-, 18
- Kugel
 - Einheitskugel, euklidisch, 19

- Lagrange-Multiplikatoren, 9
- Lemma
 - von Farkas, 30
 - von Farkas, 24
- lin, *siehe* lineare Hülle
- Linealitätsraum, 33
- lineare Hülle, 16
- lineare Optimierungsaufgabe, 3
 - natürliche Form, 3
- LP-Problem, 4
- ls, *siehe* Linealitätsraum

- Maximierungsaufgabe
 - linear-konvex, 21
- Maximierungsproblem
 - linear-konvex, 21
- Maximierungsaufgabe, 3
- Maximierungsproblem, 3
- MILP-Problem, 4
- Minimierungsaufgabe, 3
 - linear-konvex, 21
- Minimierungsproblem, 3
 - linear-konvex, 21
- Minkowski, 35
- Minkowski-Summe, 15

- nearest-point-map, 24
- Nebenbedingungen
 - aktiv, 29
- Nichtnegativkombination, 26
- Norm
 - euklidisch, 19
- Normale, 21
 - äußere, 21
- Normalenkegel, 27
- Normalenvektor, 21
 - äußerer, 21

- Optimalpunkt, 3
- Optimierung
 - restringiert, 3
 - unrestringiert, 3
- Optimierungsaufgabe
 - allgemein, 3
 - ganzzahlig, 4
 - gemischt-ganzzahlig, 4
 - linear, 3
 - linear-konvex, 21
 - unzulässig, 3
- Optimierungsproblem, 3
 - ganzzahlig, 4
 - gemischt-ganzzahlig, 4
 - linear, 4
 - linear-konvex, 21

- Parallelotop, 18
- Polyeder, 4
 - \mathcal{H} -, 4
 - \mathcal{H} -Darstellung eines, 4
 - \mathcal{V} -, 39
 - \mathcal{V} -Darstellung, 39
 - geradenfrei, 33
 - irredundante \mathcal{V} -Darstellung, 39
- Polytop, 4
 - \mathcal{H} -, 4
 - \mathcal{H} -Darstellung eines, 4
 - \mathcal{V} -Polytop, 16
- pos, *siehe* Hülle, positive Problem
 - Wire Spacing, 8
- Produktionsaufgabe, 5
 - Preisvektor, 5
 - Technologiematrix, 5
 - verfügbare Ressourcen, 5
- Produktionsplan, 5
 - optimal, 5
 - zulässig, 5
- Produktionsplanung, 4
- Produktionsproblem, 5
- Prozedur
 - Fourier-Motzkin-Elimination, 12
- Punkt, 15

- Rand, 19
 - relativ, 20
- Raum
 - Halbunter-, 22
- rec, *siehe* Rezessionskegel
- Rezessionskegel, 33
- Rundreiseproblem, 7

- Satz
 - von Carathéodory, 19
 - von Minkowski, 35
- Seite, 31
 - k -, 31
 - eigentlich, 31
 - uneigentlich, 31
 - von C , 31
 - von x in C , 31
- Simplex, 18
 - Standard-, 18

Sphere

— Einheitssphäre, euklidisch, 19

Stützhalbraum, 24

Stützhyperebene, 24

— eigentlich, 24

— uneigentlich, 24

Stützkegel, 27

Stützpunkt, 24

Standardeinheitsvektoren, 14

Standardsimplex, 18

Strahl, 22

— Extremal-, 31

Tangentialkegel, 27

Transformation

— affin, 16

Traveling Salesman Problem, 7

trennen, *siehe* Hyperebene, trennende

TSP, 7

unabhängig

— affin, 15

Unterraum

— affiner, 15

Würfel, 18

— Einheitswürfel, 18

— Standardwürfel, 18

Wire Spacing Aufgabe, 8

Wire Spacing Problem, 8

Zielfunktion, 3

Zulässigkeitsaufgabe, 3

Zulässigkeitsproblem, 3

zulässiger Bereich, 3

zulässiger Punkt, 3