KAPITEL 5 — POLYEDERTHEORIE

F. VALLENTIN, A. GUNDERT

Viele Optimierungsprobleme des Operations Research lassen sich als *lineares Programme* formulieren:

$$
\begin{array}{ll}
\underline{\text{gegeben}}: & c \in \mathbb{R}^n, \\
& a_1, \ldots, a_m \in \mathbb{R}^n, \\
& b_1, \ldots, b_m \in \mathbb{R}
\end{array}
$$

$$
\begin{array}{lll}
\underline{\text{gesucht}}: & \min & c^T x \\
& & x \in \mathbb{R}^n \\
& & a_1^T x \leq b_1, \ldots, a_m^T x \leq b_m.
\end{array}
$$

c definiert die *Zielfunktion* $\quad x \mapsto c^T x$,

$a_j^T x \leq b_j$ sind lineare Ungleichungen, welche die *Nebenbedingungen* beschreiben.

Kurzform: $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ mit

$$
A = \begin{bmatrix} — & a_1^T & — \\ & \vdots & \\ — & a_m^T & — \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \quad \text{und} \quad Ax \leq b
$$

Manchmal kommt noch eine Ganzzahligkeitsnebenbedingung $x \in \mathbb{Z}^n$ hinzu ($\rightarrow$ spätere Kapitel).

Dieses Kapitel: Wie sieht die Menge der *zulässigen Lösungen* (die Vektoren, die die Nebenbedingungen erfüllen)

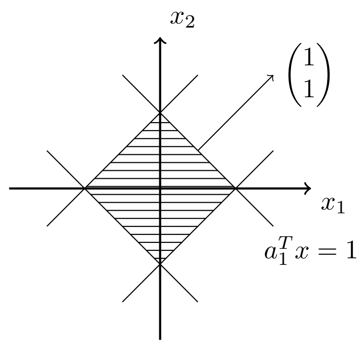
$$\{x \in \mathbb{R}^n : Ax \leq b\}$$

geometrisch aus?

Beispiel 0.1.

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

Date: 29. Juni 2014.



Dies ist ein (konvexes) Polyeder.

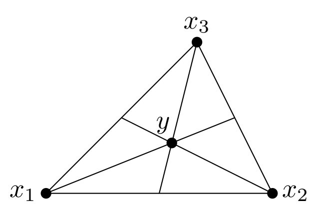
Polyedertheorie: Erweiterung der linearen Algebra über dem Körper $\mathbb{R}$ durch lineare Ungleichungen.

1. Konvexe Mengen

Definition 1.1. *Seien $x_1, \ldots, x_N \in \mathbb{R}^n$ Punkte. Dann ist $y \in \mathbb{R}^n$ eine* Konvexkombination *von $x_1, \ldots, x_N$, falls*

$$y = \sum_{i=1}^{N} \alpha_i x_i \quad \textit{mit } \alpha_1, \ldots, \alpha_N \geq 0 \quad \textit{und } \sum_{i=1}^{N} \alpha_i = 1.$$

Beispiel 1.2. *Der Schwerpunkt eines Dreiecks ist eine Konvexkombination der drei Eckpunkte.*



$y = \frac{1}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_2$

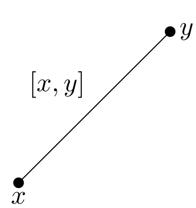
Definition 1.3. *Eine Menge $C \subseteq \mathbb{R}^n$ heißt* konvex*, wenn sie unter der Bildung von Konvexkombinationen abgeschlossen ist, das heißt*

$$\forall N \in \mathbb{N}\ \forall x_1, \ldots, x_N \in C\ \forall \alpha_1, \ldots, \alpha_N \geq 0,\ \sum_{i=1}^{N} \alpha_i = 1 \quad \Rightarrow \sum_{i=1}^{N} \alpha_i x_i \in C.$$

Definition 1.4. *Seien $x, y \in \mathbb{R}^n$ Punkte, dann ist*

$$[x, y] = \{(1-\alpha)x + \alpha y : \alpha \in [0, 1]\}$$

die Verbindungsstrecke *zwischen x und y.*



Satz 1.5. *Eine Menge $C \subseteq \mathbb{R}^n$ ist konvex* $\Leftrightarrow$ $\forall x, y \in C : [x, y] \subseteq C$.

Beweis.

„$\Rightarrow$“: Klar, denn dies ist ein Spezialfall der Definition mit $N = 2$.

„$\Leftarrow$“: Zeige per Induktion nach N: Jede Konvexkombination von höchstens N Punkten aus C liegt wieder in C.

$N = 1$: ✓

$N \to N+1$: Sei $y = \sum_{i=1}^{N+1} \alpha_i x_i$ mit $x_i \in C$, $\alpha_i \geq 0$ und $\sum_{i=1}^{N+1} \alpha_i = 1$.

1.Fall: $\alpha_{N+1} = 1$.
Dann ist $\alpha_1 = \ldots = \alpha_N = 0$ und $y = x_{N+1} \in C$.

2.Fall: $\alpha_{N+1} \neq 1$.
Nach Induktionsvoraussetzung ist

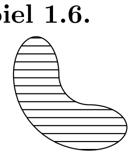
$$y' = \sum_{i=1}^{N} \frac{\alpha_i}{1-\alpha_{N+1}} x_i \in C, \quad \text{weil } \frac{\alpha_i}{1-\alpha_{N+1}} \geq 0 \text{ und } \sum_{i=1}^{N} \frac{\alpha_i}{1-\alpha_{N+1}} = 1$$

Desweiteren ist

$$y = (1-\alpha_{N+1})y' + \alpha_{N+1} x_{N+1} \in [y', x_{N+1}] \subseteq C.$$

□

Beispiel 1.6.

a)  *ist keine konvexe Menge.*

b) Sei $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ *eine Norm. Sei*

$$K = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$$

die zugehörige Einheitskugel.

Behauptung: K ist konvex.

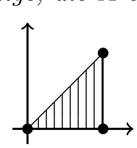
Beweis: Seien $x, y \in K$, $\alpha \in [0,1]$. Dann ist

$$\|(1-\alpha)x + \alpha y\| \leq (1-\alpha)\|x\| + \alpha\|y\| \leq (1-\alpha) \cdot 1 + \alpha \cdot 1 = 1.$$

Definition 1.7. *Sei $A \subseteq \mathbb{R}^n$. Die* konvexe Hülle *von A ist*

$$\mathit{conv}\, A = \bigcap_{\substack{B \supseteq A \\ B \text{ konvex}}} B.$$

Da der Durchschnitt von konvexen Mengen wieder konvex ist, handelt es sich bei conv A um die inklusionsminimale konvexe Menge, die A enthält.

Beispiel 1.8. $A = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$  *conv*(A)

Satz 1.9. *Sei $A \subseteq \mathbb{R}^n$. Dann gilt*

$$\text{conv } A = \{y \in \mathbb{R}^n : \exists N \in \mathbb{N}\ \exists x_1, \ldots, x_N \in A\ \exists \alpha_1, \ldots, \alpha_N \geq 0, \sum_{i=1}^{N} \alpha_i = 1, y = \sum_{i=1}^{N} \alpha_i x_i\}.$$

Das heißt, conv A besteht aus allen Konvexkombinationen der Menge A.

Beweis.

„$\subseteq$“: trivial.
„$\supseteq$“: Sei B konvex mit $B \supseteq A$. Da B abgeschlossen unter der Bildung von Konvexkombinationen ist, muss für jede Wahl von $N \in \mathbb{N},\ x_1, \ldots, x_N \in A,\ \alpha_1, \ldots, \alpha_N \geq 0, \sum_{i=1}^{N} \alpha_i = 1$ stets $y = \sum_{i=1}^{N} \alpha_i x_i \in B$ gelten.

□

Definition 1.10. *Eine Menge $P \subseteq \mathbb{R}^n$ heißt* (konvexes) Polytop*, falls es eine endliche Menge $A = \{x_1, \ldots, x_N\}$ gibt mit*

$$\begin{aligned} P &= \text{conv } A \\ &= \{\sum_{i=1}^{N} \alpha_i x_i : \alpha_i \geq 0, \sum_{i=1}^{N} \alpha_i = 1\} \end{aligned}$$

Sei $y \in \text{conv } A$, das heißt

$$y = \sum_{i=1}^{N} \alpha_i x_i \text{ für } N \in \mathbb{N},\ x_1, \ldots, x_N \in A,\ \alpha_1, \ldots, \alpha_N \geq 0, \sum_{i=1}^{N} \alpha_i = 1.$$

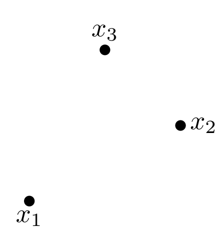
Frage: Wie groß muss N höchstens sein, damit man jedes y in conv A darstellen kann?

Definition 1.11. *Die Punkte $x_1, \ldots x_N \in \mathbb{R}^n$ sind* affin unabhängig*, falls*

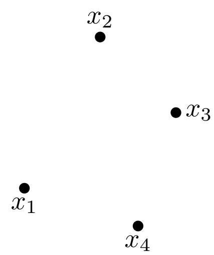
$$\forall \alpha_1, \ldots, \alpha_N : \sum_{i=1}^{N} \alpha_i = 0,\ \sum_{i=1}^{N} \alpha_i x_i = 0\ \Rightarrow \alpha_1 = \ldots = \alpha_N = 0.$$

In anderen Worten: Die Vektoren $\begin{bmatrix} 1 \\ x_1 \end{bmatrix}, \ldots, \begin{bmatrix} 1 \\ x_N \end{bmatrix} \in \mathbb{R}^{n+1}$ sind linear unabhängig.

Beispiel 1.12. *Drei Punkte in der Ebene, die nicht alle auf der selben Gerade liegen, sind affin unabhängig.*



Vier Punkte in der Ebene sind hingegen nie affin unabhängig.



Definition 1.13. *Sei $A \subseteq \mathbb{R}^n$. Die* Dimension *von A ist*

$$\textit{dim } A = \max\{N-1 : \exists x_1, \ldots, x_N \in A \text{ affin unabhängig}\}.$$

Es ist immer dim $A \leq n$, weil rank $\begin{bmatrix} 1 & \ldots & 1 \\ x_1 & \ldots & x_N \end{bmatrix} \leq n+1$.

Satz 1.14. *(Caratheodory)*
Sei $A \subseteq \mathbb{R}^n$ und $y \in$ conv A. Dann existieren $x_1, \ldots, x_N \in A$ affin unabhängig mit $y \in$ conv $\{x_1, \ldots, x_N\}$. Insbesondere ist $N \leq$ dim$(A)+1 \leq n+1$.

Beweis. Sei $y \in \operatorname{conv} A$ mit

$$y = \sum_{i=1}^{N} \alpha_i x_i, \quad x_i \in A,\ \alpha_i \geq 0,\ \sum_{i=1}^{N} \alpha_i = 1.$$

Angenommen N ist minimal und $x_1, \ldots, x_N$ sind nicht affin unabhängig. Dann gibt es $\beta_1, \ldots, \beta_N \in \mathbb{R}$ mit $\sum_{i=1}^N \beta_i = 0,\ \sum_{i=1}^N \beta_i x_i = 0$ und $(\beta_1, \ldots, \beta_N) \neq (0, \ldots, 0)$. Sei oBdA. $\beta_1 > 0$ und $\frac{\alpha_1}{\beta_1}$ so klein wie möglich. Betrachte die Darstellung

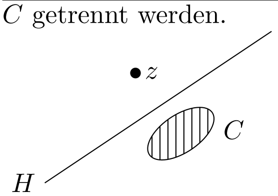
$$y = \sum_{i=1}^{N} \alpha_i x_i - \frac{\alpha_1}{\beta_1} \sum_{i=1}^{N} \beta_i x_i = \sum_{i=2}^{N} \gamma_i x_i, \quad \gamma_i = \alpha_i - \frac{\alpha_1}{\beta_1} \beta_i.$$

Da $\gamma_i \geq 0$ (klar, wenn $\beta_i \leq 0$; sonst wegen der Minimalität von $\frac{\alpha_1}{\beta_1}$) und $\sum_{i=2}^N \gamma_i = 1$, ist die Darstellung ein Widerspruch zur Minimalität von N. □

2. Trenn- und Stützhyperebenen

Sei $C \subseteq \mathbb{R}^n$ eine abgeschlossene, konvexe Menge.

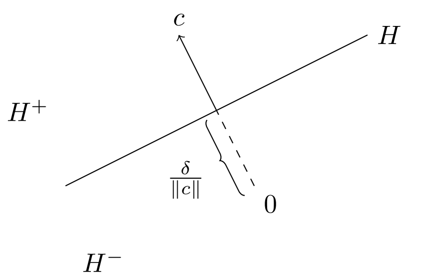
Fundamentale Eigenschaft: Jeder Punkt $z \notin C$ kann durch eine Hyperebene von C getrennt werden.



(Wichtige Verallgemeinerung in der Funktionalanalysis: Satz von Hahn-Banach)

Definition 2.1. *Eine Teilmenge $H \subseteq \mathbb{R}^n$ heißt (*affine*)* Hyperebene*, falls es einen Vektor $c \in \mathbb{R}^n \backslash \{0\}$ und ein $\delta \in \mathbb{R}$ gibt mit*

$$H = \{x \in \mathbb{R}^n : c^T x = \delta\}$$

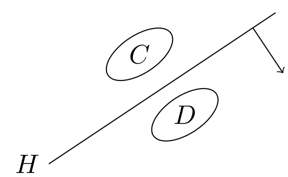


Die abgeschlossenen und konvexen Mengen

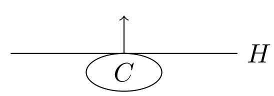
$$H^+ = \{x \in \mathbb{R}^n : c^T x \geq \delta\}$$
$$H^- = \{x \in \mathbb{R}^n : c^T x \leq \delta\}$$

heißen Halbräume

Definition 2.2. *Seien $C, D \subseteq \mathbb{R}^n$. Eine Hyperebene H heißt* Trennhyperebene *von C und D, falls $C \subseteq H^-$ und $D \subseteq H^+$ (oder umgekehrt).*

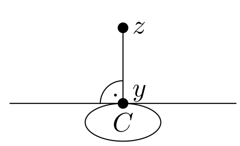


Definition 2.3. *Sei $C \subseteq \mathbb{R}^n$. Eine Hyperebene H heißt* Stützhyperebene *von C, falls $C \subseteq H^-$ und $C \cap H \neq \emptyset$.*



Frage: Wie findet man Trenn- bzw. Stützhyperebenen?

Antwort: Verwende *metrische Projektion*
(= beste Approximation von $z \notin C$ in C
= orthogonale Projektion, falls C affin linear ist).

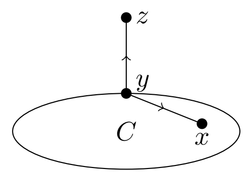


Lemma 2.4. *Sei $C \subseteq \mathbb{R}^n$ eine abgeschlossene und konvexe Menge, die nicht leer ist. Sei $z \notin C$. Dann*

$$\exists! y \in C : \quad \|y - z\| = \inf_{x \in C} \|x - z\| \quad \text{(Notation: } y = \pi_C(z))$$

und

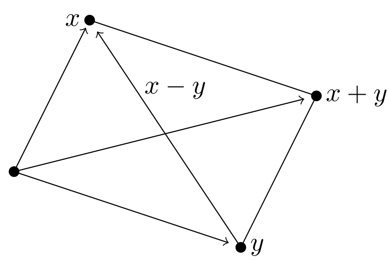
$$(z - y)^T (x - y) \leq 0 \quad \text{für alle } x \in C.$$



Beweis. (funktioniert auch in allgemeinen Hilberträumen)

Zur Erinnerung: Parallelogrammgleichung

Für $x, y \in \mathbb{R}^n$ gilt $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$.



Existenz von y: Sei $d = \inf_{x \in C} \|x - z\|$. OBdA ist $z = 0$ (durch Translation). Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge aus C mit $\lim_{n\to\infty} \|x_n\| = d$. Wir zeigen, dass $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist. Nach der Parallelogrammgleichung ist

$$\underbrace{\|\underbrace{\frac{x_n + x_m}{2}}_{\in C}\|^2}_{\Rightarrow \geq d^2} + \|\underbrace{\frac{x_n - x_m}{2}}_{\Rightarrow \to 0}\|^2 = \underbrace{\frac{1}{2}\|x_n\|^2 + \frac{1}{2}\|x_m\|^2}_{\to d^2}$$

Der zweite Summand auf der linken Seite muss gegen 0 konvergieren, da die rechte Seite gegen d^2 konvergiert und der erste Summand auf der linken Seite sicher nicht kleiner als d^2 ist (denn $d = \inf_{x\in C} \|x - z\|$). Also ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge und weil $\mathbb{R}^n$ vollständig ist, konvergiert $(x_n)_{n\in\mathbb{N}}$. Weil C abgeschlossen ist, liegt der Grenzwert $\lim_{n\to\infty} x_n$ ebenfalls in C. Dieser Grenzwert ist ein y mit der gewünschten Eigenschaft.

Eindeutigkeit von y: Angenommen es existieren $y, y' \in C$ mit $\|y\| = \|y'\| = \inf_{x\in C} \|x\|$ und $y \neq y'$. Dann ist

$$\|\underbrace{\frac{y+y'}{2}}_{\in C}\|^2 < \|\frac{y+y'}{2}\|^2 + \|\frac{y-y'}{2}\|^2 = \frac{1}{2}\|y\|^2 + \frac{1}{2}\|y'\|^2 = d^2,$$

was ein Widerspruch zur Minimalität von d ist.

Zusatz: $(z-y)^T(x-y) \leq 0$ für alle $x \in C$.

Sei $\alpha \in [0,1]$. Dann ist

$$\begin{aligned}\|z-y\|^2 &\leq \|z-((1-\alpha)y+\alpha x)\|^2\\ &= \|z-y+\alpha(y-x)\|^2\\ &= \|z-y\|^2 + 2\alpha(z-y)^T(y-x) + \alpha^2\|y-x\|^2\end{aligned}$$

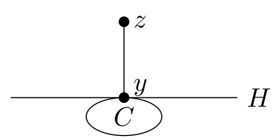
Demnach gilt für $\alpha \in (0,1]$

$$(z-y)^T(x-y) \leq \frac{\alpha}{2}\|y-x\|^2$$

und die Behauptung folgt, da die rechte Seite beliebig klein ≥ 0 werden kann. $\square$

Theorem 2.5. *Sei $C \in \mathbb{R}^n$ eine nichtleere, abgeschlossene und konvexe Menge. Sei $z \notin C$ ein Punkt außerhalb von C. Dann gibt es eine Trennhyperebene von $\{z\}$ und C.*

Beweis. Definiere die Hyperebene $H = \{x \in \mathbb{R}^n : c^T x = \delta\}$ mit $c = z - y$ und $\delta = c^T y$, wobei $y = \pi_C(z)$.



Behauptung: a) $z \in H^+$
b) $C \subseteq H^-$.

zu a)

$$\begin{aligned}&c^T z \geq \delta\\ \Leftrightarrow &(z-y)^T z \geq (z-y)^T y\\ \Leftrightarrow &(z-y)^T(z-y) \geq 0. \quad \checkmark\end{aligned}$$

zu b) Für $x \in C$ ist

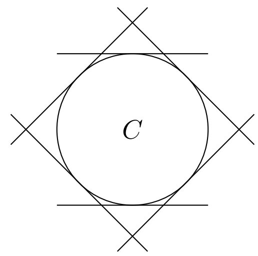
$$\begin{aligned}&c^T x \leq \delta\\ \Leftrightarrow &(z-y)^T x \leq (z-y)^T y\\ \Leftrightarrow &(z-y)^T(x-y) \leq 0. \quad \checkmark\end{aligned}$$

$\square$

Bemerkung 2.6. *Die im Beweis konstruierte Hyperebene ist eine Stützhyperebene von C. Falls $\delta \in (c^T y, c^T z)$ gewählt wäre, wäre H eine strikte Trennhyperebene von $\{z\}$ und C, das heißt $z \in H^+$ und $z \notin H$, $C \subseteq H^-$ und $C \cap H = \emptyset$.*

Korollar 2.7. *Sei $C \subseteq \mathbb{R}^n$ eine nichtleere, abgeschlossene und konvexe Menge. Dann gilt*

$$C = \bigcap_{H \text{ Stützhyperebene von } C} H^-$$



Beweis. → Aufgabe 7.1 □

Definition 2.8. *Eine Menge $P \subseteq \mathbb{R}^n$ heißt* Polyeder*, falls es eine Matrix $A \in \mathbb{R}^{m \times n}$ und einen Vektor $b \in \mathbb{R}^m$ gibt, so dass*

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

gilt.

Ein Polyeder entspricht also der Lösungsmenge eines Systems linearer Ungleichungen, was der Durchschnitt endlich vieler Halbräume ist (im Korollar 2.7 genügen endlich viele Stützhyperebenen, um ein Polyeder vollständig zu beschreiben).

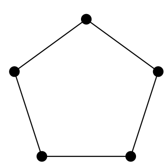
Jetzt: Verhältnis zwischen Polyedern und Polytopen (konvexe Hülle endlich vieler Punkte).

3. Extrempunkte und Ecken

Definition 3.1. *Sei $C \subseteq \mathbb{R}^n$ eine konvexe Menge. Ein Punkt $z \in C$ heißt* Extrempunkt *von C, falls für alle $x, y \in C$ mit $z = \alpha x + (1-\alpha)y$ und $\alpha \in (0,1)$ stets $x = z = y$ folgt.*

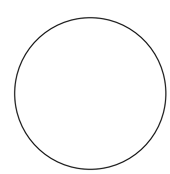
Beispiel 3.2. *:*

a)



5 Extrempunkte

b)



∞- viele Extrempunkte

Sprechweise: Extrempunkte von Polyedern heißen *Ecken.*

Notation 3.3. *Sei $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ ein Polyeder und $z \in P$. Mit A_z bezeichnen wir die Teilmatrix von A, die die Zeilen enthält mit $a_i^T z = b_i$ („*aktive Ungleichungen*“).*

Satz 3.4. *Sei $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ ein Polyeder und $z \in P$. Dann gilt*
z ist eine Ecke von $P \Leftrightarrow$ rang $A_z = n$.

Beweis.

„⇒“: Angenommen rang $A_z < n$. Dann existiert ein $c \in \mathbb{R}^n$, $c \neq 0$ mit $A_z c = 0$. Da $a_i^T z < b_i$ für alle Zeilen von A gilt, die nicht zu A_z gehören, gibt es ein $\delta > 0$ mit

$$a_i^T(z + \delta c) \leq b_i \quad \text{und} \quad a_i^T(z - \delta c) \leq b_i.$$

Also: $A(z+\delta c) \leq b$ und $A(z-\delta c) \leq b$ und $z+\delta c,\ z-\delta c \in P$. Das heißt $z = \frac{1}{2}(z+\delta c) + \frac{1}{2}(z-\delta c)$ und somit ist z keine Ecke von P.

„$\Leftarrow$“: Es sei $z = \alpha x + (1-\alpha)y$ mit $x, y \in P,\ \alpha \in (0,1)$. Für eine Zeile a_i^T von A_z gilt

$$\begin{aligned} b_i &= a_i^T z \\ &= a_i^T(\alpha x + (1-\alpha)y) \\ &= \alpha a_i^T x + (1-\alpha) a_i^T y \\ &\leq \alpha b_i + (1-\alpha) b_i = b_i \end{aligned}$$

Also $a_i^T x = a_i^T y = b_i$, da $\alpha \in (0,1)$. Da rang $A_z = n$, hat das System $a_i^T w = b_i$, das aus den Zeilen von A_z besteht, eine eindeutige Lösung. Das heißt $x = z = y$ und z ist eine Ecke von P.

□

Korollar 3.5. *Ein Polyeder $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ mit $A \in \mathbb{R}^{m \times n}$ hat höchstens $\binom{m}{n}$ Ecken.*

Theorem 3.6. *(Minkowski)*
Ein beschränktes Polyeder ist ein Polytop

Beweis. Seien $x_1, \ldots, x_t$ die Ecken des beschränkten Polyeders

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

Behauptung: $P = \operatorname{conv}\{x_1, \ldots, x_t\}$.

Beweis:

„$\supseteq$“: klar.

„$\subseteq$“: Sei $z \in P$. Wir zeigen: $z \in \operatorname{conv}\{x_1, \ldots, x_t\}$ per Induktion nach $n-$ rang(A_z).

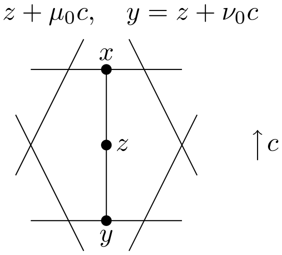
Induktionsanfang: $n-$ rang$(A_z) = 0$.
Das heißt rang$(A_z) = n \Rightarrow z$ ist eine Ecke von P, das heißt $z \in \{x_1, \ldots, x_t\}$.

Induktionsschluss: $n-$ rang$(A_z) > 0$.
Das heißt rang$(A_z) < n \Rightarrow \exists c \neq 0 : A_z c = 0$. Sei

$$\begin{aligned} \mu_0 &= \max\{\mu : z + \mu c \in P\} \\ \nu_0 &= \min\{\nu : z + \nu c \in P\}. \end{aligned}$$

μ_0 und ν_0 existieren, weil P kompakt ist. Definiere

$$x = z + \mu_0 c, \quad y = z + \nu_0 c$$



Es gilt

$$(*)\qquad \mu_0 = \min\left\{ \frac{b_i - a_i^T z}{a_i^T c} : a_i^T \text{ Zeile von } A, a_i^T c > 0 \right\},$$

weil $a_i^T(z+\mu c) \le b_i \Leftrightarrow \mu \le \frac{b_i - a_i^T z}{a_i^T c}$ und μ_0 ist das größte solche μ. In $(*)$ sei i_0 eine Zeile von A, in der das Minimum angenommen wird. Dann ist

$$(1)\qquad A_z x = A_z z + \mu_0 A_z c = A_z z$$

$$(2)\qquad a_{i_0}^T x = a_{i_0}^T (z + \mu_0 c) = b_{i_0}.$$

Somit enthält A_x alle Zeilen von A_z und außerdem die Zeile $a_{i_0}^T$. Da $A_z c = 0$, aber $a_{i_0}^T c \neq 0$, gilt $\text{rang}(A_x) > \text{rang}(A_z)$. Nach Induktionsvoraussetzung ist $x \in \text{conv}\{x_1, \ldots, x_t\}$. Genauso folgt $y \in \text{conv}\{x_1, \ldots, x_t\}$. Da $z \in [x, y]$ ist, folgt die Behauptung $z \in \text{conv}\{x_1, \ldots, x_t\}$.

□

Definition 3.7. *Sei $A \subseteq \mathbb{R}^n$. Dann heißt*

$$A^* = \{y \in \mathbb{R}^n : x^T y \le 1 \; \forall x \in A\}$$

die Polare *von A.*

Lemma 3.8.

a) Für $\alpha > 0$ gilt $(\alpha A)^* = \frac{1}{\alpha} A^*$, wobei $\beta B = \{\beta b : b \in B\}$ mit $\beta \in \mathbb{R}$, $B \subseteq \mathbb{R}^n$.
b) $A \subseteq B \Rightarrow A^* \supseteq B^*$.
c) $(B_n)^* = B_n$, wobei $B_n = \{x \in \mathbb{R}^n : x^T x \le 1\}$ die n-dimensionale Einheitskugel ist.
d) $P = conv\{x_1, \ldots, x_t\} \Rightarrow P^* = \{y \in \mathbb{R}^n : x_1^T y \le 1, \ldots, x_t^T y \le 1\}$.

Beweis. a),b) → Aufgabe 7.4.

c) „$\subseteq$“: Sei $y \in (B_n)^*$. Falls $y = 0$, dann $y \in B_n$. Falls $y \neq 0$, dann $y^T \frac{1}{\|y\|} y \le 1$. Also $\|y\| \le 1$, das heißt $y \in B_n$.

„$\supseteq$“: Seien $x, y \in B_n$. Dann gilt (nach Cauchy-Schwarz): $x^T y \le \|x\| \|y\| \le 1$, also $x \in (B_n)^*$.

d) „$\subseteq$“: klar.

„$\supseteq$“: $y \in \mathbb{R}^n$ erfülle $x_1^T y \le 1, \ldots, x_t^T y \le 1$. Sei $z \in P$. Dann ist

$$z = \sum_{i=1}^t \alpha_i x_i \text{ mit } \alpha_i \ge 0, \; \sum_{i=1}^t \alpha_i = 1.$$

Es ist

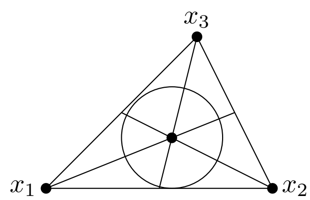
$$z^T y = (\sum_{i=1}^t \alpha_i x_i)^T y = \sum_{i=1}^t \alpha_i x_i^T y \le \sum_{i=1}^t \alpha_i = 1$$

□

Theorem 3.9. *(Weyl)*
Ein Polytop ist ein beschränktes Polyeder.

Beweis. Sei $P = \operatorname{conv}\{x_1, \dots, x_t\}$ ein Polytop im $\mathbb{R}^n$. Es sei
oBdA. $0 \in P$ (durch Translation),
oBdA. $\dim P = n$ (durch Einschränkung auf den Span von P),
oBdA. $x_1, \dots, x_{n+1}$ affin unabhängig (durch Umnumerierung).

Für den Schwerpunkt $x_0 = \frac{1}{n+1}(x_1 + \cdots + x_{n+1})$ gibt es ein $r > 0$, so dass $B(x_0, r) \subseteq P$, wobei $B(x_0, r) = \{y \in \mathbb{R}^n : \|y - x_0\| \leq r\}$.



Sei oBdA. $x_0 = 0$ nach Translation. Nach (a) - (c) von Lemma 3.8 gilt $P^* \subseteq B(x_0, \frac{1}{r})$. Nach (d) ist P^* ein beschränktes Polyeder und nach dem Theorem von Minkowski ist P^* somit ein Polytop. Das heißt

$$P^* = \operatorname{conv}\{y_1, \dots, y_s\} \text{ für } y_1, \dots, y_s \in \mathbb{R}^n.$$

<u>Behauptung</u>: $P = \{x \in \mathbb{R}^n : y_1^T x \leq 1, \dots, y_s^T x \leq 1\}$.

<u>Beweis</u>:

„$\subseteq$“: Sei $x \in P$. Dann gilt $y_i^T x \leq 1$, weil $y_i \in P^*$.

„$\supseteq$“: Sei $x \in \mathbb{R}^n$ mit $y_1^T x \leq 1, \dots, y_s^T x \leq 1$. Angenommen $x \notin P$. Dann gibt es eine Trennhyperebene von $\{x\}$ und P, das heißt

$$\exists c \neq 0, \delta \in \mathbb{R} : \quad c^T x > \delta \text{ und } c^T x' \leq \delta \; \forall x' \in P.$$

Da $0 \in P$, ist $\delta > 0$. Wähle $\delta = 1$ (durch Skalieren von c). Dann ist $c \in P^*$ und somit $c = \sum_{j=1}^s \alpha_j y_j$ mit $\alpha_j \geq 0$ und $\sum_{j=1}^s \alpha_j = 1$. Desweiteren gilt

$$1 < c^T x = (\sum_{j=1}^s \alpha_j y_j)^T x \leq \sum_{j=1}^s \alpha_j = 1.$$

Das ist ein Widerspruch. Somit gilt $x \in P$.

□

Korollar 3.10. *(Theorem von Minkowski-Weyl)*
Sei $P \subseteq \mathbb{R}^n$. Dann gilt: P ist ein beschränktes Polyeder $\Leftrightarrow$ P ist ein Polytop.

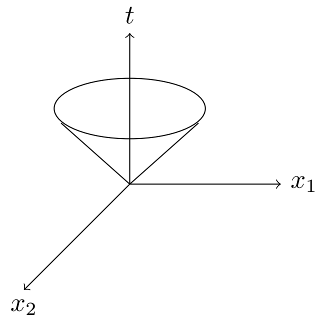
Als nächstes beschäftigen wir uns mit der Frage, wie allgemeine, eventuell unbeschränkte, Polyeder aussehen.

Definition 3.11. *Eine Menge $C \subseteq \mathbb{R}^n$ heißt* konvexer Kegel*, falls*

$$\forall \lambda, \mu \geq 0 \; \forall x, y \in C : \quad \lambda x + \mu y \in C.$$

(Insbesondere ist C tatsächlich konvex.)

Beispiel 3.12. *$\mathcal{L}^{n+1} = \{(x,t) \in \mathbb{R}^{n+1} : \|x\| \leq t\}$ ist ein konvexer Kegel („ice cream cone“).*



Definition 3.13. *(konische Hülle)*
Sei $A \subseteq \mathbb{R}^n$. Die konische Hülle *von A ist*

$$\mathit{cone}\ A = \bigcap_{\substack{B \supseteq A \\ B \text{ konvexer Kegel}}} B$$

Satz 3.14. *Sei $A \subseteq \mathbb{R}^n$, dann gilt*

$$\mathit{cone}\ A = \{y \in \mathbb{R}^n : \exists N \in \mathbb{N}, \lambda_1, \ldots, \lambda_N \geq 0, x_1, \ldots, x_N \in A : y = \sum_{i=1}^{N} \lambda_i x_i\}$$

Definition 3.15. *Ein konvexer Kegel $C \subseteq \mathbb{R}^n$ heißt* endlich erzeugt*, falls es $y_1, \ldots, y_t \in \mathbb{R}^n$ gibt mit*

$$C = \mathit{cone}\{y_1, \ldots, y_t\}.$$

Satz 3.16. *(Minkowski-Weyl für Kegel)*
Ein konvexer Kegel $C \subseteq \mathbb{R}^n$ ist endlich erzeugt $\Leftrightarrow$

$$\exists A \in \mathbb{R}^{m \times n} : C = \{x \in \mathbb{R}^n : Ax \leq 0\}.$$

Beweis. Siehe zum Beispiel Chapter 7 in Schrijver - Theory of linear and integer programming. □

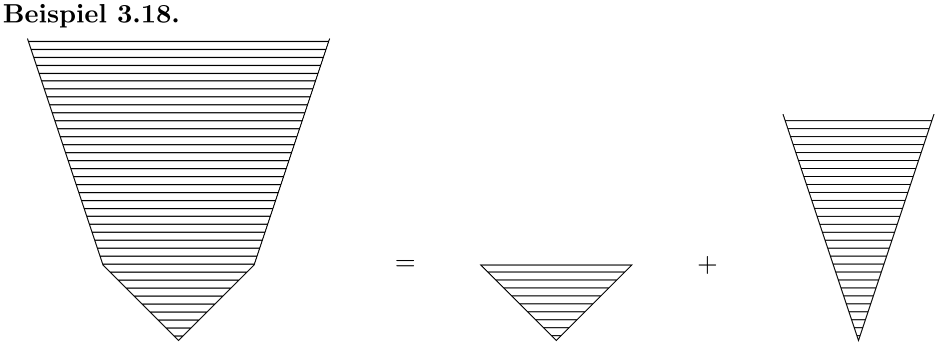
Theorem 3.17. *Sei $P \subseteq \mathbb{R}^n$. P ist ein Polyeder $\Leftrightarrow$*

$$\exists x_1, \ldots, x_s, y_1, \ldots, y_t \in \mathbb{R}^n : P = \mathit{conv}\{x_1, \ldots, x_s\} + \mathit{cone}\{y_1, \ldots, y_t\},$$

wobei $A + B = \{a + b : a \in A, b \in B\}$.

Beweis. Siehe zum Beispiel Kapitel 7 in Schrijver. □

Beispiel 3.18.



4. Lemma von Farkas

Ein Lösbarkeitskriterium für lineare Gleichungssysteme

$$Ax = b \text{ mit } A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^m$$

ist durch den folgenden Satz gegeben.

Satz 4.1. $\exists x \in \mathbb{R}^n : Ax = b \Leftrightarrow \nexists y \in \mathbb{R}^m : y^T A = 0$ *und* $y^T b \neq 0$.

Beweis.

„$\Rightarrow$“: Angenommen $\exists x \in \mathbb{R}^n : Ax = b$ und $\exists y \in \mathbb{R}^m : y^T A = 0, y^T b \neq 0$. Dann

$$\underbrace{y^T A}_{=0} x = \underbrace{y^T b}_{\neq 0}.$$

Widerspruch.

„$\Leftarrow$“: Angenommen $Ax = b$ hat keine Lösung. Dann gilt rang $[A|b] >$ rang A. Wähle $y \in \ker A^T$ mit $y^T b \neq 0$.

□

Wir erhalten ein analoges Resultat für Systeme linearer Ungleichungen.

Satz 4.2. *(Farkas Lemma)*
Seien $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. *Dann*
$\exists x \geq 0 : Ax = b \Leftrightarrow \nexists y \in \mathbb{R}^m : y^T A \geq 0$ *und* $y^T b < 0$.

Beweis.

„$\Rightarrow$“: Angenommen $\exists x \geq 0 : Ax = b$ und $\exists y \in \mathbb{R}^m : y^T A \geq 0,\ y^T b < 0$. Dann

$$y^T Ax = y^T b < 0, \text{ aber } \underbrace{(y^T A)}_{\geq 0} \underbrace{x}_{\geq 0} \geq 0.$$

Widerspruch.

„$\Leftarrow$“: Angenommen $\nexists x \geq 0 : Ax = b$. Seien $a_1, \ldots, a_n \in \mathbb{R}^m$ die Spaltenvektoren von A. Nach Voraussetzung ist $b \notin \text{cone}\{a_1, \ldots, a_n\}$. Also gibt es eine Trennhyperebene von $\{b\}$ und cone $\{a_1, \ldots, a_n\}$. Das heißt

$$\exists y \neq 0 : y^T b < 0 \text{ und } y^T x \geq 0\ \forall x \in \text{ cone}\{a_1, \ldots, a_n\}.$$

□

Interpretation: Theorie der Alternativen.

Entweder: $\exists x \geq 0 : Ax = b$

oder: $\exists y : y^T A \geq 0,\ y^T b < 0$.

Korollar 4.3. *(Variante von Farkas Lemma)*
$\exists x \in \mathbb{R}^n : Ax \leq b \Leftrightarrow \nexists y \geq 0 : y^T A = 0,\ y^T b < 0$.

Beweis. Betrachte die Matrix

$$A' = [A| - A|I_m] \in \mathbb{R}^{m \times (n+n+m)},$$

wobei I_m die $m \times m$ Einheitsmatrix ist.

$$\text{Sei } x' = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^{n+n+m} \text{ mit } x_1, x_2 \in \mathbb{R}^n, x_3 \in \mathbb{R}^m.$$

$$\begin{aligned} \exists x' \geq 0 : A'x' = b \Leftrightarrow\ & \exists x_1, x_2, x_3 \geq 0 : Ax_1 - Ax_2 + x_3 = b \\ \Leftrightarrow\ & \exists x_1, x_2, x_3 \geq 0 : A(x_1 - x_2) + x_3 = b \\ \Leftrightarrow\ & \exists x \in \mathbb{R}^n : Ax \leq b. \end{aligned}$$

Wende Farkas Lemma auf $A'x' = b$ an:

$$\exists x' \geq 0 : A'x' = b \Leftrightarrow\ \nexists y \in \mathbb{R}^m : y^T[A| - A|I_m] \geq 0 \text{ und } y^T b < 0.$$

Dabei

$$\begin{aligned} y^T[A| - A|I_m] \geq 0 \in \mathbb{R}^{n+n+m} &\Leftrightarrow y^T A \geq 0, y^T(-A) \geq 0, y^T I_m \geq 0 \\ &\Leftrightarrow y^T A = 0, y \geq 0 \end{aligned}$$

□

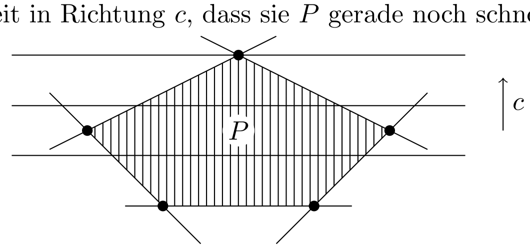
5. Lineare Programmierung

Ein *lineares Programm* (LP) in Standardform ist

$$(\text{LP}) \qquad \begin{aligned} \max\ & c^T x \\ & x \in \mathbb{R}^n \\ & Ax \leq b, \end{aligned}$$

wobei $c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ gegeben sind.

Geometrische Interpretation: Maximiere die lineare Funktion $x \mapsto c^T x$ über dem Polyeder $P = \{x \in \mathbb{R}^n : Ax \leq b\}$. Das heißt man schiebt eine zu c orthogonale Hyperebene so weit in Richtung c, dass sie P gerade noch schneidet.



Sprechweise:
x heißt *gültige/zulässige Lösung* $\Leftrightarrow x \in P$.
x heißt *optimale Lösung* $\Leftrightarrow x \in P$ und $c^T x = \max\{c^T y : y \in P\}$.
LP heißt *unbeschränkt* $\Leftrightarrow \sup\{c^T x : x \in P\} = +\infty$.
LP heißt *ungültig/unzulässig* $\Leftrightarrow \nexists x \in P$.

Satz 5.1. *Falls die Menge der zulässigen Lösungen P ein nicht-leeres, beschränktes Polyeder ist, dann gibt es eine Ecke von P, die eine optimale Lösung von LP ist.*

Beweis. Seien $x_1, \ldots, x_t$ die Ecken von P. Nach dem Theorem von Minkowski ist $P = \operatorname{conv}\{x_1, \ldots, x_t\}$. Angenommen keine Ecke ist eine optimale Lösung. Das heißt (und weil P kompakt ist):

$$\exists y \in P : c^T y > c^T x_i \quad \text{für alle } i = 1, \ldots, t.$$

Schreibe $y = \sum_{i=1}^t \alpha_i x_i$ mit $\alpha_i \geq 0, \sum_{i=1}^t \alpha_i = 1$. Dann ist

$$c^T y = c^T (\sum_{i=1}^t \alpha_i x_i) = \sum_{i=1}^t \alpha_i c^T x_i < \sum_{i=1}^t \alpha_i c^T y = c^T y.$$

Widerspruch. □

Daraus ergibt sich ein offensichtlicher (aber eventuell sehr langsamer) Algorithmus zur Lösung von (LP), falls P beschränkt ist: Bestimme alle Ecken $x_1, \dots, x_t$ von P und finde einen Index i_0 mit $c^T x_{i_0} \geq c^T x_i \; \forall i = 1, \dots, t$.

<u>Problem</u>: t (die Anzahl der Ecken) kann sehr groß werden, zum Beispiel beim Würfel:

$$C_n = \{x \in \mathbb{R}^n : -1 \leq x_i \leq 1, i = 1, \dots, n\}.$$

$2n$ Ungleichungen, aber 2^n Ecken.

Bessere Algorithmen lernen wir in späteren Kapiteln kennen. In diesem Kapitel beschäftigen wir uns zunächst mit Optimalitätsbedingungen und Dualitätstheorie.

Definition 5.2. *Seien $A \in \mathbb{R}^{m \times n}, c \in \mathbb{R}^n, b \in \mathbb{R}^m$ gegeben. Dies definiert das* primale LP

$$\text{(PLP)} \qquad \begin{aligned} p^* = \max\ & c^T x \\ & x \in \mathbb{R}^n \\ & Ax \leq b \end{aligned}$$

und das duale LP

$$\text{(DLP)} \qquad \begin{aligned} d^* = \min\ & y^T b \\ & y \in \mathbb{R}^m \\ & y \geq 0 \\ & y^T A = c^T. \end{aligned}$$

Satz 5.3. *(*schwache Dualität*)*
Sei x zulässig für (PLP) und y zulässig für (DLP), dann ist

$$c^T x \leq y^T b.$$

Insbesondere ist $p^* \leq d^*$.

Beweis. Es gilt

$$y^T b - c^T x \geq y^T (Ax) - (y^T A) x = 0$$

□

Satz 5.4. *(*Optimalitätsbedingung*)*
Seien x, y zulässig. Dann ist x optimal für (PLP) und y optimal für (DLP) $\Leftrightarrow$ $y^T (Ax - b) = 0$.
Insbesondere gelten die Komplementaritätsbedingungen*:*

$$\begin{aligned} y_j \neq 0 &\Rightarrow (Ax - b)_j = 0 \\ (Ax - b)_j \neq 0 &\Rightarrow y_j = 0, \quad \textit{für alle } j = 1, \dots, m. \end{aligned}$$

Beweis. $\rightarrow$ Blatt 9. □

Theorem 5.5. *(*starke Dualität*, von Neumann (1947))*
Falls (PLP) und (DLP) beide zulässige Lösungen besitzen, dann gilt $p^* = d^*$.

Beweis.

$p^* \leq d^*$: ✓schwache Dualität.

$p^* \geq d^*$: 1.Behauptung: $\exists x_0 : Ax_0 \leq b$ und $c^T x_0 = p^*$.

Beweis: Da (DLP) zulässig ist, folgt aus der schwachen Dualität:

$$p^* = \sup\{c^T x : Ax \leq b\} < +\infty.$$

Zu zeigen: $\exists x_0 : Ax_0 \leq b$ und $c^T x_0 \geq p^*$. Angenommen ein solches x_0 gibt es nicht. Das heißt $\nexists x_0 : \begin{bmatrix} A \\ -c^T x \end{bmatrix} x_0 \leq \begin{pmatrix} b \\ -p^* \end{pmatrix}$. Nach Farkas Lemma

$$\exists \begin{pmatrix} y \\ \lambda \end{pmatrix} \geq 0 : (y^T \ \lambda) \begin{bmatrix} A \\ -c^T x \end{bmatrix} = 0 \text{ und } (y^T \ \lambda) \begin{pmatrix} b \\ -p^* \end{pmatrix} < 0.$$

Dann

$$\begin{aligned} \lambda p^* &= \sup\{\lambda c^T x : Ax \leq b\} \\ &= \sup\{y^T Ax : Ax \leq b\} \\ &\leq y^T b \\ &< \lambda p^*. \end{aligned}$$

Widerspruch.

2. Behauptung: $\exists y_0 \geq 0 : y_0^T A = c^T$ und $y_0^T b \leq p^*$.

Beweis: Angenommen ein solches y_0 gibt es nicht. Dann hat

$$(y_0^T \ \lambda) \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} = (c^T \ p^*)$$

keine Lösung mit $(y_0^T \ \lambda) \geq (0\ 0)$. Nach Farkas Lemma

$$\exists \begin{pmatrix} z \\ \mu \end{pmatrix} \in \mathbb{R}^{n+1} \text{ mit } \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \begin{pmatrix} z \\ \mu \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ und } (c^T \ p^*) \begin{pmatrix} z \\ \mu \end{pmatrix} < 0.$$

1. Fall: $\mu = 0$

Dann ist $Az \geq 0$ und $c^T z < 0$. Nach Voraussetzung hat $Ax \leq b$ eine Lösung x_0. Dann ist für großes τ:

$$A(x_0 - \tau z) \leq b \text{ und } c^T (x_0 - \tau z) > p^*.$$

Widerspruch.

2. Fall: $\mu > 0$.

OBdA. sei $\mu = 1$ (durch Skalieren). Dann

$$Az + b \geq 0 \text{ und } c^T z + p^* < 0.$$

Also

$$A(-z) \leq b \text{ und } c^T (-z) > p^*.$$

Widerspruch.

□

Korollar 5.6. *Es gilt*

$$\max\{c^T x : x \geq 0,\ Ax \leq b\} = \min\{y^T b : y^T A \geq c^T\},$$

falls beide Mengen gültiger Lösungen nicht leer sind.

Beweis. Setze $\tilde{A} = \begin{bmatrix} A \\ -A \\ -I \end{bmatrix}$, $\tilde{b} = \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}$. Dann

$$\begin{aligned}
&\max\{c^T x : x \geq 0,\ Ax = b\} \\
&= \max\{c^T x : \tilde{A}x \leq \tilde{b}\} \\
&= \min\{z^T \tilde{b} : z \geq 0,\ z^T \tilde{A} = c^T\} \\
&= \min\{u^T b - v^T b : u, v, w \geq 0,\ u^T A - v^T A - w^T = c^T\} \quad (\text{mit } z = \begin{pmatrix} u \\ v \\ w \end{pmatrix}) \\
&= \min\{y^T b : y^T A \geq c^T\} \quad (\text{mit } y = u - v).
\end{aligned}$$

$\square$

Prof. Dr. F. Vallentin, Dr. A. Gundert, Mathematisches Institut, Universität zu Köln, Weyertal 86–90, 50931 Köln, Deutschland
E-mail address: frank.vallentin@uni-koeln.de, anna.gundert@uni-koeln.de