Lineare und ganzzahlige Optimierung

Dozent
Dr. Ulrich Brenner

Mitschrift
Josia Pietsch

Version
git: c4b738f
kompiliert: 22. Februar 2024 18:54

Inhaltsverzeichnis

Dies sind meine Notizen zur Vorlesung "Lineare und ganzzahlige Optimierung", die im Sommersemester 2022 von Dr. Ulrich Brenner an der Universität Bonn gehalten wurde.

Warning. *Dies ist kein offizielles Skript.*

Diese Mitschrift enthält sicherlich einige Fehler. Ich freue mich bei gefundenen Fehlern[1] oder anderweitigen Verbesserungsvorschlägen immer über eine Mail an *lecturenotes@jrpie.de*.

0.1 Voraussetzungen

Es wäre gut, EDM gehört zu haben.

1 Übersicht

Example 1 (Wichtige Anwedung)**.** Ein Bauer hat $10ha$ Land, auf dem er Mais oder Weizen anbauen kann. $1ha$ Mais bringt 2 Einheiten Geld ein, $1ha$ Weizen hingegen 3 Einheiten Geld.

Tabelle 1: Benötigte Ressourcen

	Mais	Weizen
Arbeit	1 Tag	3 Tage
Wasser	5	2

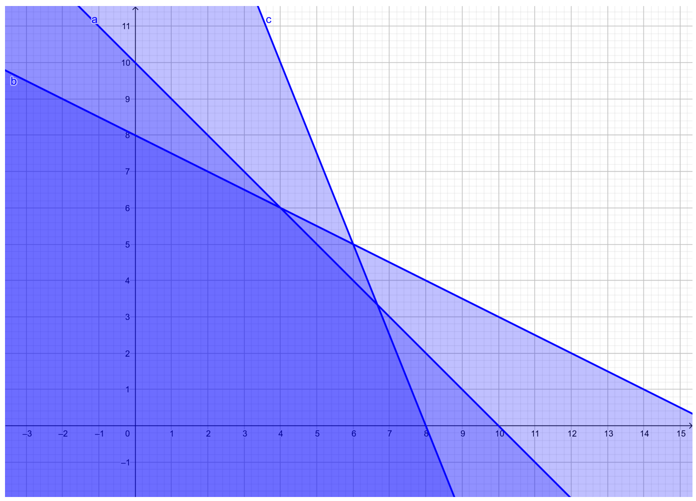
Es stehen 16 Arbeitstage und 40 Einheiten Wasser zur Verfügung. Ziel ist es, die Einnahmen zu maximieren.

Modellierung: $x :=$ Fläche mit Mais, $y :=$ Fläche mit Weizen.

$$\begin{aligned} \text{Maximiere} \quad & 2x + 3y \\ \text{so dass} \quad & \\ x + y &\leqslant 10 \\ x + 2y &\leqslant 16 \\ 5x + 2y &\leqslant 40 \\ x, y &\geqslant 0 \end{aligned}$$

Graphische Lösung:

[1] auch Rechtschreibfehlern und ähnlichen Kleinigkeiten



Es ergibt sich eine optimale Lösung von $(x, y) = (4, 6)$.

In der Vorlesung werden wir

- Probleme als Lineare Programme modellieren
- Die Struktur der Lösungen untersuchen

2 Lineare Programme

Definition 2 (Optimierungsproblem)**.** Ein **Optimierungsproblem** ist ein Paar (I, f) aus einer Menge I und einer **Zielfunktion** $f : I \to \mathbb{R}$.

Elemente von I heißen **zulässige Lösungen** von (I, f).

Falls $I = \varnothing$, so heißt (I, f) **unzulässig**, anderenfalls heißt (I, f) **zulässig**.

Wenn (I, f) ein **Maximierungsproblem** ist, dann suchen wir ein $x^* \in \operatorname{argmax}_{x \in I} f(x)$. Das Problem heißt **beschränkt**, wenn es ein $k \in \mathbb{R}$ gibt, sodass $f(x) \leqslant k$ für alle $x \in I$ gilt. Anderenfalls heißt das Problem **unbeschränkt**.

Analoges definieren wir für **Minimierungsprobleme**.

Observe. *Unzulässige Optimierungsprobleme sind beschränkt.*

Definition 3 (Lineare Optimierungsprobleme)**.** Ein **lineares Optimierungsproblem**[a] ist ein Optimierungsproblem der folgenden Form:

Gegeben eine Matrix $A \in \mathbb{R}^{m \times n}$, Vektoren $c \in \mathbb{R}^n, b \in \mathbb{R}^m$.

Gesucht ist ein Vektor $x \in \mathbb{R}^n$ mit $Ax \leqslant b$, der $c^t x$ maximiert.

[a] Aus historischen Gründen auch **lineares Programm** (LP)

Observe. *Es handelt sich um ein Optimierungsproblem mit* $I := \{x \in \mathbb{R}^n | Ax \leqslant b\}, f(x) := c^t \cdot x$.

Notation 3.1. Schreibweise für lineares Programme:

$$\begin{aligned} &\max\ c^t x \\ &\text{s.d. } Ax \leqslant b \end{aligned}$$

Diese Form heißt **Standard-Ungleichungsform**.

Kürzer, aber missverständlich: $\max\{c^t x | Ax \leqslant b\}$ bezeichnet manchmal auch das Maximierungsproblem, unter Umständen aber auch einfach nur das Maximum.

Notation 3.2. Wir schreiben immer $A = (a_{i,j})_{\substack{i \in \{1,\dots,m\} \\ j \in \{1,\dots,n\}}}$ und c als $c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = (c_1, \dots, c_n)$. Alle Vektoren sind Spaltenvektoren (außer er wird explizit transponiert).

Es bezeichne I_n die Einheitsmatrix. Für Matrizen A, B bezeichne ferner $[A|B]$ die Konkatenation der Matrizen.

Example 4 (Beispiel aus der Einleitung)**.**

$$\begin{aligned} \max \quad & (2,3) \begin{pmatrix} x \\ y \end{pmatrix} \\ \text{s.d.} \quad & \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 5 & 2 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leqslant \begin{pmatrix} 10 \\ 10 \\ 40 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

Observe. *Minimierungsprobleme können auch modelliert werden:*

$$\begin{aligned} &\min\ c^t x \\ &s.d.\ Ax \leqslant b \end{aligned}$$

ist offenbar äquivalent zu

$$\begin{aligned} \max\ & -c^t x \\ s.d.\ & Ax \leqslant b \end{aligned}$$

Außerdem sind auch Nebenbedingungen der Form

$$\sum_{j=1}^{n} a_{ij} x_j \geqslant b_i$$

möglich (da äquivalent zu $-\sum_{j=1}^{n} a_{ij} x_j \leqslant -b_i$ *). Insbesondere kann auch Gleichheit gefordert werden.*

Warning. *Strikte Ungleichungen dürfen nicht gefordert werden!*

Notation 4.1 (Standard-Gleichungsform)**.**

$$\begin{aligned} \max\ & c^t x \\ \text{s.d. } & Ax = b \\ & x \geqslant 0 \end{aligned}$$

Beide Standartformen können ineinander überführt werden:

- Gleichungsform $\rightsquigarrow$ Ungleichungsform:

  Ersetze $Ax = b$ durch $Ax \leqslant b$ und $-Ax \leqslant -b$ und $x \geqslant 0$ durch $-I_n x \leqslant 0$

- Ungleichungsform $\rightsquigarrow$ Gleichungsform:

  Sei (P):

  $$\begin{aligned} \max\ & c^t x \\ \text{s.d. } & Ax \leqslant b \end{aligned}$$

  ein LP in Standard-Ungleichungsform.

  Ersetze jede Variable x_i durch zwei Variablen $z_i, \overline{z_i}$. Für jede der Ungleichungen in $Ax \leqslant b$ führen wir eine **Slackvariable** (deutsch auch **Schlupfvariable**) $\tilde{x_i}$ ein.

  Sei $z = (z_1 \dots, z_n)$, $\overline{z} = (\overline{z_1}, \dots, \overline{z_n})$, $\tilde{x} = (\tilde{x_1}, \dots, \tilde{x_m})$. Betrachte (P'):

  $$\begin{aligned} \max\ & c^t (z - \overline{z}) \\ \text{s.d. } & [A| - A|I_m] \begin{pmatrix} z \\ \overline{z} \\ \tilde{x} \end{pmatrix} = b \\ & \begin{pmatrix} z \\ \overline{z} \\ \tilde{x} \end{pmatrix} \geqslant 0 \end{aligned}$$

Die Nebenbedingung ist offenbar äquivalent zu $Az - A\overline{z} + \tilde{x} = b$. Jede Optimallösung x von (P) lässt sich in eine Optimallösung von (P') umwandeln:

Setze $z_j := \max\{x_j, 0\}, \overline{z_j} := -\min\{x_j, 0\}$, $\tilde{x_i} := b_i - \sum_{j=1}^{n} a_{ij}x_i$.

Jede Optimallösung $z, \overline{z}, \tilde{x}$ von (P') lässt sich in eine Optimallösung von (P) umwandeln: Setze $x = z - \overline{z}$.

Observe. *Bei der Umwandlung der Standard-Ungleichungsform in die Standard-Gleichungsform erhöht sich die Zahl der Dimensionen, aus einer $m \times n$-Matrix wird eine $m \times (2n + m)$-Matrix. Allerdings hat der Lösungsraum eine bessere Struktur: Sie sind immer spitz, siehe Example 48.*

Mögliche Ergebnisse für ein LP

- **unzulässig** ($I = \varnothing$) (infeasible)
- zulässig und **unbeschränkt** (feasible but unbounded)
- zulässig und **beschränkt** (feasible and bounded)

  In diesem Fall exsitiert eine Optimallösung. Diese kann in polynomieller Zeit berechnet werden, wie wir im Verlauf der Vorlesung sehen werden.

Example 5.

$$\begin{aligned} &\max \min\{c^t x + d, e^t x + f\} \\ &\text{s.d. } Ax \leqslant b \end{aligned}$$

$A \in \mathbb{R}^{m \times n}, c, e \in \mathbb{R}^n, b \in \mathbb{R}^m, d, f \in \mathbb{R}$

kann wie folgt als LP modelliert werden

$$\begin{aligned} \max\ & \sigma \\ \text{s.d. } & \sigma - c^t x \leqslant d \\ & \sigma - e^t x \leqslant f \\ & Ax \leqslant b \end{aligned}$$

Example 6.

$$\begin{aligned} &\min |c^t x + d| \\ &\text{s.d. } Ax \leqslant b \end{aligned}$$

lässt sich als LP formulieren als

$$\begin{aligned} \max\ & -\sigma \\ \text{s.d. } & -\sigma - c^t x \leqslant d \\ & -\sigma + c^t x \leqslant -d \\ & Ax \leqslant b \end{aligned}$$

Die Nebenbedingungen sind äquivalent zu $\sigma \geqslant \max\{c^t x + d, -c^t x - d\} = |c^t x + d|$.

Definition 7 (Ganzzahliges lineares Programm)**.** Ein Problem der Form

$$\begin{aligned} \max\ & c^t x \\ & Ax \leqslant b \\ & x \in \mathbb{Z}^n \end{aligned}$$

heißt **ganzzahliges lineares Programm** (GLP).

Warning. *Ein ganzzahliges lineares Programm ist kein lineares Programm. Das Lösen ganzzahliger linearer Programme ist* NP*-schwer, lineare Programme können hingegen in polynomieller Zeit gelöst werden.*

Oft fordert man nur für einen Teil der Einträge von x Ganzzahligkeit. Dies bezeichnet man dann als **gemischt-ganzzahliges lineares Programm** (MILP).

2.1 Modellierung von Optimierungsproblemen als (MI)LPs

Definition 8. Sei G ein gerichteter Graph mit Kapazitäten $u : E(G) \to \mathbb{R}_{>0}$ und s, t Knoten von G. Ein zulässiger $s-t$-Fluss in (G, u) ist eine Abbildung $f : E(G) \to \mathbb{R}_{>0}$ mit

- $f \leqslant u$
- $\Delta_f(v) := \sum_{e \in \delta^+_G(v)} f(e) - \sum_{e \in \delta^-_G(v)} f(e) = 0$ für alle $v \in V(G) \setminus \{s, t\}$

Der Wert eines $s-t$-Fluses ist $\mathrm{val}(f) := \Delta_f(s)$.

Problem (Maximum-Flow-Problem)**.** *Gegeben sei ein gerichteter Graph G mit Kapazitäten u und Knoten $s \neq t \in V(G)$.*

Gesucht ist ein $s-t$-Fluss mit maximalem Wert.

LP-Formulierung:

$$
\begin{array}{lll}
\max & \sum_{e\in\delta^+_G(s)} x_e - \sum_{e\in\delta^-_G} x_e & \\
\text{s.d.} & x_e \geqslant 0 & \forall e \in E(G) \\
& x_e \leqslant u_e & \forall e \in E(G) \\
& \Delta_f(v) = 0 \forall v \in V(G)\backslash\{s,t\} &
\end{array}
$$

Problem (Bottleneck-Maximum-Flow-Problem mit 2 Quellen)**.** *Gegeben sei ein gerichteter Graph G, Kapazitäten u und drei Knoten $s_1, s_2, t \in V(G)$.*

Finde eine Abbildung $f : E(G) \to \mathbb{R}_{\geqslant 0}$ mit

- $f(e) \leqslant u(e)\ \forall e \in E(G)$
- $\Delta_f(v) = 0\ \forall v \in V(G)\backslash\{s_1, s_2, t\}$

sodass $\min\{\Delta_f(s_1), \Delta_f(s_2)\}$ *maximiert wird.*

Idea. *Verwende die LP-Formulierung von Maximum-Flow und Example 5.*

Andere Probleme lassen sich hingegen als ILP, aber (vermutlich) nicht als LP formulieren:

Problem (Vertex Cover)**.** *Gegeben sei ein ungerichteter Graph G mit Gewichten $c : V(G) \to \mathbb{R}_{\geqslant 0}$.*

Gesucht ist eine Menge $X \subseteq V(G)$ mit $\{v, w\} \cap X \neq \emptyset$ für alle $e = \{v, w\} \in E(G)$, so dass $c(X)$ minimal ist.

Dieses Problem ist NP-schwer und kann daher (vermutlich) nicht als LP formuliert werden. Als ILP ist dies hingegen möglich:

$$
\begin{array}{lll}
\min & \sum_{v\in V(G)} x_v c(v) & \\
\text{s.d.} & x_v + x_w \geqslant 1 & \forall \{v,w\} \in E(G) \\
& x_v \in \{0,1\} & \forall v \in V(G)
\end{array}
$$

Dabei steht $x_v = 1$ dafür, dass $v \in X$. Insbesondere folgt, dass ganzzahlige lineare Programmierung NP-schwer ist.

Durch Weglassen der Ganzzahligkeitsbedingungne erhalten wir das folgende LP:

$$
\begin{array}{lll}
\min & \sum_{v\in V(G)} x_v c(v) & \\
\text{s.d.} & x_v + x_w \geqslant 1 \forall \{v,w\} \in E(G) & \\
& x_v \geqslant 0 & \forall v \in V(G) \\
& x_v \leqslant 1 & \forall v \in V(G)
\end{array}
$$

Dieses LP wird als **LP-Relaxierung** bezeichnet. In diesem Fall liefert uns die Relaxierung eine 2-Approximation des Vertex Cover Problemes: Für jede Lösung x des relaxierten Problemes erhalten wir eine ganzzahlige Lösung $\tilde{x}$ durch

$$\tilde{x}_v := \begin{cases} x & : x_v \geqslant \frac{1}{2} \\ 0 & : x_v < \frac{1}{2} \end{cases}$$

Bei einem Minimierungsproblem kann das Relaxieren den Wert einer optimalen Lösung offenbar nur verkleinern. Das Supremum des Verhältnisses des Wertes einer optimalen Lösung eines ILP und seiner Relaxierung bezeichnen wir als **integrality gap** der Relaxierung.

Für die vorige Formulierung von Vertex Cover haben wir bereits gezeigt, dass die integrality gap höchstens 2 ist. Für die andere Abschätzung betrachte den vollständigen Graphen mit $c \equiv 1$.

Problem (Stable Set Problem)**.** *Gegeben sei ein ungerichteter Graph G mit Gewichten $c : V(G) \to \mathbb{R}_{\geqslant 0}$.*

Gesucht wird eine Menge $X \subseteq V(G)$ mit $|\{v,w\} \cap X| \leqslant 1$ für alle $e = \{v,w\} \in E(G)$, welche $c(X)$ maximiert.

Auch dieses Problem kann als ILP formuliert werden:

$$\begin{aligned} \max \quad & \sum_{v \in V(G)} x_v c(v) & \\ \text{s.d.} \quad & x_v + x_w \leqslant 1 & \forall \{v,w\} \in E(G) \\ & x_v \in \{0,1\} & \forall v \in V(G) \end{aligned}$$

Die LP-Relaxierung liefert hier jedoch keine sinnvollen Ergebnisse. Selbst wenn $G = K_n$ ist, kann $x_v = \frac{1}{2}$ für alle v gesetzt werden, folglich ist die integrality gap mindestens $\frac{n}{2} \xrightarrow{n \to \infty} \infty$.

2.2 Polyeder

Definition 9. Sei $X \subseteq \mathbb{R}^n$. X heißt **konvex**, wenn für alle $x, y \in X, t \in [0,1]$ auch $tx + (1-t)y \in X$.

Definition 10. Seien $x_1, \ldots, x_k \in \mathbb{R}^n$, sowie $\lambda_1, \ldots, \lambda_k \geqslant 0$ mit $\sum_i \lambda_i = 1$. Dann heißt $x = \sum_i \lambda_i x_i$ **Konvexkombination** von $x_1, \ldots, x_k$.

Die **konvexe Hülle** $\operatorname{conv}(X)$ ist die Menge $X \subseteq \mathbb{R}^n$ aller Konvexkombinationen von Mengen von Vektoren aus X.

Remark 10.1. Die konvexe Hülle ist die inklusionsminimale konvexe Menge, welche X enthält.

Definition 11. Sei $X \subseteq \mathbb{R}^n$.

(a) X heißt **Halbraum** (**half-space**), falls ein $a \in \mathbb{R}^n \backslash \{0\}$ und ein $b \in \mathbb{R}$ existiert, so dass $X = \{x \in \mathbb{R}^n | a^t x \leqslant b\}$. Der Vektor a heißt **Normalenvektor** von X.

(b) X heißt **Hyperebene** (**hyperplane**), wenn es einen Vektor $a \in \mathbb{R}^n \backslash \{0\}$ und ein $b \in \mathbb{R}$ gibt, sodass $X = \{x \in \mathbb{R}^n | a^t x = b\}$. a heißt dabei **Normalenvektor** von X.

(c) X heißt **Polyeder** (**polyhedron**), wenn es eine Matrix $A \in \mathbb{R}^{m \times n}$ und ein $b \in \mathbb{R}^m$ gibt, sodass $X = \{x \in \mathbb{R}^n | Ax \leqslant b\}$.

(d) X heißt **Polytop** (**polytope**), wenn es ein Polyeder ist und es eine Zahl $K \in \mathbb{R}$ gibt, sodass $\|x\| \leqslant K$ für alle $x \in X$ gilt.

Example 12. $\varnothing = \{x \in \mathbb{R}^n | 0x \leqslant -1\}$ ist ein Polytop. $\mathbb{R}^n = \{x \in \mathbb{R}^n | 0x \leqslant 0\}$ ist ein Polyeder, aber kein Polytop.

Observe. *Polyeder sind konvex und abgeschlossen.*

Lemma 13. Eine Menge $X \subseteq \mathbb{R}^n$ ist genau dann ein Polyeder, wenn eine der beiden folgenden Bedingungen gilt:

- $X = \mathbb{R}^n$
- X ist der Schnitt von endlich vielen Halbräumen.

Beweis. " $\Longleftarrow$ " trivial.

" $\Longrightarrow$ "

Sei $X = \{x \in \mathbb{R}^n | Ax \leqslant b\}$. Falls $X = \varnothing$, so ist $X = \{x \in \mathbb{R}^n | \mathbb{1}_n^t x \leqslant -1\} \cap \{x \in \mathbb{R}^n | -\mathbb{1}_n^t x \leqslant -1\}$ Anderenfalls können wir annehmen, dass A keine Zeile enthält, die der Nullvektor ist.

Dann ist $X = \bigcap_i \{x \in \mathbb{R}^n | a_i x \leqslant b_i\}$, wobei a_i die Zeilen von A seien. □

Definition 14. Die **Dimension** einer Menge $X \subseteq \mathbb{R}^n$ ist

$$\dim(X) := n - \max\{\operatorname{rank}(A) | A \in \mathbb{R}^{n \times n}, Ax = Ay \forall x, y \in X\}$$

Observe. *Eine äquivalente Definition ist, dass* $\dim(X) = d$, *wobei* d *maximal ist, so dass es* $v_0, v_1, \ldots, v_d \in X$ *gibt, für die* $v_1 - v_0, \ldots, v_d - v_0$ *linear unabhängig sind.*

Definition 15. Eine Menge $X \subseteq \mathbb{R}^n$ heißt **konvexer Kegel** (**convex cone**), wenn $X \neq \emptyset$ und für alle $x, y \in X$ und $\lambda, \mu \in \mathbb{R}_{\geqslant 0}$ gilt: $\lambda x + \mu y \in X$.

Observe. *$\emptyset \neq X \subseteq \mathbb{R}^n$ ist genau dann ein konvexer Kegel, wenn X konvex ist und $x \in X, \lambda \geqslant 0 \implies \lambda x \in X$.*

Definition 16. $X \subseteq \mathbb{R}^n$ heißt **polyedrischer Kegel** (**polyhedral cone**) wenn X ein Polyeder und ein konvexer Kegel ist.

Lemma 17. Eine Menge $X \subseteq \mathbb{R}^n$ ist genau dann ein polyedrischer Kegel, wenn es eine Matrix $A \in \mathbb{R}^{m \times n}$ gibt, sodass $X = \{x \in \mathbb{R}^n | Ax \leqslant 0\}$.[a]

[a] Wir fordern, dass $\emptyset$ kein konvexer Kegel ist, um dieses Lemma zu ermöglichen.

Beweis. " $\impliedby$ " klar.

" $\implies$ " Sei $X \subseteq \mathbb{R}^n$ ein polyedrischer Kegel. Dann ist $X = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ für eine Matrix A und einen Vektor b. Wegen $0 \in X$ gilt $b_j \geqslant 0$ für alle Einträge. Folglich ist $X \supseteq \{x \in \mathbb{R}^n | Ax \leqslant 0\}$.

Angenommen es existiert $x \in X$, sodass $(Ax)_i > 0$. Dann ist $\lambda(Ax)_i > b_i$ für hinreichend groß gewähltes λ. $\lightning$.

Daher ist schon $X = \{x \in \mathbb{R}^n | Ax \leqslant 0\}$. ☐

Notation 17.1. Es seien $x_1, \ldots, x_m \in \mathbb{R}^n$ Vektoren.

Der von $x_1, \ldots, x_m$ **erzeugte Kegel** ist

$$\operatorname{cone}(\{x_1, \ldots, x_m\}) := \left\{ \sum_{i=1}^{m} \lambda_i x_i | \lambda_i \geqslant 0 \right\}$$

Ein konvexer Kegel C heißt **endlich erzeugt**, wenn es endlich viele Vektoren $x_1, \ldots, x_m \in \mathbb{R}^n$ gibt, sodass $C = \operatorname{cone}(\{x_1, \ldots, x_m\})$.

Observe. *$\operatorname{cone}(\{x_1, \ldots, x_m\})$ ist tatsächlich ein konvexer Kegel.*

Wir werden sehen, dass ein konvexer Kegel genau dann polyedrisch ist, wenn dieser endlich erzeugt ist (siehe Theorem 52).

3 Dualität

Example 18 (Dualität)**.**

$$\begin{aligned} (P) \max\ & 12x_1 + 10x_2 \\ \text{s.d.}\ & 4x_1 + 2x_2 \leqslant 5 \\ & 8x_1 + 12x_2 \leqslant 7 \\ & 2x_1 - 3x_2 \leqslant 1 \end{aligned}$$

Es ist

$$12x_1 + 10x_2 = 2(4x_1 + 2x_2) + \frac{1}{2}(8x_1 + 12x^2) \leqslant 13.5$$

Somit haben wir eine obere Schranke gefunden.

Allgemeiner Ansatz Finde $u_1, u_2, u_3 \in \mathbb{R}_{\geqslant 0}$, sodass

$$12x_1 + 10x_2 = u_1 \cdot (4x_1 + 2x_2) + u_2 \cdot (8x_1 + 12x_2) + u_3 \cdot (2x_1 - 3x_2)$$

Dann ist $5u_1 + 7u_2 + u_3$ eine obere Schranke für den Wert jeder Lösung von (P). Wir wollen u_1, u_2, u_3 also so wählen, dass $5u_1 + 7u_2 + u_3$ minimiert wird.

Es ergibt sich das zugehörige **duale LP**:

$$\begin{aligned} \min\ & 5u_1 + 7u_2 + u_3 \\ \text{s.d.}\ & 4u_1 + 8u_2 + 2u_3 = 12 \\ & 2u_1 + 12u_2 - 3u_3 = 10 \end{aligned}$$

Definition 19 (Duales LP)**.** Für das lineare Programm (P)

$$\begin{aligned} \max\ & c^t x \\ \text{s.t.}\ & Ax \leqslant b \end{aligned}$$

in Standard-Ungleichungsform ist das **duale lineare Programm (D)** definiert als

$$\begin{aligned} \min\ & b^t z \\ \text{s.t.}\ & A^t y = c \\ & c \geqslant 0 \end{aligned}$$

In diesem Zusammenhang heißt (P) dann auch **primales LP**.

Observe. *Alle Arten von LPs lassen sich dualisieren, wenn man sie erst in*

Standard-Ungleichungsform schreibt.

Das duale LP hängt von der Zielfunktion und von der Beschreibung des linearen Programmes ab. Hinzufügen redundanter Ungleichungen ändert das duale LP.

Remark† 19.1. Wird das duale LP wieder dualisiert, so ergibt sich ein LP, das äquivalent zum primalen LP ist:

Sei $\max\{c^t x | Ax \leqslant b\}$ das primale LP. Dann ist das duale LP

$$\min\{b^t y | A^t y = c, y \geqslant 0\} = \max\left\{ b^t z | \begin{pmatrix} A^t \\ -A^t \\ I_m \end{pmatrix} z \leqslant \begin{pmatrix} c \\ -c \\ 0 \end{pmatrix} \right\}$$

Das hierzu duale LP ist

$$\max\left\{ (c, -c, 0) \begin{pmatrix} x_+ \\ x_- \\ \tilde{x} \end{pmatrix} | \begin{pmatrix} A & -A & I_m \end{pmatrix} \begin{pmatrix} x_+ \\ x_- \\ \tilde{x} \end{pmatrix} = 0, x_+, x_-, \tilde{x} \geqslant 0 \right\}$$

Eine Lösung dieses Systems entspricht einer Lösung des ursprünglichen primalen LPs, indem $x := x_+ - x_-$ definiert wird, die $\tilde{x}$ sind Schlupfvariablen.

Theorem 20 (Schwache Dualität)**.** Wenn die Ungleichungssysteme $Ax \leqslant b$ und $y^t A = c, y \geqslant 0$ beide eine Lösung haben, dann gilt:

$$\max\{c^t x | Ax \leqslant b\} \leqslant \min\{b^t y | A^t y = c, c \geqslant 0\}$$

Beweis. Für x mit $Ax \leqslant b$ und y mit $A^t y = c, y \geqslant 0$ gilt $c^t x = y^t Ax \leqslant y^t b$. □

3.1 Fourier-Motzkin-Elimination

Goal. *Finde ein Zertifikat dafür, dass ein Ungleichungssystem $Ax \leqslant b$ keine Lösung hat.*

Die **Fourier-Motzkin-Elimination** ist ein extrem ineffizientes Verfahren um zu entscheiden, ob ein LP eine Lösung hat.

Example 21.

$$\begin{array}{rrrl} 3x & +2y & +4z & \leqslant 10 \\ 3x & & +2z & \leqslant 9 \\ 2x & -y & & \leqslant 5 \\ -2x & & & \leqslant 4 \\ & 2y & +2z & \leqslant 7 \end{array}$$

Wir wollen die Variable x eliminieren. Das System ist äquivalent zu

$$\begin{array}{llll} x \leqslant \frac{10}{3} & -\frac{2}{3}y & -\frac{4}{3}z \\ x \leqslant 3 & & -\frac{2}{3}z \\ x \leqslant \frac{5}{2} & +\frac{1}{2}y & \\ x \geqslant -3 & +2y & -z \\ x \geqslant -2 & & \\ & 2y & +2z & \leqslant 7 \end{array}$$

Die ist offenbar genau dann zulässig, wenn das folgende System eine Lösung hat:

$$\min\left\{\frac{10}{3}-\frac{2}{3}y, 3-\frac{2}{3}z, \frac{5}{2}+\frac{1}{2}y\right\} \geqslant \max\{-3+2y-z, -2\}$$
$$2y+2z \leqslant 7$$

$$\begin{array}{rrrl} \frac{10}{3} & -\frac{2}{3}y & -\frac{4}{3}z & \geqslant -3 +2y -z \\ \frac{10}{3} & -\frac{2}{3}y & -\frac{4}{3}z & \geqslant -2 \\ 3 & & -\frac{2}{3}z & \geqslant -3 +2y -z \\ 3 & & -\frac{2}{3}z & \geqslant -2 \\ \frac{5}{2} & +\frac{1}{2}y & & \geqslant -3 +2y -z \\ \frac{5}{2} & +\frac{1}{2}y & & \geqslant -2 \\ & 2y & +2z & \leqslant 7 \end{array}$$

Dieses kann wieder in Standardform umgewandelt werden.

Eigenschaften des neuen Systems:

- Es gibt eine Variable weniger
- Das neue System ist genau dann zulässig, wenn das alte es war
- Jede neue Ungleichung ist nicht-negative Linearkombination von Ungleichungen des alten Systems.

Iteriere diese Schritte und entferne alle Variablen.

Remark 21.1. Im Allgemeinen ist dieses Verfahren sehr ineffizient, da die Anzahl der Ungleichungen in jedem Schritt quadratisch steigen kann.

Notation† **21.2.** Für $A \in \mathbb{R}^{m \times n}$ sei $N(A)$ die Zahl der Spalten, die keine Nullvektoren sind.

Theorem 22 (Fourier-Motzkin-Elimination)**.** Sei $A \in \mathbb{R}^{m\times n}$ mit $N(A) > 0$, $b \in \mathbb{R}^m$. Dann gibt es $\tilde{A} \in \mathbb{R}^{\tilde{m}\times n}$ und $\tilde{b} \in \mathbb{R}^{\tilde{m}}$ mit $N(\tilde{A}) < N(A)$ und $\tilde{m} \in \mathcal{O}(m^2)$, sodass

(a) Jede Ungleichung in $\tilde{A}x \leqslant \tilde{b}$ ist eine positive Linearkombination von Ungleichungen aus $Ax \leqslant b$.

(b) Das System $Ax \leqslant b$ ist genau dann lösbar wenn $\tilde{A}x \leqslant \tilde{b}$ lösbar ist.

Beweis. Sei $A = (a_{i,i})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$. O.B.d.A. sei die erste Spalte von A nicht der Nullvektor. Zerlege die Menge $\{1,\dots,m\}$ wie folgt in disjunkte Mengen U, L, N :

$$\begin{aligned} U &:= \{i \in \{1,\dots,m\} | a_{i,1} > 0\} \\ L &:= \{i \in \{1,\dots,m\} | a_{i,1} < 0\} \\ N &:= \{i \in \{1,\dots,m\} | a_{i,1} = 0\} \end{aligned}$$

O.B.d.A. sei $|a_{i,1}| = 1$ für $i \in U \cup L$ (nach Skalierung). Für Vektoren $\tilde{a}_i = (a_{i,2}, a_{i,3}, \dots, a_{i,n})$ und $\tilde{x} = (x_2,\dots,x_n)$ nehmen wir die folgenden Ungleichungen in $\tilde{A}x \leqslant \tilde{b}$ auf:

$$\begin{array}{llll} 0x_1 & +\tilde{a}_i^t\tilde{x} & +\tilde{a}_k^t\tilde{x} \leqslant b_i + b_k & i \in U, k \in L \\ 0x_1 & & +\tilde{a}_l^t\tilde{x} \leqslant b_l & l \in N \end{array}$$

Diese bilden zusammen das neue Ungleichungssystem. $\rightsquigarrow$ Jede dieser $|U| \cdot |L|$ Ungleichungen ist Summe von 2 Ungleichungen aus $Ax \leqslant b$.

Die erste Spalte von $\tilde{A}$ ist 0. Alle Spalten von A, die nur 0 enthalten, entsprechen Nullvektoren in $\tilde{A}$. Jede Lösung von $Ax \leqslant b$ ist auch eine von $\tilde{A}x \leqslant \tilde{b}$, da es sich um Linearkombinationen handelt.

Sei $x = (x_1, \ldots, x_n)$ eine Lösung von $\tilde{A}x \leqslant \tilde{b}$. Sei $\tilde{x} = (x_2, \ldots, x_n)$. Setze x_1' auf einen Wert i Intervall

$$[\max\{\tilde{a}_k^t \tilde{x} - b_k | k \in L\}, \min\{b_i - \tilde{a}_i^t \tilde{x} | i \in U\}]$$

Diese ist nach Wahl unseres Ungleichungssystem offenbar nicht leer. Folglich ist $x' = (x_1', x_2, x_3, \ldots, x_n)$ eine Lösung von $Ax \leqslant b$. □

Theorem 23 (Lemma von Farkas)**.** Für $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$ hat das System $Ax \leqslant b$ genau dann eine Lösung, wenn es keinen Vektor $u \in \mathbb{R}^m$ mit $u \geqslant 0$, $u^t A = 0^t$ und $u^t b < 0$ gibt.

Remark 23.1. "Das folgt aus dem Lemma von Farkas- ist in LGO fast immmer eine korrekte Antwort.

Beweis. "$\implies$" Falls $Ax \leqslant b$ $u \in \mathbb{R}^m$, $u \geqslant 0$, $u^t A = 0^t$, $u^t b < 0$, dann gilt

$$0 = u^t A x \leqslant u^t b < 0 \lightning$$

"$\impliedby$"

Annahme: $Ax \leqslant b$ hat keine Lösung. Sei $A^{(0)} := A, b^{(0)} := b$. Falls $N\left(A^{(0)}\right) \neq 0$, wende Fourier-Motzkin-Elimination (Theorem 22) auf $A^{(0)}x \leqslant b^{(0)}$ an.

Erhalte $A^{(1)}x \leqslant b^{(1)}$, das ebenfalls nicht lösbar ist, wobei $N(A^{(1)}) \leqslant N(A^{(0)})$, sodass jede Zeile von $A^{(1)}x \leqslant b^{(1)}$ nicht-negative Linearkombination von Zeilen in $A^{(0)}x \leqslant b^{(0)}$ ist. Iteriere, bis ein (ebenfalls unlösbares) System $A^{(k)}x \leqslant b^{(k)}$ gefunden wird, in dem $N(A^{(k)}) = 0$. Da $A^{(k)}x = 0$, gibt es einen Index i, so dass $b_i^{(k)} < 0$.

Sei $a_i^{(k)} = 0$ die entsprechende Zeile von $A^{(k)}$. Dies ist eine nicht-negative Linearkombination von Zeilen aus $A^{(k-1)}x \leqslant b^{(k-1)}$ und damit nicht-negative Linearkombination von Zeilen aus $A^{(0)}x \leqslant b^{(0)}$.

Folglich existiert $u \in \mathbb{R}^m$ mit $u \geqslant 0$, $u^t A = a_i^{(k)} = 0$, $u^t b = b_i^{(k)} < 0$. □

Theorem 24 (Lemma von Farkas, allgemeinster Fall)**.** Für $A \in \mathbb{R}^{m_1 \times n_1}, B \in \mathbb{R}^{m_1 \times n_2}, C \in \mathbb{R}^{m_2 \times n_1}, D \in \mathbb{R}^{m_2 \times n_2}, a \in \mathbb{R}^{m_1}$ und $b \in \mathbb{R}^{m_2}$ hat genau eines der beiden folgenden Systeme eine Lösung:

System 1:

$$\begin{aligned} Ax &+By \leqslant a \\ Cx &+Dy = b \\ x & \qquad \geqslant 0 \end{aligned}$$

System 2:

$$\begin{aligned} u^tA + v^tC &\geqslant 0^t \\ u^tB + v^tD &= 0^t \\ u &\geqslant 0 \\ u^ta + v^tb &< 0 \end{aligned}$$

Remark 24.1. Der einfache Fall (Theorem 23) ergibt sich, indem man sich auf B und a beschränkt. Es genügt, sich den einfachen Fall zu merken.

Beweis. Das erste System ist äquivalent zu

$$\begin{aligned} Ax+By &\leqslant a \\ Cx+Dy &\leqslant b \\ -Cx-Dy &\leqslant -b \\ -I_{n_1}x \qquad &\leqslant 0 \end{aligned}$$

Nach dem (einfachen) Lemma von Farkas (Theorem 23) hat dieses System genau dann eine Lösung, wenn das folgende System keine Lösung hat:

$$\begin{array}{ccccl} u_1^tA & +u_2^tC & -u_3^tC & -u_4^t & = 0^t \\ u_1^tB & +u_2^tD & -u_3^tD & & = 0^t \\ u_1^ta & +u_2^tb & -u_3^tb & & \leqslant 0 \\ u_1 & & & & \geqslant 0 \\ & u_2 & & & \geqslant 0 \\ & & u_3 & & \geqslant 0 \\ & & & u_4 & \geqslant 0 \end{array}$$

Wir setzen $u := u_1, v := u_2 - u_3$ und entfernen u_4, indem wir aus der ersten Gleichung eine Ungleichung machen. $\square$

Corollary 25 (Weitere Varianten von Farkas)**.** Sei $A \in \mathbb{R}^{m\times n}$,$b \in \mathbb{R}^m$. Dann gilt:

(a) Es existiert $x \in \mathbb{R}^n, x \geqslant 0$ mit $Ax = b$ genau dann, wenn kein $u \in \mathbb{R}^m$ mit $u^t A \geqslant 0^t$ und $u^t b < 0$ existiert.

(b) Es existiert $x \in \mathbb{R}^n$ mit $Ax = b$ genau dann, wenn kein $u \in \mathbb{R}^m$ existiert mit $u^t A = 0^t$ und $u^t b < 0$.

Beweis. Folgt aus Theorem 24, indem nur die Matrix C, bzw. nur die Matrix D verwendet wird. □

Remark 25.1 (Trennende Hyperebene)**.** Für Aussage (a) in Corollary 25 gibt es eine geometrische Interpretation:

Seien $a_1, \dots, a_n$ die Spalten von A. Das System hat genau dann eine Lösung, wenn $b \in \operatorname{cone}(\{a_1, \dots, a_n\})$.

Das Zertifikat entspricht einer Hyperebene, die b von $\operatorname{cone}(\{a_1, \dots, a_n\})$ trennt.

3.2 Starke Dualität

Theorem 26 (Starke Dualität)**.** Für zwei lineare Programme

$$\begin{aligned} \max\ & c^t x \quad (P) \\ \text{s.d. } & Ax \leqslant b \end{aligned}$$

und

$$\begin{aligned} \min\ & b^t y \quad (D) \\ \text{s.d. } & A^t y = c \\ & y \geqslant 0 \end{aligned}$$

gilt genau eine der folgenden Aussagen:

- (P) und (D) sind beide unzulässig.
- Eines ist unbeschränkt und das andere ist unzulässig.
- Beide sind zulässig. Dann haben beide eine Optimallösung und für jede Optimallösung $\tilde{x}$ von (P) und $\tilde{y}$ von (D) gilt:

$$c^t \tilde{x} = b^t \tilde{y}$$

Beweis. Offenbar kann nur einer der Fälle eintreten. Wenn eines der LPs unbeschränkt ist, muss auf Grund der schwachen Dualität (Theorem 20) das andere unzulässig sein.

Angenommen eines der LPs ist zulässig und beschränkt. O.B.d.A. sei dies (P).

Dann hat das System

$$Ax \leqslant b$$

eine Lösung und es existiert $B \in \mathbb{R}$, sodass

$$\begin{aligned} Ax &\leqslant b \\ -c^t x &\leqslant -B \end{aligned}$$

keine Lösung besitzt.

Nach dem Lemma von Farkas (Theorem 23) existiert daher ein $u \in \mathbb{R}^m$ und ein $z \in \mathbb{R}$, sodass $u^t A - zc^t = 0^t$ und $b^t u - zB < 0$ Hierbei ist $z > 0$, denn $z = 0$ widerspräche der Lösbarkeit von $Ax \leqslant b$.

Sei $\tilde{u} := \frac{1}{z} u$. Dann ist $A^t \tilde{u} = c$ und $\tilde{u} \geqslant 0$. Somit ist $\tilde{u}$ eine Lösung von (D).

Daher ist (D) zulässig. Auf Grund der schwachen Dualität (Theorem 20) ist (D) beschränkt.

Zu zeigen ist noch, dass es Lösungen x von (P) und y von (D) gibt, sodass $c^t x \geqslant b^t y$. ($\leqslant$ folgt schon aus Theorem 20)

Das ist genau dann der Fall, wenn das folgende System lösbar ist:

$$\begin{array}{rrl} Ax & & \leqslant b \\ & A^t y & = c \\ -c^t x & +b^t y & \leqslant 0 \\ & y & \geqslant 0 \end{array}$$

Nach Farkas ist das genau dann der Fall, wenn das folgende System keine Lösung hat:

$$\begin{array}{rrrl} u^t A & & -wc^t & = 0 \\ & v^t A^t & +wb^t & \geqslant 0 \\ u^t b & +v^t c & & < 0 \\ u & & & \geqslant 0 \\ & & w & \geqslant 0 \end{array}$$

Angenommen es gäbe eine solche Lösung u, v, w.

Wir unterscheiden zwei Fälle:

- $w = 0$. Dann wären u, v bereits ein Zertifikat dafür, dass

$$\begin{aligned} Ax &\leqslant b \\ A^t y &= c \\ y &\geqslant 0 \end{aligned}$$

nicht lösbar ist ↯.

- $w > 0$. Dann ist

$$0 > wu^t b + wv^t c \geqslant u^t(-Av) + v^t(A^t u) = 0 ↯$$

□

Remark 26.1. Theorem 26 zeigt insbesondere, dass für ein zulässiges, beschränktes LP $\max\{c^t x | Ax \leqslant b\}$ ein Vektor $\tilde{x}$ mit $A\tilde{x} \leqslant b$ existiert, so dass $c^t \tilde{x} = \sup\{c^t x | Ax \leqslant b\}$.

Remark 26.2. Mit Theorem 26 kann man zeigen, dass das Finden einer beliebigen Lösung eines LPs genau so schwer ist, wie das Berechnen einer optimalen Lösung: Betrachte zum Lösen des Programmes (P) das folgende Programm:

$$\begin{aligned} \max \quad & c^t x \\ \text{s.d.} \quad & Ax \leqslant b \\ & A^t y = c \\ & c^t x \geqslant b^t y \\ & y \geqslant 0 \end{aligned}$$

Wegen Theorem 26 ist jede zulässige Lösung dieses LPs bereits eine optimale Lösung.

Corollary 27. Seien $A, B, C, D, E, F, G, H, I$ Matrizen und a, b, c, d, e, f Vektoren passender Dimension, so dass

$$\begin{pmatrix} A & B & C \\ D & E & F \\ G & H & I \end{pmatrix} \text{ eine } m \times n\text{- Matrix ist und}$$

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \text{ ein Vektor der Länge } m \text{ sowie } \begin{pmatrix} d \\ e \\ f \end{pmatrix} \text{ ein Vektor der Länge } n \text{ ist.}$$

Dann ist

$$\max\left\{ d^t x + e^t y + f^t z \;\middle|\; \begin{array}{ccccc} Ax & +By & +Cz & \leqslant & a \\ Dx & +Ey & +Fz & = & b \\ Gx & +Hy & +Iz & \geqslant & c \\ x & & & \geqslant & 0 \\ & & z & \leqslant & 0 \end{array} \right\}$$
$$=$$
$$\min\left\{ a^t u + b^t v + c^t w \;\middle|\; \begin{array}{ccccc} A^t u & +D^t v & +G^t w & \geqslant & d \\ B^t u & +E^t v & +H^t w & = & e \\ C^t u & +F^t v & +I^t w & \leqslant & f \\ u & & & \geqslant & 0 \\ & & w & \leqslant & 0 \end{array} \right\}$$

sofern beide Mengen nicht leer sind.

Beweis. Transformiere das LP in Standardungleichungsform und wende Theorem 26 an. □

3.3 Übersicht Dualisierung

Tabelle 2: Primales und duales LP

(P) \ (D)	zulässig, beschränkt	zulässig, unbeschränkt	unzulässig
zulässig, beschränkt	✓		
zulässig, unbeschränkt			✓
unzulässig		✓	✓

Tabelle 3: Dualisierung

	primal	dual
Variablen	$x_1, \dots, x_n$	$y_1, \dots, y_m$
Matrix	A	A^t
RHS	b	c
Zielfunktion	$\max c^t x$	$\min b^t y$
Nebenbedingungen	$\sum_{j=1}^n a_{ij} x_j \leqslant b_i$	$y_i \geqslant 0$
	$\sum_{j=1}^n a_{ij} x_j \geqslant b_i$	$y_i \leqslant 0$
	$\sum_{j=1}^n a_{ij} x_j = b_i$	$y_i \in \mathbb{R}$
	$x_j \geqslant 0$	$\sum_{i=1}^m a_{ij} y_i \geqslant c_j$
	$x_j \leqslant 0$	$\sum_{i=1}^m a_{ij} y_i \leqslant c_j$
	$x_j \in \mathbb{R}$	$\sum_{i=1}^m a_{ij} y_i = c_j$

Tabelle 4: Dualisierung - Spezialfälle

primal	dual
$\max\{c^t x \| Ax \leqslant b\}$	$\min\{b^t y \| x^t A = c, y \geqslant 0\}$
$\max\{c^t x \| Ax \leqslant b, x \geqslant 0\}$	$\min\{b^t y \| x^t A \geqslant c, y \geqslant 0\}$
$\max\{c^t x \| Ax \geqslant b, x \geqslant 0\}$	$\min\{b^t y \| y^t A \geqslant c, y \leqslant 0\}$
$\max\{c^t x \| Ax = b, x \geqslant 0\}$	$\min\{b^t y \| y^t A \geqslant c\}$

3.4 Komplementärer Schlupf

Theorem 28 (Komplementärer Schlupf für Ungleichungen)**.**
Sei $\max\{c^t x | Ax \leqslant b\}$ (P) und $\min\{b^t y | A^t y = c, y \geqslant 0\}$ (D) ein Paar von primalem und dualem LP. Dann sind für $x \in \mathbb{R}^n$ mit $Ax \leqslant b$ und $y \in \mathbb{R}^m$ mit $A^t y = c, y \geqslant 0$ die folgenden Aussagen äquivalent:

(a) x ist eine optimale Lösung von (P) und y eine optimale Lösung von (D).

(b) $c^t x = b^t y$

(c) $y^t(b - Ax) = 0$, d.h. $\forall i : y_i = 0 \vee (b - Ax)_i = 0$ wegen $Ax \leqslant b, y \geqslant 0$.

Beweis. (a) $\iff$ (b) folgt aus Theorem 26.

(b) $\iff$ (c) Es ist

$$\begin{aligned} y^t(b - Ax) &= y^t b - y^t Ax \\ &= y^t b - c^t x \end{aligned}$$

□

Theorem 29 (Komplementärer Schlupf für Ungleichungen mit nicht-negativen Variablen)**.** Betrachte $\max\{c^t x | Ax \leqslant b, x \geqslant 0\}$ und $\min\{b^t < | A^t y \geqslant c, y \geqslant 0\}$.

Dann sind für $x \in \mathbb{R}^n$ mit $Ax \leqslant b, x \geqslant 0$ und $y \in \mathbb{R}^m$ mit $A^t y \geqslant c$, $y \geqslant 0$ äquivalent:

(a) x und y sind Optimallösungen.

(b) $c^t x = b^t y$

(c) $y^t(b - Ax) = 0$ und $x^t(A^t y - c) = 0$

Beweis. (a) $\iff$ (b) folgt wieder aus Theorem 26.

(b) $\iff$ (c): Es gilt

$$\begin{aligned} 0 &\leqslant y^t(b - Ax) \\ 0 &\leqslant x^t(A^t y - c) \end{aligned}$$

Folglich ist

$$y^t(b - Ax) + x^t(A^t y - c) = y^t b - x^t c$$

genau dann 0 wenn $y^t(b - Ax) = 0$ und $x^t(A^t y - c) = 0$. □

Lemma 30. Sei $\max\{c^t | Ax \leqslant b\}$ ein zulässiges LP. Dann ist dieses genau dann beschränkt, wenn $c \in \operatorname{cone}(\{a_1, \ldots, a_m\})$, wobei a_i die Zeilen von A seien.

Beweis. Das LP ist genau dann beschränkt, wenn $\min\{b^t y | A^t y = c, y \geqslant 0\}$ zulässig ist (Theorem 26). Das ist äquivalent zu $c \in \operatorname{cone}(\{a_1, \ldots, a_m\})$. □

Dies können wir mit dem komplementären Schlupf wie folgt verschärfen:

Corollary 31. Es sei x eine Optimallösung von $\max\{c^t x | Ax \leqslant b\}$ und y eine Optimallösung des dualen LPs $\min\{b^t y | A^t y = c, y \geqslant 0\}$. Es seien $a_1, \ldots, a_m$ die Zeilen von A.

Dann liegt c in dem Kegel, der von den Zeilen a_i von A erzeugt wird, für die $a_i^t x = b_i$ gilt.

Beweis. Es gilt

$$c = \sum_{i=1}^{m} y_i a_i = \sum_{i \in \{1,\ldots,m\}, a_i^t x = b} y_i a_i$$

wobei die zweite Gleichheit aus Theorem 28 ($a_i^t x < b_i \implies y_i = 0$) folgt. □

Theorem 32 (Starker komplementärer Schlupf)**.**

Betrachte $\max\{c^t x | Ax \leqslant b\}$ (P) und $\min\{b^t y | A^t y = c, y \geqslant 0\}$ (D).

Dann gilt für jede Ungleichung $a_i^t x \leqslant b_i$ in $Ax \leqslant b$ genau eine der folgenden Aussagen:

(a) (P) eine eine Optimallösung x^* mit $a_i^t x^* < b_i$.

(b) (D) hat eine Optimallösung y^* mit $y_i^* > 0$.

Beweis. Aus Theorem 28 folgt, dass nicht beide Aussagen erfüllt sein können.

Angenommen (a) gilt nicht. Sei δ der Wert einer optimalen Lösung. Betrachte

$$\begin{aligned} \max\ & -a_i^t x \\ \text{s.d.}\ & \quad Ax \leqslant b \\ & -c^t x \ \leqslant\ -\delta \end{aligned}$$

Dieses Problem hat einen optimalen Lösungswert von $-b_i$. Das dazu duale LP

$$\begin{array}{rll} \min & b^t y & -\delta u \\ & A^t y & -uc = -a_i \\ & y & \geqslant 0 \\ & & u \geqslant 0 \end{array}$$

hat daher eine Optimallösung $y \geqslant 0, u \geqslant 0$ mit Wert $-b_i$.

Es ist $y^t A - uc^t = -a_i^t$ und $y^t b - u\delta = -b_i$. Sei $\tilde{y} = y + e_i$.

- Wenn $u = 0$, dann gilt $\tilde{y}^t A = y^t A + a_i = 0$ und $\tilde{y}^t b = y^t b + b_i = 0$. Wenn y^* eine optimale Lösung von (D) ist, dann ist $y^* + \tilde{y}$ eine optimale Lösung von (D) mit positivem i-ten Eintrag.
- Wenn $u > 0$, dann ist $\frac{1}{u}\tilde{y}$ eine optimale Lösung von (D), da $\frac{1}{u}\tilde{y}^t A = \frac{1}{u} y^t A + \frac{1}{u} a_i^t = c^t$ und $\frac{1}{u}\tilde{y}^t b = \frac{1}{u} y^t b + \frac{1}{u} b_i \leqslant \delta$. Ferner hat $\frac{1}{u}\tilde{y}$ einen positiven i-ten Eintrag.

□

Durch Linearkombination solcher Lösungen für alle i kann ein Paar von Lösungen gefunden werden, für die in jedem Eintrag (a) oder (b) gilt:

Theorem 33. Sei wieder $\max\{c^t x | Ax \leqslant b\}$ und $\min\{b^t y | A^t y = c, y \geqslant 0\}$ ein Paar von primalem und dualem LP, welche beide zulässig und beschränkt seien.

Dann existieren optimale Lösungen x^* und y^*, sodass für alle i entweder $a_i^t x_i^* < b_i$ oder $y_i^* > 0$.

3.5 Anwendung: Max-Flow-Min-Cut

Problem (Maximum-Flow-Problem)**.** *Gegeben sei ein gerichteter Graph G mit Kapazitäten $u : E(G) \to \mathbb{R}_{\geqslant 0}$ und Knoten $s \neq t \in V(G)$.*

Gesucht ist ein s-t-Fluss $f : E(G) \to \mathbb{R}_{\geqslant 0}$ mit maximalem Wert.

Dies lässt sich als LP formulieren: (Wir gehen davon aus, dass es keine eingehenden Kanten in s und keine ausgehenden Kanten aus t gibt.)

$$
\begin{array}{lrl}
\max & \displaystyle\sum_{e \in \delta_G^+(s)} x_e & \\
\text{s.d.} & x_e \geqslant 0 & \forall e \in E(G) \\
& x_e \leqslant u(e) & \forall e \in E(G) \\
& \displaystyle\sum_{e \in \delta_G^+(v)} x_e - \sum_{e \in \delta_G^-(v)} x_e = 0 & \forall v \in V(G) \backslash \{s,t\}
\end{array}
$$

Als duales LP ergibt sich

$$
\begin{array}{lrl}
\min & \displaystyle\sum_{e \in E(G)} u(e) y_e & \\
\text{s.d.} & y_e \geqslant 0 & \forall e \in E(G) \\
& y_e + z_v - z_w \geqslant 0 & \forall e = (v,w) \in E(G), \{s,t\} \cap \{v,w\} = \varnothing \\
& y_e + z_v \geqslant 0 & \forall e = (v,t) \in E(G), v \neq s \\
& y_e - z_w \geqslant 1 & \forall e = (s,w) \in E(G), w \neq t \\
& y_e \geqslant 1 & \forall e = (s,t) \in E(G)
\end{array}
$$

Dies lässt sich vereinfachen zu:

$$
\begin{array}{lrl}
\min & \displaystyle\sum_{e \in E(G)} u(e) y_e & \\
\text{s.d.} & y_e \geqslant 0 & \forall e \in E(G) \\
& y_e + z_v - z_w \geqslant 0 & \forall e = (v,w) \in E(G) \\
& z_s = -1 & \\
& z_t = 0 &
\end{array}
$$

(z_s und z_t wurden hinzugefügt, um das System zu vereinfachen).

Es folgt das bereits bekannte Max-Flow-Min-Cut-Theorem:

Theorem 34 (Max-Flow-Min-Cut)**.** Sei G ein gerichteter Graph mit Kapazitäten $u : E(G) \to \mathbb{R}_{\geqslant 0}$, $s, t \in V(G)$. Dann ist das Minimum aller s-t-Schnitte gleich dem maximalen Wert eines s-t-Flusses.

Beweis. Offenbar ist der Flusswert durch die minimale Kapazität eines Schnittes beschränkt.

Sei $\tilde{x}$ eine optimale primale Lösung und $\tilde{y}, \tilde{z}$ eine optimale Duallösung.

Betrachte $R := \{v \in V(G) | \tilde{z}_v \leqslant -1\}$. Dann ist $s \in \mathbb{R}$ und $t \notin R$.

Falls $e = (v, w) \in \delta_G^+(R)$, dann ist $\tilde{z}_w > \tilde{z}_v$. Folglich ist $\tilde{y}_e \geqslant \tilde{z}_w - \tilde{z}_v > 0$.

Mit dem komplementären Schlupf folgt $\tilde{x}_e = u(e)$. Falls $e = (v, w) \in \delta_G^-(R)$, dann gilt $\tilde{z}_v > \tilde{z}_w$ und daher $\tilde{y}_e + \tilde{z}_v - \tilde{z}_w > 0$. Nach komplementärem Schlupf ist $\tilde{x}_e = 0$.

Folglich ist $\text{val}(f) = u(\delta^+(R))$ mindestens so groß wie die Kapazität eines minimalen Schnittes. □

4 Die Struktur von Polyedern

4.1 Abbildungen von Polyedern

Proposition 35. Sei $A \in \mathbb{R}^{m \times (n+k)}$ und $b \in \mathbb{R}^m$. Dann ist die Menge

$$P = \{x \in \mathbb{R}^n | \exists y \in \mathbb{R}^k : A \begin{pmatrix} x \\ y \end{pmatrix} \leqslant b\}$$

ein Polyeder.

Beweis. Induktion nach k. Für $k = 0$ ist die Aussage trivial. Sei nun $k > 0$ und die Aussage für $k' = k - 1$ bereits gezeigt. Wende einen Schritt der Fourier-Motzkin-Elimination auf $A \begin{pmatrix} x \\ y \end{pmatrix} \leqslant b$ an, um eine der Variablen aus y zu eliminieren. Hierdurch erhält man $A' \in \mathbb{R}^{m' \times (n+k-1)}, b' \in \mathbb{R}^{m'}$, so dass

$$P = \{x \in \mathbb{R}^n | \exists y \in \mathbb{R}^k : A \begin{pmatrix} x \\ y \end{pmatrix} \leqslant b\} = \{x \in \mathbb{R}^n | \exists y' \in \mathbb{R}^{k-1} : A' \begin{pmatrix} x \\ y' \end{pmatrix} \leqslant b'\}$$

□

Notation 35.1. Die Menge $P = \{x \in \mathbb{R}^n | \exists y \in \mathbb{R}^k : A \begin{pmatrix} x \\ y \end{pmatrix} \leqslant b\}$ heißt **Projektion** von $\{z \in \mathbb{R}^{n+k} | Az \leqslant b\}$ auf $\mathbb{R}^n$.

Corollary 36. Es seien $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, D \in \mathbb{R}^{k \times n}$ und $d \in \mathbb{R}^k$. Dann ist

$$\{y \in \mathbb{R}^k | \exists x \in \mathbb{R}^n : Ax \leqslant b \wedge y = Dx + d\}$$

ein Polyeder.

Beweis. Es gilt

$$\begin{aligned}&\{y \in \mathbb{R}^k | \exists x \in \mathbb{R}^n : Ax \leqslant b \wedge y = Dx + d\}\\ = &\{y \in \mathbb{R}^k | \exists x \in \mathbb{R}^n : \begin{pmatrix} A & 0 \\ D & -I_k \\ -D & I_k \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leqslant \begin{pmatrix} b \\ -d \\ d \end{pmatrix} \}\end{aligned}$$

Die Aussage folgt damit aus Proposition 35. □

4.2 Flächen

Definition 37. Sei $P = \{x \in \mathbb{R}^n | Ax \leqslant b\} \neq \emptyset$ ein Polyeder und $c \in \mathbb{R}^n \backslash \{0\}$.

- Für $\delta := \max\{c^t x | x \in P\} < \infty$ heißt $\{x \in \mathbb{R}^n | c^t x = \delta\}$ die **Stützhyperebene** (**supporting hyperplane**) von P.
- Eine Menge $X \subseteq \mathbb{R}^n$ heißt **Fläche** (**face**) von P, falls $X = P$, oder falls eine Stützhyperebene H von P existiert, sodass $X = P \cap H$.
- Falls $\{x'\}$ eine Fläche von P ist, heißt x' **Ecke** (**vertex**) von P. bzw. **Basislösung** (**basic solution**) des Systems $Ax \leqslant b$.

Observe. *Nicht jedes Polyeder hat Ecken.*

Proposition 38. Sei $P = \{x \in \mathbb{R}^n | Ax \leqslant b\} \neq \emptyset$ und $F \subseteq P$. Dann sind äquivalent:

(a) F ist Fläche von P.

(b) Es existiert $c \in \mathbb{R}^n$, sodass $\delta := \max\{c^t x | x \in P\} < \infty$ und $F = \{x \in P | c^t x = \delta\}$.

(c) Es existiert ein Teilsystem $A'x \leqslant b'$ von $Ax \leqslant b$, sodass $F = \{x \in P | A'x = b'\} \neq \emptyset$.

Beweis.

- (a) $\Longrightarrow$ (b): Sei F Fläche von P. Falls $F = P$, setze $c = 0$. Falls $F \neq P$ existiert ein solches c nach Definition.
- (b) $\Longrightarrow$ (c): Seien c, δ, F wie in (b). Sei $J \subseteq \{1, \ldots, m\}$ so gewählt, dass $A_J x \leqslant b_J$ ein maximales Teilsystem mit $A_J x = b_J$ für alle $x \in F$ ist. Zu zeigen ist, dass $F = \{x \in P | A_J x = b_J\}$,

  Nach Theorem 33 existiert eine optimale duale Lösung y^* zum primalen LP $\max\{c^t x | Ax \leqslant b\}$, sodass $y_j^* > 0$ für alle $j \in J$ und $y_j^* = 0$ für $j \notin J$. Wegen Theorem 28 ist $x \in P$ genau dann optimal (d.h. $\in F$) wenn $y^*(b - Ax) = 0$, d.h. genau dann wenn $A_J x = b_J$.
- (c) $\Longrightarrow$ (a): Sei $A'x \leqslant b'$ ein Teilsystem und $F = \{x \in P | A'x = b'\}$. Sei c^t die Summe aller Zeilenvektoren von A' und sei δ die Summe aller Einträge

von b'. Dann gilt $c^t x \leqslant \delta$ für alle $x \in P$. Ferner ist $F = P \cap \{x \in \mathbb{R}^n | c^t x = \delta\}$.

$\square$

Man folgert leicht:

Corollary 39. Sei $P \neq \emptyset$ ein Polyeder und F eine Fläche von P.

(a) Sei $c \in \mathbb{R}^n$ mit $\max\{c^t x | x \in P\} < \infty$. Dann ist $\operatorname{argmax}\{c^t x | x \in P\}$ eine Fläche von P.

(b) F ist ein Polyeder.

(c) Eine Teilmenge $F' \subseteq F$ ist genau dann eine Fläche von F, wenn F' eine Fläche von P ist.

(d) Wenn P von der Form $P = \{x \in \mathbb{R}^n | Ax = b\}$, dann hat P genau eine Fläche, nämlich P selbst.

4.3 Facetten

Definition 40. Sei P ein Polyeder. Eine **Facette** (**facet**) von P ist eine inklusionsmaximale Fläche $F \subsetneq P$.

Eine Ungleichung $c^t x \leqslant \delta$ heißt **facettenbestimmend** (**facet-defining**) für P, wenn $c^t x \leqslant \delta$ für alle $x \in P$ gilt und $\{x \in P | c^t x = \delta\}$ eine Facette ist.

Theorem 41. Es sei $P \subseteq \{x \in \mathbb{R}^n | Ax = b\}$ ein nicht-leeres Polyeder der Dimension $n - \operatorname{rank}(A)$. Es sei $A'x \leqslant b'$ ein kleinstes Ungleichungssystem, sodass $P = \{x \in \mathbb{R}^n | Ax = b, A'x \leqslant b'\}$. Dann ist jede Ungleichung in $A'x \leqslant b'$ facettenbestimmend für P und jede Facette von P wird durch eine Ungleichung von $A'x \leqslant b'$ gegeben.

Beweis. Falls $P = \{x \in \mathbb{R}^n | Ax = b\}$, dann hat P keine Facette.

Sei daher $P \subsetneq \{x \in \mathbb{R}^n | Ax = b\}$. Sei $A'x \leqslant b'$ ein kleinstes System von Ungleichungen, sodass $P = \{x \in \mathbb{R}^n | Ax = b, A'x \leqslant b'\}$.

Sei $a^t x \leqslant \beta$ eine Ungleichung in $A'x \leqslant b'$ und sei $A''x \leqslant b''$ der Rest des Systems $A'x \leqslant b'$.

Behauptung 1. *$a^t x \leqslant \beta$ ist facettenbestimmend.*

Unterbeweis. Sei $y \in \mathbb{R}^n$ so, dass $Ay = b$ und $A''y \leqslant b''$, $a^t y > \beta$. Ein solches y existiert, das $a^t y \leqslant \beta$ nicht redundant ist.

Sei ferner $\tilde{y} \in P$ mit $A'\tilde{y} < b'$. $\tilde{y}$ existiert, weil $\dim(P) = n - \operatorname{rank}(A)$.

Sei $z = \tilde{y} + \frac{\beta - a^t\tilde{y}}{a^t y - a^t\tilde{y}}(y - \tilde{y})$. Es ist $a^t z = \beta$. Ferner ist $0 < \frac{\beta - a^t\tilde{y}}{a^t y - a^t\tilde{y}} < 1$, folglich ist $z \in \operatorname{conv}(\{y, \tilde{y}\})$. Sei $F := \{x \in P | a^t x = \beta\}$. Es ist $z \in F$, folglich $F \neq \emptyset$. Da $\tilde{y} \in P \backslash F$ ist $F \neq P$.

Da $a^t x \leqslant \beta$ die einzige Ungleichung von $A'x \leqslant b'$ ist, welche von allen Elementen von F mit Gleichheit erfüllt wird (betrachte z), ist F eine Facette. ■

Andererseits ist nach Proposition 38 jede Facette durch eine Ungleichung aus $A'x \leqslant b'$ definiert. □

Corollary 42. Sei $P \subseteq \mathbb{R}^n$ ein Polyeder

(a) Jede Fläche F von P mit $F \neq P$ ist ein Schnitt von Facetten.

(b) Die Dimension jeder Facette von P ist $\dim(P) - 1$.

Remark 42.1. Insbesondere liefern uns Polyederbeschreibungen mit facettenbestimmenden Ungleichungen kleinste Darstellungen des Polyeders.

4.4 Minimale Flächen

Neben den Facetten (größte Flächen, welche nicht das ganze Polyeder sind) interessieren uns auch die minimalen Flächen:

Definition 43. Eine Fläche F eines Polyeders heißt **minimale Fläche**, wenn es keine Fläche F' von P mit $F' \subsetneq F$ gibt.

Theorem 44. Sei $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ ein Polyeder. Eine nichtleere Menge $F \subseteq P$ ist genau dann eine minimale Fläche von P, wenn es ein Teilsystem $A'x \leqslant b'$ von $Ax \leqslant b$ mit $F = \{x \in \mathbb{R}^n | A'x = b'\}$ gibt.[a]

[a]Anders als bei der Definition einer Fläche, wird hier nicht mit P geschnitten!

Beweis. „ $\Longrightarrow$ " Sei F minimale Fläche von P. Dann gibt es ein Teilsystem $A'x \leqslant b'$ von $Ax \leqslant b$ mit $F = \{x \in P | A'x = b'\}$. Wähle $A'x \leqslant b'$ maximal mit dieser Eigenschaft.

Sei $\tilde{A}x \leqslant \tilde{b}$ ein kleinstes Teilsystem von $Ax \leqslant b$, so dass $F = \{x \in \mathbb{R}^n | \tilde{A}x \leqslant \tilde{b}, A'x = b'\}$.

Behauptung 1. *$\tilde{A}x \leqslant \tilde{b}$ ist ein leeres System.*

Unterbeweis. Angenommen das System ist nicht leer. Sei $a^t x \leqslant \beta$ eine Ungleichung aus $\tilde{A}x \leqslant \tilde{b}$. Sei $F' = \{x \in \mathbb{R}^n | A'x = b', \tilde{A}x \leqslant \tilde{b}, a^t x = \beta\}$. Es ist $F' \neq \emptyset$, da sonst $a^t x \leqslant \beta$ aus $\tilde{A}x \leqslant \tilde{b}$ eliminiert werden könnte.

F' ist eine Fläche von F und $F' \neq F$ (sonst könnte $a^t x = \beta$ bereits zu $A'x = b'$ hinzugefügt werden), im Widerspruch zur Minimalität von F. ■

„ $\Longleftarrow$ " Sei $F = \{x \in \mathbb{R}^n | A'x = b'\} \subseteq P$ für ein Teilsystem $A'x \leqslant b'$ von $Ax \leqslant b$ mit $F \neq \varnothing$.

Dann ist F ein Polyeder und F ist die einzige Fläche von F. Ferner ist $F = \{x \in \mathbb{R}^n | A'x = b'\} = \{x \in P | A'x = b'\}$, insbesondere eine Fläche von P. Daher ist F minimale Fläche von P. □

Corollary 45. Sei $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ ein Polyeder. Dann hat jede minimale Fläche von P Dimension $n - \operatorname{rank}(A)$.

Beweis. Sei F eine minimale Fläche von P. Dann gibt es ein Teilsystem $A'x \leqslant b'$ von $Ax \leqslant b$, sodass $F = \{x \in \mathbb{R}^n | A'x = b'\}$.

Es ist $\dim F = n - \operatorname{rank} A'$. Offenbar ist $\operatorname{rank} A' \leqslant \operatorname{rank} A$.

Falls $\operatorname{rank} A' < \operatorname{rank} A$, so können Nebenbedingungen $a^t x = \beta$ aus $Ax \leqslant b$ zu $A'x = b'$ hinzugefügt werden, sodass a^t von den Zeilen von A' linear unabhängig ist. Es folgt, dass $F' = \{x \in \mathbb{R}^n | A'x = b', a^t x = \beta\} \subsetneq F$ und $F' \neq \varnothing$. Dies widerspricht der Minimalität von F. □

Proposition 46. Es sei $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ ein Polyeder und $x' \in P$. Dann sind die folgenden Aussagen äquivalent:

(a) x' ist eine **Ecke** von P (einziges Element einer Fläche)

(b) Es gibt ein Teilsystem $A'x \leqslant b'$ von $Ax \leqslant b$ aus n Ungleichungen, sodass die Zeilen von A' linear unabhängig sind und $x' = \{x \in P | A'x = b'\}$ gilt.

(c) x' kann nicht als Konvexkombination von Vektoren in $P \backslash \{x'\}$ geschrieben werden.

(d) Es gibt keinen von Nullvektor verschiedenen Vektor $d \in \mathbb{R}^n$, für den $\{x' + d, x' - d\} \subseteq P$ gilt.

Beweis. $(a) \iff (b)$: x' ist genau dann eine Ecke von P, wenn es ein Teilsystem $A'x \leqslant b'$ von $Ax \leqslant b$ mit $\{x'\} = \{x \in \mathbb{R}^n | A'x = b'\}$. Wegen $\dim\{x'\} = 0$ ist das äquivalent zu (b).

$(b) \implies (c)$ Seien A', b', x' wie in (b). Angenommen x' kann als Konvexkombination $x' = \sum_{i=1}^k \lambda_i x^{(i)}$ von Vektoren $x^{(i)} \in P \backslash \{x'\}$ geschrieben werden.

Wenn $a^t x^{(i)} < \beta$ für eine Ungleichung $a^t x \leqslant \beta$ aus $A'x \leqslant b'$ und $i \in \{1, \ldots, k\}$, dann wäre $a^t y' = \sum_{i=1}^k \lambda_i a^t x^{(i)} < \beta$. Es folgt $a^t x^{(i)} = \beta$ für alle $i \in \{1, \ldots, k\}$ im Widerspruch dazu, dass $x^{(i)} \in P \backslash \{x'\}$.

$(c) \implies (d)$ trivial

$(d) \implies (b)$ Sei $A'x \leqslant b'$ maximales Teilsystem von $Ax \leqslant b$ mit $A'x' = b'$. Angenommen in A' gibt es keine n linear unabhängige Zeilenvektoren. Dann gibt es $d \in \mathbb{R}^n$, sodass d orthogonal zu allen Zeilen von A' steht. Für jedes $\varepsilon > 0$ gilt $A'(x' + \varepsilon d) = A'(x' - \varepsilon d) = b'$. Für jede Ungleichung $a^t x \leqslant \beta$, die in $Ax \leqslant b$, aber nicht in $A'x \leqslant b'$ liegt, gilt $a^t x' < \beta$.

Insbesondere können wir auf Grund der Stetigkeit linearer Abbildungen ε so wählen, dass $\{x' + \varepsilon d, x' - \varepsilon d\} \subseteq P$.

□

Definition 47. Ein Polyeder heißt **spitz** (**pointed**), wenn es leer ist oder alle seine minimalen Flächen Dimension 0 haben.

Example 48.

- Polytope sind spitz: Sei $P = \{x \in \mathbb{R}^n | Ax \leqslant b\} \neq \varnothing$ ein Polytop. Falls $\operatorname{rank}(A) < n$, so gibt es einen Vektor $\tilde{x}$, der senkrecht auf allen Zeilen in A steht. Insbesondere ist $x_0 + \mathbb{R} \cdot \tilde{x} \subseteq P$ für jedes $x_0 \in P$ $\lightning$.
- Polyeder, die als $P = \{x \in \mathbb{R}^n | Ax = b, x \geqslant 0\}$ geschrieben werden können, sind spitz, denn $P = \{x \in \mathbb{R}^n | \begin{pmatrix} A \\ -A \\ -I_n \end{pmatrix} x \leqslant \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}\}$, wobei der Rang der Matrix offenbar n ist.

Corollary 49. Wenn das LP $\max\{c^t x | Ax \leqslant b\}$ zulässig und beschränkt und $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ spitz ist, dann gibt es eine Ecke x', so dass $c^t x' = \max\{c^t x | x \in P\}$.

Beweis. Die Menge der Optimallösungen bildet eine Fläche. Insbesondere ist diese spitz und wir erhalten induktiv eine Ecke x', welche eine Optimallösung ist. □

Theorem 50 (Carathéodorys Theorem)**.** Wenn $X \subseteq \mathbb{R}^n$ eine endliche Menge von Vektoren und $c \in \operatorname{cone}(X)$ ist, dann gibt es linear unabhängige Vektoren $a_1, \ldots, a_k \in X$, sodass $c \in \operatorname{cone}(\{a_1, \ldots, a_k\})$.

Beweis. Sei $\{a_1, \ldots, a_k\} \subseteq X$ minimal mit $c \in \operatorname{cone}(\{a_1, \ldots, a_k\})$. Dann existieren $\lambda_1, \ldots, \lambda_k \geqslant 0$ so dass $c = \sum_{i=1}^k \lambda_i a_i$. Angenommen $a_1, \ldots, a_k$ sind nicht linear unabhängig. Dann gibt es Zahlen $\gamma_1, \ldots, \gamma_k$ mit $\sum_{i=1}^k \gamma_i a_i = 0$. O.B.d.A. ist mindestens ein γ_i positiv. Sei σ maximal, sodass $\lambda_i - \sigma\gamma_i \geqslant 0$ für alle $i \in \{1, \dots, k\}$. Insbesondere gibt es ein $i \in \{1, \dots, k\}$, so dass $\lambda_i - \sigma\gamma_i = 0$. Jedoch ist dann $c = \sum_{i=1}^{k} (\lambda_i - \sigma\gamma_i) a_i$ eine Darstellung von c mit weniger Vektoren $\lightning$

$\square$

Theorem 51 (Fundamentalsatz der linearen Ungleichungen)**.**
Seien $a_1, \dots, a_m, c \in \mathbb{R}^n$ und t die Dimension des Unterraums von $\mathbb{R}^n$, der von $a_1, \dots, a_m, c$ aufgespannt wird (d.h $t = \operatorname{rank} (a_1 \ \dots \ a_m \ c)$). Dann gilt genau eine der folgenden Aussagen:

(a) c kann als nicht-negative Kombination von linear unabhängigen Vektoren aus $a_1^t, \dots, a_m^t$ geschrieben werden.[a]

(b) Es gibt eine Hyperebene $\{x \in \mathbb{R}^n | u^t x = 0\}$ für ein $u \in \mathbb{R}^n \backslash \{0\}$, die $t-1$ linear unabhängige Vektoren aus $a_1, \dots, a_m$ enthält, sodass $a_i^t u \geqslant 0$ für $i \in \{1, \dots, m\}$ und $c^t u < 0$.

[a] Eine beliebige nicht-negative Kombination genügt, siehe Theorem 50

Beweis. Offenbar können nicht beide Aussagen gelten.

Zu zeigen ist, dass mindestens eine der Assagen gilt. Falls $c \in \operatorname{cone}(\{a_1, \dots, a_m\})$, folgt (a) aus Carathéodorys Theorem (Theorem 50).

Falls $c \notin \operatorname{cone}(\{a_1, \dots, a_m\})$, dann gibt es keinen Vektor $v \in \mathbb{R}^m, v \geqslant 0$ mit $c^t = v^t A$, wobei A die Matrix mit Zeilen a_i^t sei.

Mit Farkas folgt, dass ein $\tilde{u} \in \mathbb{R}^n$ existiert mit $A\tilde{u} \geqslant 0, c^t \tilde{u} < 0$. Somit ist das folgende LP zulässig (offenbar ist es auch beschränkt).

$$\begin{aligned} \max \quad & c^t u \\ \text{s.d.} \quad & c^t u = -1 \\ & -Au \leqslant 0 \end{aligned}$$

Insbesondere existiert eine minimale Fläche $F = \{u \in \mathbb{R}^n | A'u = b'\}$ des zugehörigen Polyeders $\{u \in \mathbb{R}^n | c^t u = -1, -Au \leqslant 0\}$, die nur aus Optimallösungen besteht, wobei $A'u \leqslant b'$ ein Teilsystem von $c^t u \leqslant -1, -c^t u \leqslant 1, -Au \leqslant 0$ ist (Theorem 44).

Es gilt $n - \operatorname{rank} A' = \dim(F) = n - \operatorname{rank} \begin{pmatrix} A \\ c^t \\ -c^t \end{pmatrix} = n - t$ (Corollary 45). A' enthält daher t linear unabhängige Zeilen. Mindestens $t-1$ dieser Zeilen gehören zu A. Somit erfüllt jeder Vektor aus F (b). $\square$

4.5 Kegel

Theorem 52 (Farkas-Minkowski-Weyl)**.** Ein Kegel ist genau dann polyedrisch, wenn er endlich erzeugt ist.

Beweis. " $\Longleftarrow$ " Seien $a_1, \dots, a_m \in \mathbb{R}^n$. Zu zeigen ist, dass $\text{cone}(\{a_1, \dots, a_m\})$ polyedrisch ist.

Wir können annehmen, dass $a_1, \dots, a_m$ den $\mathbb{R}^n$ aufspannen (anderenfalls betrachten wir einen Unterraum).

Sei $\mathcal{H}$ die Menge aller Halbräume $H_u = \{x \in R^n | u^t x \leqslant 0\}$, so dass für jedes $H_u \in \mathcal{H}$ folgende Bedingungen gelten:

- $\{a_1, \dots, a_m\} \subseteq H_u$
- Es gibt $n-1$ linear unabhängige Vektoren $a_{i_1}, \dots, a_{i_{n-1}}$ in $\{a_1, \dots, a_m\}$ mit $u^t a_{i_j} = 0$ für $j \in \{1, \dots, n-1\}$

Nach Theorem 51 ist $\text{cone}(\{a_1, \dots, a_m\})$ der Schnitt der Halbräume in $\mathcal{H}$. Insbesondere ist der Kegel als Schnitt endlich vieler Halbräume ein Polyeder.

" $\Longrightarrow$ " Sei $C = \{x \in \mathbb{R}^n | Ax \leqslant 0\}$ ein polyedrischer Kegel. Zu zeigen ist, dass C endlich erzeugt ist. Seien $a_1^t, \dots, a_m^t$ die Zeilen von A. Sei $C_A := \text{cone}(\{a_1, \dots, a_m\})$. Aus " $\Longleftarrow$ " folgt, dass C_A ein polyedrischer Kegel ist. Insbesondere existieren $d_1, \dots, d_k \in \mathbb{R}^n$, so dass $C_A = \{x \in \mathbb{R}^n | d_1^t x \leqslant 0, \dots, d_k^t x \leqslant 0\}$. Sei $C_D := \text{cone}(\{d_1, \dots, d_k\})$.

Behauptung 1. $C = C_D$

Unterbeweis. $C_D \subseteq C$: Es gilt $\{a_1, \dots, a_n\} \subseteq C_A$. Folglich ist $Ad_i \leqslant 0$ für $i \in \{1, \dots, k\}$. Daher liegen alle $d_i \in C$ und somit $C_D \subseteq C$.

$C \subseteq C_D$: Angenommen es existiert ein $c \in C \backslash C_D$. C_D ist endlich erzeugt, also nach dem bereits gezeigten Teil polyedrisch. Daher existiert $u \in \mathbb{R}^n$ mit $u^t d_i \leqslant 0$ ($i \in \{1, \dots, k\}$) und $u^t y > 0$. Es ist $u \in C_A$ und daher $u^t x \leqslant 0$ für alle $x \in C$. Dies ist allerdings ein Widerspurch zu $u^t y > 0$ und $y \in C$. ■

□

Remark 52.1. Für $S \subseteq \mathbb{R}^n$ heißt $S^\circ := \{x \in \mathbb{R}^n | x^t y \leqslant 0 \forall y \in S\}$ **Polarkegel** von S.

Wenn $S = \{x \in \mathbb{R}^n | Ax \leqslant 0\}$ ein polyedrischer Kegel ist, dann wird S° von den Zeilen von A erzeugt. (Übung)

Wir haben also gerade für polyedrische Kegel S gezeigt, dass $S^{\circ\circ} = S$ ist.

4.6 Polytope

Theorem 53. Eine Menge $X \subseteq \mathbb{R}^n$ ist genau dann ein Polytop, wenn sie die konvexe Hülle einer endlichen Menge von Vektoren in $\mathbb{R}^n$ ist.

Beweis. "$\implies$" Sei $X = \{x \in \mathbb{R}^n | Ax \leqslant b\} \neq \emptyset$ ein Polytop.

Schreibe X als $X = \{x \in \mathbb{R}^n | \begin{pmatrix} x \\ 1 \end{pmatrix} \in C\}$, wobei

$$C := \left\{ \begin{pmatrix} x \\ \lambda \end{pmatrix} \in \mathbb{R}^{n+1} | \lambda \geqslant 0, \forall x - \lambda b \leqslant 0 \right\}.$$

C ist ein polyedrischer Kegel.[2] Somit existieren $\begin{pmatrix} x_1 \\ \lambda_1 \end{pmatrix}, \ldots, \begin{pmatrix} x_k \\ \lambda_k \end{pmatrix} \in \mathbb{R}^{n+1}$ so dass $C = \operatorname{cone}\left(\left\{\begin{pmatrix} x_1 \\ \lambda_1 \end{pmatrix}, \ldots, \begin{pmatrix} x_k \\ \lambda_k \end{pmatrix}\right\}\right)$. X ist beschränkt, folglich kann c keinen Vektor $\begin{pmatrix} x \\ \lambda \end{pmatrix}$ mit $x \neq 0$ und $\lambda = 0$ enthalten. O.B.d.A. können wir daher $\lambda_i > 0$ für alle i annehmen. Nach Skalierung somit o.B.d.A. $\lambda_i = 1$.

Dann gilt

$$x \in X \iff \exists \mu_1, \ldots, \mu_k \geqslant 0 | \begin{pmatrix} x \\ 1 \end{pmatrix} = \mu_1 \begin{pmatrix} x_1 \\ 1 \end{pmatrix} + \ldots + \mu_k \begin{pmatrix} x_k \\ 1 \end{pmatrix}$$

und somit $X = \operatorname{conv}(\{x_1, \ldots, x_k\})$.

"$\impliedby$" Sei $X = \operatorname{conv}(\{x_1, \ldots, x_k\})$. Zu zeigen ist, dass X ein Polytop ist.

Sei $C := \operatorname{cone}\left(\left\{\begin{pmatrix} x_1 \\ 1 \end{pmatrix}, \ldots, \begin{pmatrix} x_k \\ 1 \end{pmatrix}\right\}\right)$. Es gilt $x \in X \iff \begin{pmatrix} x \\ 1 \end{pmatrix} \in C$. C ist endlich erzeugt und daher polyedrisch.

C kann geschrieben werden als $C = \left\{\begin{pmatrix} x \\ \lambda \end{pmatrix} | Ax + b\lambda \leqslant 0\right\}$. Insbesondere ist $X = \{x \in \mathbb{R}^n | Ax \leqslant -b\}$ ein Polyeder. Sei $M := \max\{\|x_i\| \, | i \in \{1, \ldots, k\}\}$. Dann gilt für alle $x \in X$: $x = \sum_{i=1}^k \lambda_i x_i$ für geeingete $\lambda_1, \ldots, \lambda_k \geqslant 0$ mit $\sum_{i=1}^k \lambda_i = 1$. Folglich ist $\|x\| \leqslant \sum_{i=1}^k \lambda_i \|x_i\| \leqslant M \sum_{i=1}^k \lambda_i = M$ und daher X beschränkt. $\square$

Corollary 54. Ein Polytop ist die konvexe Hülle seiner Ecken.

Beweis. Sei P ein Polytop mit Eckenmenge X. Da P konvex und $X \subseteq P$ ist, gilt $\operatorname{conv}(X) \subseteq P$. Zu zeigen ist noch die andere Inklusion. Wir wissen, dass $\operatorname{conv}(X)$ ein Polyeder ist. Angenommen es gibt $y \in P \setminus \operatorname{conv}(X)$. Dann gibt es $c \in R^n \setminus \{0\}$, sodass für $H = \{x \in \mathbb{R}^n | c^t x \leqslant \delta\}$ gilt $\operatorname{conv}(X) \subseteq H, y \notin H$. Dann gilt $c^t y > c^t x$ für alle $x \in X$. Allerdings gibt es stets eine Ecke, an der das Maximum $\max\{c^t x | x \in P\}$ angenommen wird $\lightning$. $\square$

[2] Dieser Trick ist auch für andere Beweise nützlich.

4.7 Zerlegung von Polyedern

Notation 54.1. Für zwei Vektormengen $\emptyset \neq X, Y \subseteq \mathbb{R}^n$ ist die **Minkowski-Summe** definiert als

$$X + Y := \{z \in \mathbb{R}^n | \exists x \in X, y \in Y : z = x + y\}$$

Ferner definieren wir $X + \emptyset = \emptyset + X = X$.

Theorem 55. Es sei $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ ein Polyeder. Dann gibt es endliche Mengen $V, E \subseteq \mathbb{R}^n$, sodass

$$P = \operatorname{conv}(V) + \operatorname{cone}(E)$$

Beweis. Der Kegel

$$C := \left\{ \begin{pmatrix} x \\ \lambda \end{pmatrix} | x \in \mathbb{R}^n, \lambda \in \mathbb{R}, \lambda \geqslant 0, Ax - \lambda b \leqslant 0 \right\}$$

ist polyedrisch, wird also von endlich vielen Vektoren $\begin{pmatrix} x_1 \\ \lambda_1 \end{pmatrix}, \dots, \begin{pmatrix} x_k \\ \lambda_k \end{pmatrix}$ erzeugt.

$x \in P$ gilt genau dann, wenn $\begin{pmatrix} x \\ 1 \end{pmatrix} \in C$, was genau dann der Fall ist, wenn $\begin{pmatrix} x \\ 1 \end{pmatrix} \in \operatorname{cone}\left(\{ \begin{pmatrix} x_1 \\ \lambda_1 \end{pmatrix} \}, \dots, \begin{pmatrix} x_k \\ \lambda_k \end{pmatrix} \right)$. Es gilt $\lambda_i \geqslant 0$ und nach Skalierung können wir $\lambda_i \in \{0, 1\}$ annehmen.

Setze $V := \{x_i | i \in \{1, \dots, k\}, \lambda_i = 1\}$ und $E := \{x_i | i \in \{1, \dots, k\}, \lambda_i = 0\}$. Dann ist $P = \operatorname{conv}(V) + \operatorname{cone}(E)$.

□

5 Der Simplex-Algorithmus

Der Simplex-Algorithmus war das erste Verfahren zur Lösung allgemeiner linearer Programme. Er wurde 1951 von G. Dantzig entwickelt. Eine polynomielle Laufzeit konnte nicht nachgewiesen werden, aber in der Praxis ist er oft sehr schnell.

Idea.

- *Starte in einer Ecke des Lösungspolyeders*
- *Solange es noch Nachbarecken mit besserem Zielfunktionswert gibt, laufe zu einer solchen.*

Kanonisches Bild von Wikipedia kopieren

Voraussetzungen

- Das Polyeder muss spitz sein

- Wir müssen eine Startlösung finden [3]

Um ein spitzes Lösungspolyeder zu haben, betrachten wir LPs in Standard-Gleichungsform. Um das Finden einer Startlösung werden wir uns später kümmern.

Wir wollen das folgende LP lösen ($A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$):

$$\begin{aligned} \max\, & c^t x \\ \text{s.d. } & Ax = b \\ & x \geqslant 0 \end{aligned}$$

Dabei können wir annehmen:

- $\operatorname{rank}(A) = m$ (weitere Zeilen sind entweder redundant oder liefern Widersprüche)
- $Ax = b$ ist lösbar
- $m < n$ (sonst ist die Lösung eindeutig)

Notation 55.1 (zur Vermeidung von Doppelindices)**.**

- Indexmenge der Spalten einer Matrix $A \in \mathbb{R}^{m \times n}$ mit $\{1, \dots, n\}$
- Für $B \subseteq \{1, \dots, n\}$ bezeichne A_B die Teilmatrix von A aus den Spalten mit Index in B
- Ebenso bezeichnen wir für einen Vektor $x \in \mathbb{R}^n$ mit x_B den Teilvektor x aus Einträgen mit Index in B. x_B ist dann ein Vektor der Länge $|B|$, dessen Einträge allerdings nicht mit $\{1, \dots, |B|\}$, sondern mit B indiziert sind.

Example 56.

$$A = \begin{pmatrix} 2 & 4 & 0 & -3 & 1 \\ 1 & 3 & 2 & 1 & 5 \end{pmatrix}, x = \begin{pmatrix} 3 \\ -2 \\ 0 \\ 1 \\ -1 \end{pmatrix}$$

$$A_{\{1,3,4\}} = \begin{pmatrix} 2 & 0 & -3 \\ 1 & 2 & 1 \end{pmatrix}, x_{\{1,3,4\}} = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$$

[3] Wir haben bereits gesehen, dass dies genau so schwer ist, wie das Lösen eines LPs (starke Dualität). Der Algorithmus ist trotzdem hilfreich, wie wir sehen werden.

Die Spalten von A_B und x_B sind mit $1, 3, 4$ indiziert, also

$$A_B x_B = x_1 a_1 + x_3 a_3 + x_4 a_4$$

Definition 57 (Zulässige Basislösungen)**.** Es sei $A \in \mathbb{R}^{m \times n}$ eine Matrix mit Rang m und $b \in \mathbb{R}^m$ ein Vektor. Sei $B \subseteq \{1, \dots, n\}, |B| = m$, sodass A_B regulär ist. Setze $N := \{1, \dots, n\} \setminus B$.[a]

(a) Wir nennen B eine **Basis** von A. Der Vektor x mit $x_B = A_B^{-1} b$ und $x_N = 0$ heißt **Basislösung von $Ax = b$ zur Basis B**.

(b) Wenn x eine Basislösung von $Ax = b$ zur Basis B ist, dann heißen die Variablen x_j mit $j \in B$ **Basisvariablen** und die Variablen x_j mit $j \in N$ heißen **Nicht-Basisvariablen**.

(c) Eine Basislösung x heißt **zulässig** (feasible), wenn $x \geqslant 0$. Eine Basis heißt **zulässig**, wenn die zugehörige Basislösung zulässig ist.

(d) Eine zulässige Basislösung x für eine Basis B heißt **nicht-degeneriert**, falls $A_B^{-1} b > 0$. Andernfalls heißt sie **degeneriert**.

[a] Vorstellung: $A = [A_B | A_N]$; die Spalten können dabei natürlich anders permutiert sein.)

Example 58.

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Für $B = \{1, 2\}$ ist $A_B = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ mit Basislösung $(1, 0, 0, 0)$, die zulässig und degeneriert ist.

- Für $B = \{1, 3\}$ und $B = \{1, 4\}$ ergeben sich die selben Basislösungen.
- Für $B = \{2, 3\}$ ist $A_B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ mit nicht zulässiger Basislösung $(0, 2, -1, 0)$.
- Für $B = \{2, 4\}$ ergibt sich $A_B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ mit zulässiger nicht-degenerierter Basislösung $(0, 1, 0, 1)$.

Um Beispiele im 2 und 3-dimensionalen zeichnen zu können, wollen wir diese Begriff auf Systeme der Form $\tilde{A}x \leqslant b, x \geqslant 0$ übertragen.

Definition 59. Es sei $\tilde{A} \in \mathbb{R}^{m \times \tilde{n}}$. Wir nennen einen Vektor $x^* \in \mathbb{R}^{\tilde{n}}$ mit

$\tilde{A}x^* \leqslant b$ mit $x^* \geqslant 0$ eine (zulässsige/degenerierte) Basislösung, falls x^*, s^* mit $s^* := b - \tilde{A}x^*$ eine solche von $\tilde{A}x + I_m s = b, x \geqslant 0, s \geqslant 0$ (mit $n := \tilde{n} + m$ Variablen). In einer zulässigen Basislösung von $\tilde{A}x \leqslant b, x \geqslant 0$ muss die Zahl an Nebenbedingungen, die mit Gleichheit erfüllt sind (inklusive Nicht-Negativitäts-Nebenbedingungen) folglich mindestens $n - m = \tilde{n}$ sein. In einer nicht-degenerierten Basislösung muss die Zahl der mit Gleichheit erfüllten Nebenbedingungen genau $\tilde{n}$ sein.

Remark 59.1. Ob eine Basislösung degeneriert ist, hängt zunächst einmal davon ab, welche Ungleichungen wir gewählt haben. In Dimensionen $\geqslant 3$ existieren jedoch Basislösungen, die unabhängig von der Wahl der Ungleichungen stets degeneriert sind.

Theorem 60. Sei $P = \{x \in \mathbb{R}^n | Ax = b, x \geqslant 0\}$ ein Polyeder mit $\operatorname{rank}(A) = m < n$. Dann ist der Vektor $x' \in P$ genau dann eine Ecke von P, wenn er eine zulässige Basislösung ist.

Beweis. Der Vektor x' ist genau dann eine Ecke von P, wenn er alle Ungleichungen des folgenden Systems erfüllt und davon n linear unabhängige mit Gleichheit:

$$\begin{aligned} Ax &\leqslant b \\ -Ax &\leqslant -b \\ -I_n x &\leqslant 0 \end{aligned}$$

Das ist genau dann der Fall, wenn $x' \geqslant 0$, $Ax' = b$ und $x'_N = 0$ für eine Menge $N \subseteq \{1, \ldots, n\}$ mit $|N| = n - m$, sodass mit $B = \{1, \ldots, n\} \backslash N$ die Matrix A_B vollen Rang hat. Das ist äquivalent dazu, eine zulässige Basislösung zu sein.

□

Example 61 (Simplex Algorithmus 1)**.** Wir betrachten

$$\begin{array}{lrrrrrl} \max & x_1 + x_2 & & & & & \\ \text{s.d.} & -x_1 & +x_2 & +x_3 & & & = 1 \\ & x_1 & & & +x_4 & & = 3 \\ & & x_2 & & & +x_5 & = 2 \\ & x_1, & x_2, & x_3, & x_4, & x_5 & \geqslant 0 \end{array}$$

Eine Startlösung ist in diesem Fall leicht ablesbar als $B = \{3, 4, 5\}$.

Wir stellen ein **Simplex-Tableau** auf (Gleichungssystem aufgelöst nach den Basisvariablen, und Zielfunktion aufgelöst als Funktion der Nicht-

Basisvariablen)

$$\begin{array}{rcl} x_3 & = & 1 + x_1 - x_2 \\ x_4 & = & 3 - x_1 \\ x_5 & = & 2 - x_2 \\ \hline z & = & x_1 + x_2 \end{array} \tag{1}$$

(Zu einem Simplex-Tableau denken wir uns immer genau die Lösung, für die alle Nicht-Basisvariablen 0 sind).

Wir wollen nun genau eine der Nicht-Basis Variablen, welche in der Zielfunktion einen positiven Koeffizienten hat, erhöhen. Wir wählen dafür x_2. Aus jeder Gleichung können wir ablesen, um wie viel wir x_2 erhöhen können: 1. Nebenbedingung: $x_2 \leqslant 1$, 2. Nebenbedingung: Keine Einschränkung an x_2, 3. Nebenbedingung: $x_2 \leqslant 2$.

Wir erhöhen x_2 also auf 2 und machen x_2 hiermit zu einer Basisvariable. x_3 wird hierdurch aus der Basis entfernt.

Wir erhalten folgendes neues Tableau:

$$\begin{array}{rcl} x_2 & = & 1 + x_1 - x_3 \\ x_4 & = & 3 - x_1 \\ x_5 & = & 1 - x_1 + x_3 \\ \hline z & = & 1 + 2x_1 - x_3 \end{array} \tag{2}$$

Nun können wir nur noch x_1 wählen: Es ist $x_4 = 1 - x_1 + x_3$ kritisch. Wir ersetzen 5 in B durch 1. Es ist $x_1 = 1 + x_3 - x_5$.

$$\begin{array}{rcl} x_1 & = & 1 + x_3 - x_5 \\ x_2 & = & 2 \\ x_4 & = & 2 - x_3 + x_5 \\ \hline z & = & 3 + x_3 - 2x_5 \end{array} \tag{3}$$

$$\begin{array}{rcl} x_1 & = & 1 + x_3 - x_5 \\ x_2 & = & 2 \\ x_4 & = & 2 - x_3 + x_5 \\ \hline z & = & 3 + x_3 - 2x_5 \end{array} \tag{4}$$

Unsere aktuelle Lösung ist $x = (1, 2, 0, 2, 0)$.

Der einzige Kandidat für eine Ersetzung ist x_3

Es ergibt sich

$$
\begin{array}{rcl}
x_1 & = & 3 + x_4 \\
x_2 & = & 2 - x_5 \\
x_3 & = & 2 - x_4 + x_5 \\
\hline
z & = & 5 - x_4 - x_5
\end{array}
\tag{5}
$$

Die zugehörige Lösung ist $x = (3, 2, 2, 0, 0)$. Dies ist eine Optimallösung.

Example 62 (Simplex-Algorithmus 2, unbeschränkt)**.**

$$
\begin{aligned}
\max x_1 & \\
\text{s.d.} x_1 - x_2 + x_3 &= 1 \\
-x_1 + x_2 + x_4 &= 2 \\
x_1, x_2, x_3, x_4 &\geqslant 0
\end{aligned}
$$

Startbasis: $B = \{3, 4\}$

Simplex Tableau

$$
\begin{array}{rcl}
x_3 & = & 2 - x_1 + x_2 \\
x_4 & = & 2 + x_1 - x_2 \\
\hline
z & = & x_1
\end{array}
\tag{6}
$$

Wir können nur x_1 erhöhen. $x_3 = 1 - x_1 + x_2$ ist kritisch.

$$
\begin{array}{rcl}
x_1 & = & 1 + x_2 - x_3 \\
x_4 & = & 3 - x_3 \\
\hline
z & = & 1 + x_2 - x_3
\end{array}
\tag{7}
$$

Der einzige Kandidat zur Erhöhung ist x_2. Für x_2 gibt es keine obere Schranke. Das LP ist daher unbeschränkt.

Example 63 (Simplex-Algorithmus 3, Degeneriertheit)**.** Bei degenerierten Lösungen gehört die selbe Lösung zu mehreren Basen. Dies ist ein Problem, da der Simplex-Algorithmus Basen (nicht Basis-Lösungen) durchläuft.

$$
\begin{aligned}
&\max x_2 \\
&\quad \text{s.d.} -x_1 + x_2 + x_3 = 0 \\
&\quad\quad x_1 + x_4 \qquad\quad = 2 \\
&\quad\quad x_1, x_2, x_3, x_4 \ \geqslant 0
\end{aligned}
$$

Initiale Basis $B = \{3, 4\}$.

Simplex-Tableau:

$$
\begin{array}{rcl}
x_3 & = & x_1 - x_2 \\
x_4 & = & 2 - x_1 \\
\hline
z & = & x_2
\end{array}
\tag{8}
$$

Wir wollen x_2 erhöhen. $x_3 = x_1 - x_2$ ist kritisch. Wir ersetzen daher in B die 3 durch 2.

x_2 kann nicht erhöht werden! Es ergibt sich

$$
\begin{array}{rcl}
x_2 & = & x_1 - x_3 \\
x_4 & = & 2 - x_1 \\
\hline
z & = & x_1 - x_3
\end{array}
\tag{9}
$$

Wir assoziieren mit dem neuen Tableau die selbe Lösung.

Wir können jetzt allerdings x_1 in die Basis aufnehmen. $x_4 = 2 - x_1$ ist kritisch.

$$
\begin{array}{rcl}
x_1 & = & 1 + x_4 \\
x_2 & = & 2 - x_3 - x_4 \\
\hline
z & = & 2 - x_3 - x_4
\end{array}
\tag{10}
$$

Wir erhalten eine Optimallösung $x = (2, 2, 0, 0)$.

Warning. *Bei degenerierten Lösungen müssen wir aufpassen, dass wir nicht im Kreis laufen.*

Definition 64 (Simplex-Tableau)**.** Für eine zulässige Basis B ist das **Simplex-Tableau** ein System $T(B)$ von $m + 1$ linearen Gleichungen mit Variablen

$x_1, \ldots, x_n$ und z der F

$$\begin{array}{ccc} x_B & = & p + Qx_N \\ \hline z & = & z_0 + r^t x_N \end{array} \tag{11}$$

Dabei ist x_B der Vektor der Basisvariablen, $N := \{1, \ldots, n\} \backslash B$ und x_N ist der Vektor der Nicht-Basisvariablen. p ist ein Vektor der Länge m, Q eine $m \times (n-m)$-Matrix, r ein Vektor der Länge $n-m$ und $z_0 \in \mathbb{R}$.

$T(B)$ soll dabei die gleiche Lösungsmenge wie das System $Ax = b, z = c^t x$ haben.

Man beachte, dass p durch B indiziert ist (und entsprechend für r). Insbesondere sind die Zeilen von Q durch B indiziert und die Spalten durch N. Bezeichne die Einträge von Q mit q_{ij} für $i \in B, j \in N$.

Wichtige Eigenschaften eines Simplex-Tableau

- $x \in \mathbb{R}^n$ ist genau dann eine Lösung von $Ax = b$, wenn $x_B = p + Qx_N$
- Für jede Lösung $x \in \mathbb{R}^n$ von $Ax = b$ gilt $c^t x = z_0 + r^t x_N$

Lemma 65. Für jede zulässige Basis B gibt es ein Simplex-Tableau $T(B)$

Beweis. Setze

- $p = A_B^{-1} b$
- $Q = -A_B^{-1} A_N$
- $r = c_N c_N - (c_B^t A_B^{-1} A_N)^t$
- $z_0 = c_B^t A_B^{-1} b$

Dann gilt

$$\begin{aligned} & x_B = A_B^{-1} b - A_B^{-1} A_N x_n \\ \iff\ & A_B x_B = b - A_N x_N \\ \iff\ & Ax = b \end{aligned}$$

Außerdem

$$\begin{aligned} z &= c_B^t A_B^{-1} b + (c_N^t - (c_B^t A_B^{-1} A_N)) x_N \\ &= c_B^t A_B^{-1} (b - A_N x_N) + c_N^t x_N \\ &= c_B^t A_B^{-1} A_B x_B + c_N^t x_N \\ &= c_B^t x_B + c_N^t x_N \\ &= c^t x \end{aligned}$$

$\square$

Notation 65.1. r heißt auch Vektor der **reduzierten Kosten**.

Lemma 66 (Optimalitätskriterium)**.** Sei $T(B)$:

$$\begin{array}{ccll} x_B & = & p & +Qx_N \\ \hline z & = & z_0 & +r^t x_N \end{array}$$

ein Simplextableau für eine zulässige Basis B. Falls $r \leqslant 0$, dann ist die Basislösung zu B optimal.

Beweis. Sei x die Basislösung von B. Weben $x_N = 0$ gilt $c^t x = z_0 (= c_B^t A_B^{-1} b)$. Wenn x^* irgendeine zulässige Lösung mit Wert $z^* = c^t x^*$ ist, dann bilden x^* und z^* auch eine Lösung von $T(B)$ und es gilt (wegen $r \leqslant 0$ und $x_N^* \geqslant 0$):

$$z^* = z_0 + r^t x_N^* \leqslant z_0 = c^t x$$

$\square$

Lemma 67 (Unbeschränkte Instanzen)**.** Sei $T(B)$:

$$\begin{array}{ccll} x_B & = & p & +Qx_N \\ \hline z & = & z_0 & +r^t x_N \end{array}$$

ein Simplextableau für eine zulässige Basis B. Wenn es ein $\alpha \in N$ mit $r_\alpha > 0$ gibt, sodass die Spalte von Q mit Index α nur nicht-negative Einträge enthält, dann ist das zugehörige LP unbeschränkt.

Beweis. Sei x die zulässige Basislösung für B. Sei $K \in \mathbb{R}$ mit $K > c^t x$ eine Konstante. Definiere wie folgt eine neue zulässige Lösung $\tilde{x}$:

$$\tilde{x}_i := \begin{cases} \frac{K - c^t x}{r_\alpha} & \text{für } i = \alpha \\ x_i = 0 & \text{für } i \in N \backslash \{\alpha\} \\ p_i + q_{i_\alpha} \tilde{x}_\alpha & \text{für } i \in B \end{cases}$$

Dann ist $\tilde{x}$ eine zulässige Lösung mit $c^t \tilde{x} \geqslant K$ und das LP somit unbeschränkt.
$\square$

Lemma 68 (Austauschschritt)**.** Sei $T(B)$:

$$\begin{array}{ccll} x_B & = & p & +Qx_N \\ \hline z & = & z_0 & +r^t x_N \end{array}$$

ein Simplextableau für eine zulässige Basis B. Sei $\alpha \in N$ ein Index mit

$r_\alpha > 0$ und $\beta \in B$ mit $q_{\beta\alpha} < 0$ und $\frac{p_\beta}{q_{\beta\alpha}} = \max\{\frac{p_i}{q_{i_\alpha}} | q_{i_\alpha} < 0, i \in B\}$. Dann ist $\tilde{B} = (B \cup \{\alpha\}) \backslash \{\beta\}$ eine zulässige Basis.

Beweis.

- $A_{\tilde{B}}$ ist regulär:

  Wir zeigen, dass $A_B^{-1} A_{\tilde{B}}$ regulär ist, daraus folgt die zu zeigende Aussage. Bis auf eine Spalte stimmen die Spalten von A_B und $A_{\tilde{B}}$ überein. Insbesondere entsteht $A_B^{-1} A_{\tilde{B}}$ aus der $m \times m$-Einheitsmatrix, indem eine Spalte ersetzt wird:

$$A_B^{-1} A_{\tilde{B}} = \begin{pmatrix} 1 & & & \vdots & & & \\ & \ddots & & \vdots & & & \\ & & 1 & \vdots & & & \\ & & & \vdots & & & \\ & & & \vdots & 1 & & \\ & & & \vdots & & \ddots & \\ & & & \vdots & & & 1 \end{pmatrix}$$

  Die einzige Spalte von $A_B^{-1} A_{\tilde{B}}$, die kein Einheitsvektor ist, ist $A_B^{-1} a_{\cdot\alpha}$, wobei $a_{\cdot\alpha}$ die Spalte von A mit Index α sei. Das ist die Spalte mit Index α von $-Q = A_B^{-1} A_N$. Es gilt $q_{\beta\alpha} \neq 0$, daher hat die Matrix $A_B^{-1} A_{\tilde{B}}$ an der Stelle mit Zeilenindex β und Spaltenindex α einen von 0 verschiedenen Eintrag.

- Die zugehörige Basislösung ist nichtnegativ:

  Wir erhöhen x_α auf $-\frac{p_\beta}{q_{\beta\alpha}}$ und setzen die Basisvariablen x_β auf $p - q_{\cdot\alpha} \frac{p_\beta}{q_{\beta\alpha}}$, wobei $q_{\cdot\alpha}$ die Spalte von Q mit Index α sei. Für $i \in B$ mit $q_{i\alpha} \geqslant 0$ (insbesondere $i \neq \beta$) gilt:

$$p_i - q_{i\alpha} \frac{p_\beta}{q_{\beta\alpha}} \geqslant p_i \geqslant 0$$

  Für $i \in B$ mit $q_{i\alpha} < 0$ gilt:

$$\frac{p_\beta}{q_{\beta\alpha}} \geqslant \frac{p_i}{q_{i\alpha}} \implies p_i \geqslant q_{i\alpha} \frac{p_\beta}{q_{\beta\alpha}}$$

  mit Gleichheit bei $i = \beta$.

  Folglich ist $x_\beta = 0$ und $x_{\tilde{B}} \geqslant 0$. Dies ist somit eine zulässige Basislösung zu $\tilde{B}$.

□

Algorithmus 1 : Simplex Algorithmus

Data : $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$
Result : $\tilde{x} \in \{x \in \mathbb{R}^n | Ax = b, x \geqslant 0\}$ maximizing $c^t x$ or `infeasible`, `unbounded`

Berechne eine zulässige Basis B.
Falls keine solche existiert **return** `infeasible`
while *true* **do**

    $N := \{1, \dots, n\} \backslash B$ und berechne eine zulässige Basislösung x für B
    Berechne das Simplex-Tableau

$$\begin{array}{rcll} x_B & = & p & +Qx_N \\ \hline z & = & z_0 & +r^t x_N \end{array}$$

    für B.
    if $r \leqslant 0$ **then**
        return $\tilde{x} = x$
    Wähle $\alpha \in N$ mit $r_\alpha > 0$.
    if $q_{i_\alpha} \geqslant 0$ *für alle* $i \in B$ **then**
        return `unbounded`
    Wähle $\beta \in B$ mit $q_{\beta\alpha} < 0$ und $\frac{p_\beta}{q_{\beta\alpha}} = \max\{\frac{p_i}{q_{i_\alpha}} | q_{i_\alpha} < 0, i \in B\}$.
    $B := (B \backslash \{\beta\}) \cup \{\alpha\}$

Question 68.1. Wie finden wir eine Startlösung?

Das Finden einer Startlösung ist genau so schwer wie das Lösen eines LPs. Wir können hierzu allerdings den Simplex-Algorithmus zunächst auf ein LP anwenden, für das wir eine Startlösung kennen. Sei o.B.d.A. $b \geqslant 0$. Setze $\tilde{A} := (A|I_m)$ und füge neue Variablen $x_{n+1}, \dots, x_{n+m}$ hinzu und löse mit $\tilde{x} = (x_1, \dots, x_{n+m})$ das folgende LP ($\tilde{P}$):

$$\begin{aligned} \max\ & -(x_{n+1} + x_{n+2} + \ldots + x_{n+m}) \\ \text{s.t. } & \tilde{A}\tilde{x} = b \\ & \tilde{x} \geqslant 0 \end{aligned}$$

Für ($\tilde{P}$) ist $\{n+1, \dots, n+m\}$ eine zulässige Basis. Das LP kann also mit dem Simplex-Algorithmus gelöst werden. Falls der Lösungswert von ($\tilde{P}$) negativ ist, hat das ursprüngliche LP keine zulässige Lösung. Anderenfalls liefert der Simplex-Algorithmus eine Basislösung. Da für diese $x_{n+1} = x_{n+2} = \ldots = x_{n+m} = 0$ gilt, kann man daraus leicht eine zulässige Basislösung des ursprünglichen LPs konstruieren.

5.1 Pivotregeln

Question 68.2. Wie wählen wir α und β?

Regeln, die dies festlegen, heißen **Pivotregeln**.

Example 69 (Pivotregeln).

- **Largest coefficient rule**: Wähle α so, dass r_α maximiert wird. (ursprüngliche Wahl von Dantzig)
- **Largest increase rule**: Wähle alle Variablen so, dass der Anstieg der Zielfunktion maximiert wird. (Aufwendig zu überprüfen).
- **Steepest edge rule**: Wähle α so, dass

  $$\frac{c^t(x_{\text{new}-x_{\text{old}}})}{\|x_{\text{new}-x_{\text{old}}}\|}$$

  maximiert wird. (Aufwändig, aber gut in der Praxis)

Für alle dieser Regeln ist bekannt, dass es Instanzen gibt, auf denen die Laufzeit nicht polynomiell ist.

5.1.1 Terminierung

Goal. *Finde Pivotregeln, die erzwingen, dass der Simplex-Algorithmus terminiert.*

Wenn der Algorithmus nicht terminiert, betrachtet er eine Basis B zweimal (und damit auch unendlich oft). Diese Verhalten heißt **Kreiseln**. Die Berechnung zwischen zwei Vorkommen von B nennen wir **Kreis**. Seien $F \subseteq \{1, \ldots, n\}$ die Indizes der Variablen, die während eines Kreises zur Basis hinzugefügt (und daher auch wieder entfernt werden). Wir nennen x_F **Kreiselvariablen**.

Lemma 70. Wenn der Simplex-Algorithmus kreiselt, sind alle Basislösungen während des Kreiselns identisch und alle Kreisvariablen sind 0.

Beweis. Der Lösungswert wird während des Algorithmus nie kleiner, kann während des Kreiselns also auch nicht größer werden.

Sei B eine Basis während des Kreiselns und $(B \cup \{\alpha\}) \backslash \{\beta\}$ die nächste Basis.

Unter den Nicht-Basisvariablen könnte dabei höchstens x_α größer werden. Wegen $r_\alpha > 0$ würde das allerdings den Lösungswert erhöhen. Folglich bleiben alle Nicht-Basisvariablen 0. Da die Nicht-Basisvariablen die gesamte Lösung bestimmen bleiben alle Variablen unverändert. □

Definition 71 (Bland-Regel). Wählen α unter allen Elementen von N mit $r_\alpha > 0$ so, dass α minimal ist. Wähle β ebenfalls minimal.

Remark 71.1. Das ist keine sonderlich intelligente Wahl von α und β, genügt aber um eine Terminierung zu erzwingen.

Theorem 72. Mit der Bland-Regel als Pivotregel terminiert der Simplex-Algorithmus nach endlich vielen Schritten.

Beweis. Angenommen der Algorithmus kreiselt mit der Bland-Regel. Sei F die Indexmenge der Kreiselvariablen. Sei $\pi := \max(F)$ und B die Basis unmittelbar, bevor π sie betritt.

Sei $T(B)$ wie folgt:

$$\begin{array}{rcll} x_B & = & p & +Qx_N \\ \hline z & = & z_0 & +r^t x_N \end{array}$$

Sei B' die Basis unmittelbar, bevor π sie verlässt und $T(B')$:

$$\begin{array}{rcll} x_{B'} & = & p' & +Q'x_{N'} \\ \hline z & = & z_0' & +r'^t x_{N'} \end{array}$$

Nach der Bland-Regel wählen wir sets den kleinsten möglichen Index. Wenn $\pi = \max(F)$ ausgesucht wird, ist π der einzige Kandidat in F, der B betreten kann. Also:

$$r_\pi > 0 \text{ und } r_j \leqslant 0 \text{ für } j \in N \cap (F \backslash \{\pi\})(1)$$

Sei α der Index, der B' betritt. Sei ist $\pi = \max(F)$ der einzige Kandidat aus F, der B' verlassen kann. Wegen $p_j' = 0$ für alle $j \in B' \cap F$ folgt:

$$q_{\pi\alpha}' < 0 \text{ und } q_{j\alpha}' \geqslant 0 \text{ für } j \in B' \cap (F \backslash \{\pi\})(2)$$

Betrachte das folgende Hilfs-LP ($\tilde{P}$):

$$\begin{aligned} \max\ & c^t x \\ \text{s.t. } & Ax = b \\ & x_{F \backslash \{\pi\}} \geqslant 0 \\ & x_\pi \leqslant 0 \\ & x_{N \backslash F} = 0 \end{aligned}$$

Beachte: Es werden keine Bedingungen an die Vorzeichen der Variablen in $x_{B \backslash F}$ gestellt.

Behauptung 1. *($\tilde{P}$) hat eine Optimallösung.*

Unterbeweis. Sei $\tilde{x}$ die zulässige Basislösung des ursprünglichen LPs zur Basis B. Es gilt $\tilde{x}_\pi = 0$, somit ist $\tilde{x}$ eine zulässige Lösung von $(\tilde{P}\ )$.

Die Kosten einer Lösung x von $Ax = b$ sind $c^t x = z_0 + r^t x_N$. Für jede Lösung x von $(\tilde{P}\ )$ gilt:

$$x_j \begin{cases} \geqslant 0 & \text{falls } j \in F \backslash \{\pi\} \\ \leqslant 0 & \text{falls} j = \pi \end{cases}$$

Somit ist $r_j x_j \leqslant 0$ für alle $j \in F$. Wegen $x_{N \backslash F} = 0$ folgt $r^t x_N \leqslant 0$ für jede Lösung x von $(\tilde{P})$. Der Wert jeder Lösung von $(\ \tilde{P}\ )$ ist daher höchstens z_0. Insbesondere ist $\tilde{x}$ eine Optimallösung von $(\tilde{P}\ )$. ■

Behauptung 2. *$(\tilde{P}\ )$ ist unbeschränkt.*

Unterbeweis. Wir haben beim Kreiseln immer dieselbe Basislösung. Eine Basislösung $\tilde{x}$ des ursprünglichen LP zur Basis B ist daher auch eine zulässige Basislösung zur Basis B'.

Wähle $K > 0$ und setze $x'_\alpha = K$. Für alle $j \in N' \backslash \{\alpha\}$ setze $x'_j = \tilde{x}_j = 0$. Ferner setzen wir $x_{B'} = p' + Q' x'_{N'}$. Dies definiert wegen (2) eine zulässige Lösung von $(\tilde{P})$.

Weil α die Basis B' betreten konnte, gilt $r'_\alpha > 0$. Folglich haben wir eine Lösung mit Wert $c^t x' = z'_0 + r'^t x'_{N'} = z'_0 + K \cdot r'_\alpha$.

K kann beliebig große gewählt werden, somit ist $(\tilde{P})$ unbeschränkt und die Behauptung folgt.

■

Dies liefert offenbar einen Widerspruch. □

Example 73 (Klee-Minty-Würfel)**.** Betrachte das folgende LP:

$$\begin{aligned} \max\ & x_n \\ \text{s.t.}\ & -x_1 \leqslant 0 \\ & x_1 \leqslant 1 \\ & \varepsilon x_{j-1} - x_j \leqslant 0 \quad \forall j \in \{2, \dots, n\} \\ & \varepsilon x_{j-1} + x_j \leqslant 1 \quad \forall j \in \{2, \dots, n\} \end{aligned}$$

Der Lösungsraum dieses LPs heißt **Klee-Minty-Würfel**. Mit der Bland-Regel betrachtet der Simplex-Algorithmus (beginndend mit der Ecke (0, ..., 0)) 2^n Basen.

Remark 73.1. Für jede bisher angegebenen Pivotregel konnten Beispiele gefunden haben, auf der der Algorithmus exponentielle Laufzeit hat.

Question 73.2. Kann man vielleicht allgemein zeigen, dass der Algorithmus nicht polynomiell sein kann?

Definition 74. Der **kombinatorische Durchmesser** eines spitzen Polyeders P ist der Durchmesser (größeter Abstand zwischen zwei Knoten) des ungerichteten Graphen G_P, wobei $V(G_P)$ die Menge der Ecken von P ist, und zwei Knoten genau dann durch eine Kante in G_P verbunden sind, wenn es eine Fläche der Dimension 1 gibt, die diese Ecken enthält.

Ohne Annahmen über die Startlösung ist der kombinatorische Durchmesser eine untere Schranke für die Laufzeit des Simplex-Algorithmus.

Remark 74.1. Es ist nicht bekannt, ob der kombinatorische Durchmesser polynomiell (oder sogar linear) beschränkt in der Eingabegröße ist.

Definition 75 (Revidierter Simplex-Algorithmus)**.** Die explizite Berechnung des Simplex-Tableaus ist recht zeitaufwendig und kann vermieden werden. Die $m \times (n-m)$-Matrix Q muss nicht vollständig gespeichert werden. Es reicht, die Spalte von Q für den Index $\alpha \in N$ mit $r_\alpha > 0$, den man auswählt, zu berechnen (**Spaltenerzeugung**), da $Q = -A_B^{-1} A_N$. Die Matrix A_B^{-1} muss nicht explizit bestimmt werden. Es reicht, Gleichungssysteme der Form $A_B y = d$ zu lösen. Dazu kann man z.B. eine LU-Zerlegung von A_B verwalten und geeignet aktualisieren.

Der Simplex-Algorithmus mit diesen Anpassungen heißt **revidierter Simplex-Algorithmus**.

Definition 76 (Dualer Simplex-Algorithmus)**.** Sei $T(B)$:

$$\begin{array}{rcll} x_{B'} & = & p' & +Q' x_{N'} \\ \hline z & = & z'_0 & +r'^t x_{N'} \end{array}$$

ein Simplex-Tableau. Während des (primalen) Simplex-Algorithmus gilt die Invariante $p \geqslant 0$, das ist äquivalent zur (primalen) Zulässigkeit. Sobald $r \leqslant 0$ gilt, ist die Lösung optimal. Dann ist $\tilde{y} = A_B^{-1\,t} c_B$ eine optimale Lösung des dualen LPs $\min\{b^t y \geqslant 0\}$. Dabei ist $r \leqslant 0$ äquivalent zur dualen Zulässigkeit.

Im **dualen Simplex-Algorithmus** bleibt die duale Zulässigkeit (d.h. $r \leqslant 0$) stets garantiert, während die primale Zulässigkeit ($p \geqslant 0$) erst am Ende erreicht wird.

Der duale Algorithmus hat dieselbe Effizienz wie der primale Algorithmus. Ein Vorteil ist, dass beim nachträglichen Hinzufügen von Nebenbedingungen die duale Lösung zulässig bleibt.

6 Der Netzwerk-Simplex-Algorithmus

Idea. *Wende den Simplex-Algorithmus an, um Min-Cost-Flow-Probleme zu lösen. Die Laufzeit ist wie im allgemeinen Fall: Es ist keine polynomielle Laufzeit beweisbar, in der Praxis ist der Algorithmus allerdings recht gut. Der Algorithmus kann auch als rein kombinatorischer Algorithmus angesehen werden.*

Definition 77. Sei G ein gerichteter Graph mit Kapazitäten $u : E(G) > \mathbb{R}_{>0}$ und Zahlen $b : V(G) \to \mathbb{R}$ mit $\sum_{v \in V(G)} b(v) = 0$.

Ein **zulässiger b-Fluss** in (G, u, b) ist eine Abbildung $f : E(G) \to \mathbb{R}_{\geqslant 0}$ mit

- $\forall e \in E(G) : f(e) \leqslant u(e)$
- $\forall v \in V(G) : \sum_{e \in \delta^+_G(v)} f(e) - \sum_{e \in \delta^-_G(v)} f(e) = b(v)$.

Notation 77.1. Es heißt $b(v)$ die **Balance** von v. Wenn $b(v) > 0$, nennen wir $b(v)$ **Angebot**, falls $b(v) < 0$ **Nachfrage**.

Knoten v von G mit $b(v) > 0$ heißen **Quellen**, Knoten mit $b(v) < 0$ heißen **Senken**.

Problem (Minimum-Cost-Flow-Problem)**.** *Gegeben sei ein gerichteter Graph G, Kapazitäten $u : E(G) \to \mathbb{R}_{>0}$, Balancen $b : B(G) \to \mathbb{R}$ mit $\sum_v b(v) = 0$ und Kantenkosten $c : E(G) \to \mathbb{R}$.*

Gesucht ist ein b-Fluss f, der $\sum_{e \in E(G)} c(e) \cdot f(e)$ minimiert.

Definition 78. Sei G ein gerichteter Graph.

- Für $e = (v, w)$ sei $\overleftarrow{e} = (w, v)$ die **Rückwärtskante**.
- Definiere $\overleftrightarrow{G}$ durch $V(\overleftrightarrow{G}) := V(G)$ und $E(\overleftrightarrow{G}) := E(G) \sqcup \{\overleftarrow{e} | e \in E(G)\}$
- Kantenkosten werden durch $c(\overleftarrow{e}) := -c(e)$ erweitert.
- Sei (G, u, b, c) eine Instanz von Minimum-Cost-Flow und f ein b-Fluss. Der **Residualgraph** $G_{u,f}$ ist definiert durch $V(G_{u,f}) := V(G)$ und $E(G_{u,f}) := \{e \in E(G) | f(e) < u(e)\} \sqcup \{\overleftarrow{e} \in E(\overleftrightarrow{G}) | f(e) > 0\}$.
- Für $e \in E(G)$ definieren wir die **Residualkapazität** durch $u_f(e) := u(e) - f(e)$ und $u_f(\overleftarrow{e}) := f(e)$.

Definition 79. Wenn P ein Teilgraph des Residualgraphs $G_{u,f}$ ist, dann heißt **Augmentieren** von f entlang P und $\gamma > 0$, dass man f auf den Vorwärtskanten in P um γ erhöht und auf den Rückwärtskanten in P um γ verringert.

Definition 80. Sei (G, u, b, c) eine Instanz des Min-Cost-Flow-Problems. Ein b-Fluss f in (G, u) heißt **Baumlösung**, wenn der Graph $(V(G), \{e \in E(G) | 0 < f(e) < u(e)\})$ keinen ungerichteten Kreis enthält.

Lemma 81. Sei (G, u, b, c) eine Instanz des Min-Cost-Flow-Problems. Ein b-Fluss f ist genau dann eine Baumlösung, wenn $\tilde{x} \in \mathbb{R}^{E(G)}$ mit $\tilde{x}_e = f(e)$ eine Ecke des Polytops

$$\left\{ x \in \mathbb{R}^{E(G)} \;\middle|\; \begin{array}{ll} \forall e \in E(G): & 0 \leqslant x_e \leqslant u(e), \\ \forall v \in V(G): & \sum_{e \in \delta^+(v)} x_e - \sum_{e \in \delta^-(v)} x_e = b(v) \end{array} \right\}$$

Beweis. „ $\Longrightarrow$ " Sei f eine Baumlösung und $\tilde{x} \in \mathbb{R}^{E(G)}$ mit $\tilde{x}_e = f(e)$ für alle $e \in E(G)$.

Für jede Kante $e \in E(G)$ mit $f(e) = 0$ betrachte die Bedingung $x_e \geqslant 0$. Für jede Kante $e \in E(G)$ mit $f(e) = u(e)$ betrachte die Bedingung $x_e \leqslant u(e)$. Für jede Zusammenhangskomponente von $(V(G), \{e \in E(G) | 0 < f(e) < u(e)\}$ betrachte für alle Knoten bis auf einen die Gleichung $\sum_{e \in \delta^+_G(v)} x_e - \sum_{e \in \delta^-_G(v)} x_e = b(v)$.

Dies sind $|E(G)|$ linear unabhängige Nebenbedingungen, die alle von $\tilde{x}$ mit Gleichheit erfüllt werden. Somit ist $\tilde{x}$ eine Ecke des Polyeders.

„ $\Longleftarrow$ " Sei f ein b-Fluss und sei $\tilde{x} \in \mathbb{R}^{E(G)}$ mit $\tilde{x}_e = f(e)$. Sei $\tilde{x}$ keine Baumlösung.

Dann enthält $(V(G), \{e \in E(G) | 0 < f(e) < u(e)\}$ einen ungerichteten Kreis C.

Wähle $\varepsilon > 0$ mit $\varepsilon \leqslant \min\{\min\{f(e), u(e) - f(e)\} | e \in E(G)\}$. Fixiere eine der beiden möglichen Orientierungen von C.

$x' \in \mathbb{R}^{E(G)}$ entstehen, indem $\tilde{x}$ entlang C in dieser Orientierung um ε augmentiert werde. $x'' \in \mathbb{R}^{E(G)}$ entstehe, indem $\tilde{x}$ entlang C in der umgekehrten Orientierung von ε augmentiert werde.

Dann ist $\tilde{x} = \frac{1}{2}(x' + x'')$, somit keine Ecke des Polyeders. □

Corollary 82. Sei (G, u, b, c) eine Instanz des Min-Cost-Flow-Problems. Wenn es einen b-Fluss (G, u) gibt, dann gibt es eine Optimallösung von (G, u, b, c), die eine Baumlösung ist.

Beweis. In spitzen Polyedern wird das Minimum einer linearen Funktion in einer Ecke angenommen. □

Definition 83. Sei (G, u, b, c) eine Minimum-Cost-Flow-Instanz mit zusammenhängendem G.

Eine **Spannbaum-Struktur** ist ein Quadrupel (r, T, L, U), wobei $r \in V(G)$, $E(G) = T \sqcup L \sqcup U$, $|T| = |V(G)| - 1$ und $(V(G), T)$ keinen ungerichteten Kreis enthalte.

Der zu (r, T, L, U) **assoziierte b-Fluss** f ist definiert durch

- $f(e) = 0$ für $e \in L$
- $f(e) = u(e)$ für $e \in U$
- $f(e) = \sum_{v \in C_e} b(v) + \sum_{e' \in U \cap \delta^-(C_e)} u(e') - \sum_{e' \in U \cap \delta^+(C_e)} u(e')$ für $e \in T$, wobei C_e die Knotenmenge der Zusammenhangskomponente von $(V(G), T \backslash \{e\})$ ist, die v enthalte (für $e = (v, w)$).

Die Struktur (r, T, L, U) heißt **zulässig**, wenn $0 \leqslant f(e) \leqslant u(e)$ für alle $e \in T$ gilt. Eine Kante $(v, w) \in T$ heißt **Abwärtskante**, wenn v auf dem ungerichteten $r - w$-Weg in T liegt. Sonst heißt sie **Aufwärtskante**.

Observe. *Der zu (r, T, L, U) assoziierte b-Fluss erfüllt immer die Flusserhaltungsgleichungen, ist aber nicht unbedingt zulässig.*

Definition 84. Eine zulässige Spannbaum-Struktur (r, T, L, U) heißt **stark zulässig**, wenn $0 < f(e)$ für jede Abwärtskante $e \in T$ und $f(e) < u(e)$ für jede Aufwärtskante $e \in T$.

Observe. *In einer stark zulässigen Spannbaum-Struktur kann stets in Richtung von r augmentiert werden.*

Definition 85. Sei (r, T, L, U) eine Spannbaum-Struktur. Die eindeutig bestimmte Funktion $\pi . V(G) \to \mathbb{R}$ mit $\pi(r) = 0$ und $c_\pi(e) := c(e) - \pi(v) + \pi(w) = 0$ für alle $e \in (v, w) \in T$ heißt das zu (r, T, L, U) **assoziierte Potential**.

Observe. *Die π-Werte in einem zu (r, T, L, U) assoziierten Potential geben den Abstand zu r in $(V(G), T)$ an.*

Theorem 86. Zu gegebenen (G, u, b, c) und gegebener Spannbaum-Struktur (r, T, L, U) können der b-Fluss f und das Potential π zu (r, T, L, U) in Zeit $\mathcal{O}(m)$ berechnet werden.

Beweis. Verwende Breitensuche, um π zu berechnen. Von den Blättern ausgehend können ferner in linearer Zeit die Flusswerte auf T berechnet werden. □

Observe.

- *Für eine Kante $e = (v, w) \in L$ sind $c_\pi(e)$ die reduzierten Kosten der zugehörigen Nicht-Basisvariablen.*
- *Für eine Kante $e \in U$ sind $-c_\pi(e)$ die reduzierten Kosten der zugehörigen Nicht-Basisvariablen (d.h. der Slack-Variablen).*

Theorem 87. Sei (r, T, L, U) eine zulässige Spannbaum-Struktur und π das zugehörige Potential. Wenn $c_\pi(e) \geqslant 0$ für alle $e \in L$ und $c_\pi(e) \leqslant 0$ für alle $e \in U$, dann ist der zu (r, T, L, U) gehörige b-Fluss optimal.

Beweis. In diesem Fall haben alle Nicht-Basisvariablen nicht-negative Kosten. In Minimierungsproblemen bedeutet das, dass die Lösung optimal ist. □

Definition 88. Für eine Kante $e = (v, w) \in E(\overleftrightarrow{G}) \backslash T$ mit $\overleftarrow{e} \notin T$ heißt e zusammen mit dem w-v-Weg, der nur aus Kanten in T und Rückwärtskanten von Kanten in T besteht, **Fundamentalkreis** von e.

Der Knoten im Kreis, der am nächsten zu r liegt, heißt **Gipfel** von e.

Observe. *Es sind $\pi(w)$ die Kosten des Weges in T von r zu w und $\pi(v)$ die Kosten des Weges von r zu v. Insbesondere ist $c_\pi(e)$ gleich der Kosten des Fundamentalkreises von e.*

Theorem 89. Der Netzwerk-Simplex-Algorithmus terminiert nach endlich vielen Schritten und berechnet eine Optimallösung.

Beweis. Die Optimalität der Lösung folgt aus dem vorigen Satz.

Übung: Zeigen, dass f und π nach einer Iteration weiter der zu (r, T, L, U) gehörende Fluss und das zugehörige Potential sind.

Behauptung 1. *(r, T, L, U) bleibt stark zulässig.*

Unterbeweis. Die Zulässigkeit folgt aus der Wahl von γ. Da e_1 stets als letzte Kante aus dem Fundamentalkreis gewählt wird, ist (r, T, L, U) außerdem stark zulässig.

■

Behauptung 2. *Der Algorithmus terminiert.*

Unterbeweis. Wir zeigen, dass nie dieselbe Baumstruktur zweimal betrachtet wird. Die Kosten des Flusswertes ändern sich um $\gamma \cdot |\rho|$, falls $\gamma > 0$ sind wir daher fertig (der Flusswert wird nie verschlechtert).

Algorithmus 2 : Netzwerk Simplex Algorithmus

Data : Eine Instanz (G, u, b, c) von Min-Cost-Flow und eine stark zulässige Spannbaumstruktur (r, T, L, U)

Result : Ein kostenminimaler Fluss f

```
Berechne den zu (r, T, L, U) gehörenden b-Fluss f sowie π.
while true do
    Sei e_0 ∈ {e ∈ L | c_π(e_0) < 0} ∪ {e ∈ U | c_π(e) > 0}
    if keine solche Kante existiert then
        return f
    else
        Sei C der Fundamentalkreis von e (bzw. ←e falls e ∈ U) und
          ρ := c_π(e)
        γ := min_{e'∈E(C)} u_f(e')
        e' := letzte Kante auf C mit u_f(e') = γ (vom Gipfel aus)
        e_1 := zu e' zugehörige Kante in G (d.h. e' = e_1 oder e' = ←e_1)
        Entferne e aus L oder U
        T := (T ∪ {e_0})\{e_1}
        if e' = e_1 then
            U := U ∪ {e_1}
        else
            L := L ∪ {e_1}
        Augmentiere f entlang C um γ.
        Sei X die Zusammenhangskomponente von (V(G), T\{e_0}), welche r
          enthält.
        if e_0 ∈ δ^+(X) then
            π(v) := π(v) + ρ für alle v ∈ V(G)\X
        if e_0 ∈ δ^-(X) then
            π(v) := π(v) − ρ für alle v ∈ V(G)\X
```

Angenommen $\gamma = 0$. Falls $e_0 \neq e_1$, dann gilt $e_0 \in L \cap \delta^-(X)$ oder $e_0 \in U \cap \delta^+(X)$. Folglich wird $\sum_{v \in V(G)} \pi(v)$ größer. Solange keine Iteration mit $\gamma > 0$ durchgeführt wird, wird dieser Wert nicht kleiner.

Angenommen $e_0 = e_1$. Dann ist $X = V(G)$ und $\sum_{v \in V(G)} \pi(v)$ bleibt gleich. Aber $|\{e \in L | c\pi(e) < 0\}| + |\{e \in U | c\pi(e) > 0\}$ wird um 1 kleiner. ■

□

Question 89.1. Wie finden wir eine stark zulässige Startlösung?

Verbinde r mit hinreichend teuren Kanten mit jedem Knoten. Wähle den Fluss entsprechend (die Richtung der Kanten seien je nach b-Wert des Knotens gewählt). Die Kapazitäten seien dabei so gewählt, dass $u((v, r)) = b(v)+1$, bzw. $u((r, v)) = -b(v)$.

T bestehe aus den neuen Kanten. Außerdem sei $L := E(G)$ und $U := \emptyset$.

7 Größen von Lösungen

Definition 90. Sei $n \in \mathbb{Z}$. Wir definieren

$$\operatorname{size}(n) := 1 + \lceil \log(|n| + 1) \rceil$$

Für $p, q \in \mathbb{Z}$ teilerfremd sei

$$\operatorname{size}\left(\frac{p}{q}\right) := \operatorname{size}(p) + \operatorname{size}(q)$$

Für $A \in \mathbb{Q}^{m \times n}$ sei

$$\operatorname{size}(A) := nm + \sum_{i,j} \operatorname{size}(a_{i,j})$$

Man sieht leicht:

Theorem 91. Für $r_1, \dots, r_n \in \mathbb{Q}$ gilt:

(a) $\operatorname{size}\left(\prod_i r_i\right) \leqslant \sum_i \operatorname{size}(r_i)$

(b) $\operatorname{size}\left(\sum_i r_i\right) \leqslant 2 \sum_i \operatorname{size}(r_i)$

Theorem 92. Für $x, y \in \mathbb{Q}^n$ gilt:

(a) $\operatorname{size}(x + y) \leqslant 2(\operatorname{size}(x) + \operatorname{size}(y))$

(b) $\operatorname{size}(x^t y) \leqslant 2(\operatorname{size}(x) + \operatorname{size}(y))$

Theorem 93. Für jede Matrix $A \in \mathbb{Q}^{n \times n}$ gilt $\operatorname{size}(\det(A)) \leqslant 2 \operatorname{size}(A)$.

Beweis. Sei $A = (\frac{p_{ij}}{q_{ij}})_{i,j}$, wobei $\operatorname{ggT}(p_{ij}, q_{ij}) = 1$. Sei ferner $\det(A) = \frac{p}{q}$ (vollständig gekürzt).

Dann ist $|\det(A)| \leqslant \prod_{i=1}^n \prod_{j=1}^n (|p_{ij}| + 1)$ und $|q| \leqslant \prod_{i=1}^n \prod_{j=1}^n |q_{ij}|$. Folglich ist

$$\operatorname{size}(q) \leqslant \operatorname{size}(A)$$

und

$$|p| = |\det(A)| \cdot |q| \leqslant \prod_{i=1}^n \prod_{j=1}^n ((|p_{ij}| + 1)|q_{ij}|)$$

Folglich ist $\operatorname{size}(p) \leqslant \sum_{i=1}^n \sum_{j=1}^n (\operatorname{size}(p_{ij}) + 1 + \operatorname{size}(q_{ij})) = \operatorname{size}(A)$ und es folgt die Aussage. $\square$

Theorem 94. Sei $\max\{c^t x | Ax \leqslant b\}$ ein zulässiges und beschränktes lineares Programm mit $A \in \mathbb{Q}^{m \times n}$ und $b \in \mathbb{Q}^m$. Dann gibt es eine optimale (rationale) Lösung x mit $\text{size}(x) \leqslant 4n(\text{size}(A) + \text{size}(b))$. Wenn $b = e_i$ oder $b = -e_i$ für einen Einheitsvektor e_i gilt, dann gibt es eine reguläre Teilmatrix A' von A und eine Optimallösung x mit $\text{size}(x) \leqslant 4n\,\text{size}(A')$.

Beweis. Das Optimum wird in einer minimalen Fläche F von $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ angenommen.

Wir können F schreiben als $F = \{x \in \mathbb{R}^n | \tilde{A}x = \tilde{b}\}$ wobei $\tilde{A}$ ein Teilsystem mit linear unabhängigen Zeilen ist.

Wähle $B \subseteq \{1, \dots, n\}$, so dass $\tilde{A}_B$ eine reguläre, quadratische Matrix ist. x mit $x_B = \tilde{A}_B^{-1}\tilde{b}$ und $x_N = 0$ mit $N = \{1, \dots, n\} \backslash B$ ist eine optimale LP-Lösung.

Nach der Cramerschen Regel können die Einträge von x_B als

$$x_j = \frac{\det\left(\tilde{A}_j\right)}{\det(\tilde{A}_B)}$$

wobei $\tilde{A}_j$ aus $\tilde{A}_B$ entsteht, indem die j-te Spalte durch $\tilde{b}$ ersetzt wird.

Folglich ist

$$\text{size}(x) \leqslant n + 2\sum_j (\text{size}(\tilde{A}_j) + \text{size}(\tilde{A}_B)) \leqslant n + 4n(\text{size}(A) + \text{size}(b))$$

Falls $b \in \{e_i, -e_i\}$ ist, dann ist $|\det(\tilde{A}_j)|$ der Absolutbetrag einer Determinante einer Teilmatrix von $\tilde{A}_B$.

□

Corollary 95. Sei $\max\{c^t x | Ax \leqslant b\}$ ein zulässiges und beschränktes lineares Programm mit $A \in \mathbb{Q}^{m \times n}$ und $b \in \mathbb{Q}^m$. Dann gibt es eine optimale (rationale) Lösung x, sodass für jeden von 0 verschiedenen Eintrag x_j von x gilt:

$$|x_j| \geqslant 2^{-4n(\text{size}(A)+\text{size}(b))}$$

Beweis. Es gibt eine Optimallösung x, sodass für jeden Eintrag x_j von x gilt:

$$\text{size}(x_j) \leqslant 4n(\text{size}(A) + \text{size}(b))$$

Da jede positive Zahl, die kleiner als $2^{-4n(\text{size}(A)+\text{size}(b))}$ ist, eine size hat, die größer als $4n(\text{size}(A) + \text{size}(b))$ ist, folgt die Behauptung. □

7.1 Gauß-Elimination

Wir wollen ein Gleichungssystem $Ax = b$ lösen.

Transformiere A dazu in eine obere rechte Dreiecksmatrix mit folgenden erlaubten Schritten:

- Addiere ein Vielfaches einer Zeile zu einer anderen.
- Vertausche zwei Spalten
- Vertausche zwei Zeilen

Aus AlMa 1 ist bekannt, dass hierfür $\mathcal{O}(mn(\text{rank}(A)+1))$ elementare Operationen ausreichen. Um eine polynomielle Laufzeit beweisen zu können, muss jedoch noch gezeigt werden, dass alle auftretenden Zahlen mit polynomiell vielen Bits geschrieben werden können.

Sei $\tilde{A} = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$ eine Matrix, die während der Gauß-Elimination betrachtet wird. B ist dann eine rechte obere $k \times k$-Dreiecksmatrix. Zur Abschätzung der Größe der Zahlen sind Zeilen- und Spaltenvertauschungen nicht relevant. Für jeden Eintrag d_{ij} von D gilt $\det(\tilde{A}^{1\dots k,k+1}_{1\dots k,k+j}) = d_{ij} \cdot \det(\tilde{A}^{1\dots k}_{1\dots k})$ wobei $M^{i_1\dots i_k}$ die Untermatrix einer Matrix M mit Zeilenindizes $i_1, \dots, i_k$ und Spaltenindizes $j_1, \dots, j_k$ ist.

Das Addieren eines Vielfachen einer Zeile zu einer anderen ändert nichts an den Determinanten.

Folglich kann d_{ij} (und damit auch jede andere auftretende Zahl) als Quotient von Determinanten von Untermatrizen von A geschrieben werden.

Folglich ist $\text{size}(d_{ij}) \leqslant 4\,\text{size}(A)$. Da alle während des Gauß-Algorithmus auftretenden Zahlen von dieser Form sind, folgt:

Theorem 96. Die Gauß-Elimination hat polynomielle Laufzeit.

Corollary 97. Folgende Probleme können in polynomieller Laufzeit gelöst werden:

- Lösen eines Gleichungssystems
- Determinantenberechung
- Berechnung des Rangs einer Matrix
- Invertieren einer regulären Matrix
- Überprüfen von linearer Unabhängigkeit

8 Die Ellipsoid-Methode

8.1 Idealisierte Ellipsoid-Methode

Die Ellipsoid-Methode ist der erste Algorithmus zur LP-Lösung mit beweisbar polynomieller Laufzeit.

Der Algorithmus wurde von Khachiyan 1979 entwickelt. In der Grundversion wird nur entschieden, ob ein Polyeder leer ist. Der Algorithmus kann verwendet werden, ohne dass alle Nebenbedingungen aufgelistet werden müssen. Er ist auch für andere (nicht-lineare) Optimierungsprobleme geeignet.

Definition 98. Eine Menge $E \in \mathbb{R}^n$ ist ein **Ellipsoid**, wenn es einen Vektor $s \in \mathbb{R}^n$ und eine reguläre Matrix $M \in \mathbb{R}^{n \times n}$ gibt, sodass

$$E = \{Mx + s | x \in B^n\}$$

wobei $B^n = \{x \in \mathbb{R}^n | x^t x \leqslant 1\}$ die n-dimensionale Einheitskugel ist.

Kurznotation: $E = s + MB^n$.

Definition 99. Eine symmetrische Matrix A heißt **positiv definit**, wenn $x^t Ax > 0$ für alle $x \neq 0$ gilt. Sie heißt **positiv semidefinit**, wenn $x^t Ax \geqslant 0$ für alle x gilt.

Remark 99.1. Eine $n \times n$-Matrix Q ist genau dann positiv definit, wenn es eine nicht-singuläre Matrix M mit $Q = MM^t$ gibt (Cholesky-Zerlegung).

Lemma 100. Eine Menge $E \subseteq \mathbb{R}^n$ ist genau dann ein Ellipsoid, wenn es eine symmetrische positiv definite $n \times n$-Matrix Q und einen Vektor $s \in \mathbb{R}^n$ mit $E = \{x \in \mathbb{R}^n | (x - s)^t Q^{-1} (x - s) \leqslant 1\}$ gibt.

Beweis. $E \subseteq \mathbb{R}^n$ ist genau dann ein Ellipsoid, wenn es eine reguläre Matrix $M \in \mathbb{R}^{n \times n}$ und einen Vektor $s \in \mathbb{R}^n$ gibt, mit

$$\begin{aligned} E &= (Mx + s | x \in B^n) \\ &= \{y \in \mathbb{R}^n | M^{-1}(y - s) \in B^n\} \\ &= \{y \in \mathbb{R}^n | (y - s)^t (M^{-1})^t M^{-1} (y - s) \leqslant 1\} \end{aligned}$$

Das ist äquivalent zur Existenz einer positiv definiten Matrix Q mit $E = \{y \in \mathbb{R}^n | (y - s)^t Q^{-1} (y - s) \leqslant 1\}$. $\square$

Goal. *Finde ein Ellipsoid mit kleinstem Volumen, das eine Halbkugel der Einheitskugel B^n enthält.*

Ansatz: Wähle als Mittelpunkt des Ellipsoids eine Position $c \cdot e_1$. Kandidaten für das Ellipsoid sind gegeben durch Matrizen Q^{-1} der Form

$$Q^{-1} = \begin{pmatrix} \alpha^2 & 0 & \dots & 0 \\ 0 & \beta^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \beta^2 \end{pmatrix}$$

$$E = \left\{ x \in \mathbb{R}^n \mid \alpha(x_1 - c) + \beta \sum_{i=2}^{n} x_i \leqslant 1 \right\}$$

wobei α, β und c zu bestimmen sind.

Wunsch 1: e_1 soll auf dem Rand von E liegen. Fordere: $\alpha^2(1-c)^2 = 1$, d.h.

$$\alpha^2 = \frac{1}{(1-c)^2} \tag{12}$$

Wunsch 2: Die Punkte $\{x \in \mathbb{R}^n | x_1 = 0\}$, die auf dem Rand von B^n liegen, sollen auf dem Rand von E liegen. Fordere $\alpha^2 c^2 + \beta^2 = 1$ und somit:

$$\beta^2 = 1 - \alpha^2 c^2 = 1 - \frac{c^2}{(1-c)^2} = \frac{1-2c}{(1-c)^2} \tag{13}$$

Lemma 101. Das Volumen eines Ellipsoids $E = \{x \in \mathbb{R}^n | (x-s)Q^{-1}(x-s) \leqslant 1\}$ ist $\operatorname{vol}(E) = \sqrt{\det(Q)} \cdot \operatorname{vol} B^n$

Beweis. Aus der Maßtheorie bekannt. □

Das Ziel ist daher, die Determinante von Q^1 zu minimieren. Es ist

$$\sqrt{\det Q^{-1}} = \sqrt{\alpha^{-1}\beta^{-(n-1)}}$$

Wir wollen folglich ein c finden, welches $\frac{(1-c)^{2n}}{(1-2c)^{n-1}}$ minimiert. Es ist $\frac{\partial}{\partial c} \frac{(1-c)^{2n}}{(1-2c)^{n-1}} = \frac{2(n-1)(1-c)^{2n}}{(1-2c)^n} - \frac{2n(1-c)^{2n-1}}{(1-2c)^{n-1}}$ genau dann 0, wenn $\frac{2(n-1)(1-c)}{1-2c} = 2n$. Folglich muss $2(n-1) - 2c(n-1) = 2n - 4cn$ und $c(2n-(n-1)) = 1$ gelten. Daher wird das Volumen durch $c := \frac{1}{n+1}$ minimiert und wir erhalten $\alpha^2 = \frac{(n+1)^2}{n^2}$ sowie $\beta^2 = \frac{n^2-1}{n^2}$.

Lemma 102 (Halbkugel-Lemma)**.** $B^n \cap \{x \in \mathbb{R}^n | x_1 \geqslant 0\} \subseteq E$ mit

$$E = \left\{ x \in \mathbb{R}^n y \frac{(n+1)^2}{n^2} \left(x_1 - \frac{1}{n+1} \right)^2 + \frac{n^2-1}{n^2} \sum_{l=2}^{i} x_i^2 \leqslant 1 \right\}$$

Außerdem ist $\frac{\operatorname{vol}(E)}{\operatorname{vol} B^n} \leqslant e^{-\frac{1}{2(n+1)}}$.

Beweis. Sei $x \in B^n \cap \{x \in \mathbb{R}^n | x_1 \geqslant 0\}$. Es gilt

$$\sum_{i=2}^{n} x_i^2 \leqslant 1 - x_1^2$$

Daher genügt es zu zeigen, dass

$$g(x_1)_i := \frac{(n+1)^2}{n^2}\left(x_1 - \frac{1}{n+1}\right)^2 + \frac{n^2-1}{n^2}\left(1 - x_1^2\right) \leqslant 1$$

Für $x_1 = 0$ gilt: $g(0) = \frac{(n+1)^2}{n^2}\frac{1}{(n+1)^2} + \frac{n^2-1}{n^2} = 1$. Für $x_1 = 1$ gilt: $g(1) = \frac{(n+1)^2}{n^2}\left(\frac{n}{n+1}\right)^2 = 1$. g ist eine quadratische Funktion und der Koeffizient von x_1^2 ist $\frac{(n+1)^2}{n^2} - \frac{n^2-1}{n^2} > 0$, folglich handelt es sich um eine konvexe Funktion. Daher gilt $g(x_1) \leqslant 1$ für alle $x_1 \in [0,1]$.

Für das Volumen gilt:

$$\begin{aligned}
\frac{\operatorname{vol}(E)}{\operatorname{vol} B^n} &= \sqrt{\det Q} \\
&= \frac{1}{\alpha \cdot \beta^{(n-1)}} \\
&= \frac{n}{n+1}\left(\frac{n^2}{n^2-1}\right)^{\frac{n-1}{2}} \\
&= \left(1 - \frac{1}{n+1}\right)\left(1 + \frac{1}{n^2-1}\right)^{\frac{n-1}{2}} \\
&\overset{1+x \leqslant e^x}{\leqslant} e^{-\frac{1}{n+1}} \cdot e^{\frac{n-1}{2(n^2-1)}} \\
&= e^{-\frac{1}{2(n+1)}}
\end{aligned}$$

□

Wir haben nicht gezeigt, dass dies das kleinste Ellipsoid ist (obwohl das der Fall ist), es genügt uns zu zeigen, dass das Ellipsoid klein genug ist.

Lemma 103 (Halb-Ellipsoid-Lemma)**.** Sei $E = p + \{x \in \mathbb{R}^n | x^2 Q^{-1} x \leqslant 1\}$ ein Ellipsoid und $a \in \mathbb{R}^n$ mit $a^t Q a = 1$. Dann gilt

$$E \cap \{x \in \mathbb{R}^n | a^t x \geqslant a^t p\} \subseteq E'$$

mit

$$E' = p + \frac{1}{n+1} Qa + \left\{x \in \mathbb{R}^n | \frac{n^2-1}{n^2} x^t \left(Q^{-1} + \frac{2}{n-1} aa^t\right) x \leqslant 1\right\}$$

und

$$\frac{\text{vol}(E')}{\text{vol}(E)} \leqslant e^{-\frac{1}{2(n+1)}}$$

Beweis. Lang und nervig

Sei M eine reguläre $n \times n$-Matrix mit $Q = MM^t$. O.B.d.A ist $a^t M = e_1^t$ und somit $Qa = MM^t a = M(a^t)M)^4 = Me_1$ (sonst multipliziere M mit einer Rotationsmatrix, die $a^t M$ auf e_1 abbildet. Dann gilt:

$$\begin{aligned}
& E \cap \{x \in \mathbb{R}^n | a^t x \geqslant a^t p\} \\
= \quad & (p + MB^n) \cap \{x \in \mathbb{R}^n | a^t x \geqslant a^t p\} \\
= \quad & p + (MB^n \cap \{x \in \mathbb{R}^n | a^t(x + p \geqslant a^t p\}) \\
= \quad & p + (MB^n \cap \{x \in \mathbb{R}^n | a^t x \geqslant 0\}) \\
= \quad & p + M(B^n \cap M^{-1}\{x \in \mathbb{R}^n | a^t x \geqslant 0\}) \\
= \quad & p + M(B^n \cap \{x \in \mathbb{R}^n | a^t Mx \geqslant 0\}) \\
= \quad & p + M(B^n \cap \{x \in \mathbb{R}^n | a^t Mx \geqslant 0\}) \\
= \quad & p + M(B^n \cap \{x \in \mathbb{R}^n | e_1^t x \geqslant 0\}) \\
\overset{\text{Lemma 102}}{\subseteq} \quad & p + \frac{1}{n+1} Me_1 + M\{x \in \mathbb{R}^n | \frac{n^2-1}{n^2} x^t \left(I_N + \frac{2}{n-1} e_1 e_1^t\right) x \leqslant 1\} \\
= \quad & p + \frac{1}{n+1} Me_1 + M\{x \in \mathbb{R}^n | \frac{n^2-1}{n^2} x^t \left(I_N + \frac{2}{n-1} e_1 e_1^t\right) x \leqslant 1\} \\
= \quad & p + \frac{1}{n+1} Qa + \{x \in \mathbb{R}^n | \frac{n^2-1}{n^2} x^t \left(Q^{-1} + \frac{2}{n-1} aa^t\right) x \leqslant 1\}
\end{aligned}$$

Wir können E' in Standardform schreiben als

$$E' = p + \frac{1}{n+1} Qa + \{x \in \mathbb{R}^n | x^t \tilde{Q}^{-1} x \leqslant 1\}$$

mit

$$\tilde{Q} = \frac{n^2}{n^2-1} \left(Q - \frac{2}{n+1} Qaa^t Q^t\right)$$

. Denn

$$\begin{aligned}
& \frac{n^2-1}{n^2} \left(Q^{-1} + \frac{2}{n-1} aa^t\right) \frac{n^2}{n^2-1} \left(Q - \frac{2}{n+1} Qaa^t Q^t\right) \\
&= I_n - \frac{2}{n+1} aa^t Q^t + \frac{2}{n-1} aa^t Q - \frac{4}{n^2-1} a \underbrace{a^t Qa}_{=1} a^t Q^t \\
&= I_n
\end{aligned}$$

Es ist $\frac{\text{vol}(E')}{\text{vol}(E)} = \sqrt{\frac{\det(\tilde{Q})}{\det(Q)}}$ und außerdem

$$
\begin{aligned}
\frac{\det(\tilde{Q})}{\det(Q)} &= \det\left(\frac{n^2}{n^2-}1\left(I_n - \frac{2}{n+1}aa^tQ^t\right)\right) \\
&= \left(\frac{n^2}{n^2-1}\right)^n \det\left(I_n - \frac{2}{n+1}aa^tQ^t\right) \\
&= \left(\frac{n^2}{n^2-1}\right)^n \left(1 - \frac{2}{n+1}\right)
\end{aligned}
$$

Eigenwerte von $I_n - \frac{2}{n+1}aa^tQ^t$ sind $1 - \frac{2}{n+1}$ (zum Eigenvektor a) und 1. Die Determinante ist Produkt der Eigenwerte. Daher folgt

$$
\begin{aligned}
\sqrt{\frac{\det(\tilde{Q})}{\det(Q)}} &\leqslant \left(\frac{n^2}{n^2-1}\right)^{\frac{n}{2}} \left(1 - \frac{2}{n+1}\right)^{\frac{1}{2}} \\
&= \frac{n}{n+1}\left(\frac{n^2}{n^2-1}\right)^{\frac{n-1}{2}} \\
&\overset{\text{Wie beim Halbkugellemma}}{\leqslant} e^{-\frac{1}{2(n+1)}}
\end{aligned}
$$

□

Remark 103.1. Das Ellipsoid E' wird auch **Löwner-John Ellipsoid** genannt. Es ist das kleinste Ellipsoid, welches $E \cap \{x \in \mathbb{R}^n | a^t x \geqslant a^t p\}$ enthält.

Definition 104. Ein **Separationsorakel** für eine konvexe Menge $K \subseteq \mathbb{R}^n$ ist ein Black-Box-Algorithmus, der zu gegebenem $x \in \mathbb{R}^n$ entweder einen Vektor $a \in \mathbb{R}^n$ mit $a^t y > a^t x$ für alle $y \in K$ zurückgibt oder ausgibt, dass $x \in K$ gilt.

Observe. *Zu einem gegebenen $A \in \mathbb{Q}^{m\times n}$ und $b \in \mathbb{Q}^m$ kann ein Separationsorakel für $\{x \in \mathbb{R}^n | Ax \leqslant b\}$ in $\mathcal{O}(mn)$ arithmetischen Operationen implementiert werden.*

Theorem 105. Zu einem durch ein Separationsorakel gegebenen $K \subseteq \mathbb{R}^n$, $\varepsilon > 0$ und R mit $K \subseteq \{x \in \mathbb{R}^n | x^t x \leqslant R\}$ kann man mit $\mathcal{O}(n(n\ln(R) + \ln(\frac{1}{\varepsilon})))$ Iterationen des idealisierten Ellipsoid-Verfahrens ein $x \in K$ berechnen oder korrekterweise $\text{vol}(K) < \varepsilon$ ausgeben. Jede Iteration benötigt einen Orakelaufruf, $\mathcal{O}(n^2)$ arithmetische Standardoperationen und die Berechnung einer Quadratwurzel von einer reellen Zahl.

Beweis.

Behauptung 1. *In jeder Iteration ist K enthalten in $p_k + \{x \in \mathbb{R}^n | x^t A_k^{-1} x \leqslant 1\}$*

Algorithmus 3 : Idealized Ellipsoid Algorithm

Data : Separationsorakel für $K \subseteq \mathbb{R}^n$ (konvex, abgeschlossen), $R > 0$ mit $K \subseteq \{x \in \mathbb{R}^n | x^t x \leqslant \mathbb{R}^2\}$, $\varepsilon > 0$

Result : $x \in K$ oder "vol$(K) < \varepsilon$"

$p_0 := 0$, $A_0 := R^2 I_n$

$N(R, \varepsilon) := \left\lfloor 2(n+1)(n \ln(2R) + \ln(\frac{1}{\varepsilon})) \right\rfloor$

for $k = 0, \dots, N(R, \varepsilon)$ **do**

  if $p_k \in K$ **then**

    return p_k

  Sei $\overline{a} \in \mathbb{R}^n$ ein Vektor mit $\overline{a}^t y > \overline{a^t} p_k$ für alle $y \in K$.

  $b_k := \frac{A_k \overline{a}}{\sqrt{\overline{a^t} A_k \overline{a}}}$

  $p_{k+1} := p_k + \frac{1}{n+1} b_k$

  $A_{k+1} := \frac{n^2}{n^2 - 1} \left(A_{k-1} - \frac{2}{n+1} b_k b_k^t \right)$

return "vol$(K) < \varepsilon$"

Unterbeweis. Für $k = 0$ ist dies trivial. Die Aussage folgt durch induktives Anwenden des Halb-Ellipsoid-Lemma (Lemma 103) auf $Q = A_k$ und $a = \frac{\overline{a}}{\sqrt{\overline{a^t} A_k \overline{a}}}$. ■

Behauptung 2. *Nach* $N(R, \varepsilon)$ *Iterationen ist* vol$(p_k + \{x \in \mathbb{R}^n | x^t A_k^{-1} x \leqslant 1\}) < \varepsilon$.

Unterbeweis. Es gilt vol$(\{x \in \mathbb{R}^n | x^t x \leqslant \mathbb{R}\}) \leqslant$ vol$([-R, R]^n) = 2^n R^n$. In jeder Iteration wird das Volumen von

$$E_k = p_k + \{x \in \mathbb{R}^n | x^t A_k^{-1} x \leqslant 1\}$$

mindestens um den Faktor $e^{-\frac{1}{2(n+1)}}$ verringert.

Folglich ist

$$\text{vol}(E_k) \leqslant e^{-\frac{K}{2(n+1)}} 2^n R^n$$

Mit $k = N(R, \varepsilon)$ folgt die Aussage. ■

□

8.2 Fehleranalyse

Problem. *Wir können in* $b_k := \frac{A_k \overline{a}}{\sqrt{\overline{a}^t A_k \overline{a}}}$ *die Wurzel nicht exakt ausrechnen.*

Wir müssen daher mit gerundeten Zwischenlösungen rechnen.

Es seien $\tilde{p}_k$ und $\tilde{A}_k$ die exakten Werte (die allerdings aus gerundeten Werten der vorigen Iteration berechnet werden) und p_k, A_k gerundetete Werte.

Seien $\tilde{E}_k$ und E_k die zugehörigen Ellipsoide.

Sei δ eine obere Schranke für den absoluten Rundungsfehler, also $\|p_k - \tilde{p}_k\|_\infty \leqslant \delta$ und $\|A_k - \tilde{A}_k\|_\infty \leqslant \delta$

Beim Runden der Einträge in $\tilde{A}_k$ sorgen wir dafür, dass die Matrix symmetrisch bleibt.

Sei $\Gamma_k = A_k - \tilde{A}_k$ und $\Delta_k = p_k - \tilde{p}_k$.

Es sei $\|\cdot\|$ für Vektoren die Euklidische Norm und für Matrizen die induzierte Operator-Norm.

Wir können annehmen, dass für jedes $x \in K$ gilt:

$$(x - \tilde{p}_k)^t \tilde{A}_k^{-1}(x - \tilde{p}_k) \leqslant 1$$

Für p_k und A_k muss das aber nicht gelten. Daher vergrößern wir das Ellipsoid in jeder Iteration leicht durch Skalierung von $\tilde{A}_k$ um den Faktor $\mu = 1 + \frac{1}{2n(n+1)}$. Wir ersetzen $\tilde{A}_k$ durch $\mu\tilde{A}_k$ (das Ergebnis heiße im Folgenden wieder $\tilde{A}_k$).

Dann gilt für $x \in K$:

$$(x - \tilde{p}_k)^t \tilde{A}_k^{-1}(x - \tilde{p}_k) \leqslant \frac{1}{1 + \frac{1}{2n(n+1)}} = \frac{2n^2 + 2n}{2n^2 + 2n + 1} < 1 - \frac{1}{4n^2}$$

Mit einer langen und nervigen Rechnung folgt, dass das genügt, um die ursprüngliche Bedingung auch nach Runden noch zu erfüllen.

Theorem 106. Es sei in Iteration k der Ellipsoid-Methode $\delta \leqslant \frac{1}{12n_4^k}$ gewählt. Dann gilt:

(a) A_k ist positiv definit.

(b) $\|p_k\| \leqslant R_2^k$, $\|\tilde{p}_k\| \leqslant R_2^k$

(c) $\|A_k\| \leqslant R^2 2^k$, $\|\tilde{A}_k\| \leqslant R^2 2^k$

(d) $\|A_k^{-1}\| \leqslant R^{-2} 4^k$, $\|\tilde{A}_k^{-1}\| \leqslant R^{-2} 4^k$

Goal. *Wähle δ so, dass*

- $$2\sqrt{n}\delta\|\tilde{A}_k^{-1}\|(R + \|\tilde{p}_k\|) + n\delta^2\|\tilde{A}_k^{-1}\| + (R + \|pk\|)^2\|A_k^{-1}\|\|\tilde{A}_k^{-1}\|n\delta \leqslant \frac{1}{4n^2}$$
- $\delta\|\tilde{A}_{k+1}^{-1}\| \leqslant \frac{1}{4(n+1)^3}$

Algorithmus 4 : Ellipsoid Algorithm

Data : Separationsorakel für $K \subseteq \mathbb{R}^n$ (konvex, abgeschlossen), $R > 0$ mit $K \subseteq \{x \in \mathbb{R}^n | x^t x \leqslant \mathbb{R}^2\}$, $\varepsilon > 0$

Result : $x \in K$ oder "vol$(K) < \varepsilon$"

$p_0 := 0$, $A_0 := R^2 I_n$

$N(R, \varepsilon) := \lfloor 8(n+1)(n \ln(2R) + \ln(\frac{1}{\varepsilon})) \rfloor$

for $k = 0, \dots, N(R, \varepsilon)$ **do**

  if $p_k \in K$ **then**

    return p_k

  Sei $\overline{a} \in \mathbb{R}^n$ ein Vektor mit $\overline{a}^t y > \overline{a^t} p_k$ für alle $y \in K$.

  $b_k := \frac{A_k \overline{a}}{\sqrt{\overline{a}^t A_k \overline{a}}}$

  p_{k+1} eine Approximation von $\tilde{p_{k+1}} := p_k + \frac{1}{n+1} b_k$ mit Fehler $\leqslant \delta$

  A_{k+1} eine symmetrische Approximation von

  $\tilde{A_{k+1}} := \frac{n^2}{n^2-1} \left(A_{k-1} - \frac{2}{n+1} b_k b_k^t \right)$ mit Fehler $\leqslant \delta$

return "vol$(K) < \varepsilon$"

Theorem 107. Es sei in Iteration k der Ellipsoid-Methode $\delta \leqslant \frac{1}{12 n_4^k}$ gewählt. Dann gilt:

- A_k ist positiv definit
- $\|p_k\| \leqslant R \cdot 2^k$, $\|\tilde{p}_k\| \leqslant R \cdot 2^k$
- $\|A_k\|, \|\tilde{A}_k\| \leqslant R^2 \cdot 2^k$
- $\|A_k^{-1}\|, \|\tilde{A}_k^{-1}\| \leqslant R^{-2} 4^k$

Insbesondere kann $\delta < (2^{6N(R,\varepsilon)+1)} 16 n^3)^{-1}$ gewählt werden.

Beweis. Wir führen Induktion nach k. Für $k = 0$ sind diese Aussagen trivial. Sei die Aussage für k bereits gezeigt. A_k^{-1} ist positiv definit, wegen

$$\forall x \neq 0: \; x^t A_k^{-1} x = x^t A_k^{-1} A_k A_k^{-1} x > 0$$

Ferner ist

$$\tilde{A}_{k+1}^{-1} = \frac{n^2 - 1}{n^2} \frac{1}{\mu} \left(A_k^{-1} + \frac{2}{n+1} \frac{\overline{a a}^t}{\overline{a}^t A_k \overline{a}} \right)$$

als Summe einer positiv definiten und einer positiv semidefiniten Matrix positiv definit.

Somit ist auch $\tilde{A}_{k+1} = \frac{n^2}{n^2-1} \mu \left(A_k - \frac{2}{n+1} b_k b_k^t \right)$ positiv definit.

Es gilt

$$\left\| \frac{\overline{a a}^t}{\overline{a}^t A_k \overline{a}} \right\| = \frac{\overline{a}^t \overline{a}}{\overline{a}^t A_k \overline{a}} \leqslant \|A_k^{-1}\|$$

(Übungsblatt 7)

Hieraus ergibt sich

$$\|\tilde{A}_{k+1}^{-1}\| \leqslant \frac{n^2-1}{n^2}\frac{1}{\mu}\left(\|A_k^{-1}\| + \frac{2}{n-1}\|\frac{\overline{a}\overline{a}^t}{\overline{a}^t A_k\|a\|}\|\right) \leqslant \|A_k^{-1}\|$$

Es sei λ kleinster Eigenwert von A_{k+1} und v ein Vektor mit Länge 1 und $\lambda = v^t A_k v$.

Folglich ist

$$\begin{aligned}
v^t A_{k+1} v &\geqslant v^t \tilde{A}_{k_1} v - n\delta \\
&\geqslant \min\{n^t \tilde{A}_{k+1} n | n \in \mathbb{R}^n, \|n\| = 1\} - n\delta \\
&\geqslant \frac{1}{\|\tilde{A}_{k+1}^{-1}\|} - n\delta \\
&\geqslant \frac{1}{3\|A_k^{-1}\|} - n\delta \\
&\overset{I.V.}{\geqslant} \frac{1}{3R^{-2}4^k} - n\delta \\
&\geqslant \frac{1}{R^{-2}4^{k+1}}
\end{aligned}$$

Wobei im letzten Schritt $n\delta \leqslant (\frac{1}{3} - \frac{1}{4})\frac{R^2}{4^k}$ verwendet wurde.

Da der kleinste Eigenwert positiv ist, ist A_{k+1} positiv definit und es folgt (a).

Da λ kleinster Eigenwert von A_{k+1} ist, gilt:

$$\frac{1}{\lambda} = \|A_{k+1}^{-1}\|$$

Daraus folgt

$$\|A_{k+1}^{-1}\| = \frac{1}{\lambda} \leqslant R^{-2}4^{k+1}$$

(1. Teil von (d))

Wegen $\|\tilde{A}_{k+1}^{-1}\| \leqslant 3\|A_k^{-1}\|$ folgt $\|\tilde{A}_{k+1}^{-1}\| \leqslant R^2 4^{k+1}$ (d).

Es gilt $\|\tilde{A}_{k+1}\| \leqslant \frac{n^2}{n^2-1}\mu\|A_k\|$, denn $\|A\| \leqslant \|A+B\|$ für positiv semidefinite Matrizen. Ferner ist $\|A_0\| = R^2$ und daher

$$\|A_{k+1}\| = \|\tilde{A}_{k+1}\| + \|\Gamma_{k+1}\| \leqslant \frac{n^2}{n^2-1}\mu\|A_k\| + n\delta \leqslant R^2 2^{k+1}$$

Somit ist $\|\tilde{A}_{k+1}\| \leqslant \frac{n^2}{n-1}\mu\|A_n\| \leqslant R^2 2^{k+1}$ und (c) folgt.

Schreibe $A_k = MM^t$ mit M regulär. Dann ist

$$\|b_k\| = \frac{A_k\overline{a}}{\sqrt{\overline{a}^t A_k \overline{a}}} = \sqrt{\frac{\overline{a}^t A_k^t A_k \overline{a}}{\overline{a}^t A_k \overline{a}}} = \frac{(\mu^t\overline{a})^t A_k (\mu^t \overline{a})}{(\mu^t\overline{a}^t)(\mu^t\overline{a})} \leqslant \sqrt{\|A_k\|} \leqslant R \cdot 2^{\frac{k}{2}}$$

und $p_0 = 0$.

Es ist $\|p_{k+1}\| = \|p_k\| + \frac{1}{n+1}\|b_k\| + \sqrt{n}\delta = \|p_k\| + R \cdot 2^{\frac{k}{2}} + \sqrt{n}\delta \leqslant R \cdot 2^{k+1}$.

Ebenso folgt $\sqrt{\tilde{p}_{k+1}} \leqslant R \cdot 2^{k+1}$ und es folgt (b). □

Lemma 108. Sei $0 < \delta < \left(2^{6N(R,\varepsilon)+1)} 16n^3\right)^{-1}$, wobei

$$N(R, \varepsilon) := \left\lceil 7(n+1)(n \ln(2R) + \ln(\frac{1}{\varepsilon}) \right\rceil.$$

Dann ist in der k-ten Iteration des Ellipsoid Algorithmus $K \subseteq p_k + E_k$ und $\text{vol}(E_k) < e^{-\frac{k}{8(n+1)}} 2^n R^n$

TODO: Hier fehlt etwas

8.3 Die Ellipsoid-Methode für LPs

Wir wollen den Ellipsoid-Algorithmus verwenden, um zu überprüfen, ob ein Polyeder P leer ist. Hierfür müssen wir sicherstellen, dass P ein Polytop ist und, falls $P \neq \emptyset$, das Volumen nicht beliebig klein sein kann.

Theorem 109. Sei $A \in \mathbb{Q}^{m \times n}$ und $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$. Für $R = 1 + 2^{4n(\text{size}(A)+\text{size}(B))}$ und $\varepsilon = \left(2n2^{4n(\text{size}(A)+\text{size}(b))}\right)^{-1}$ sei

$$P_{R,\varepsilon} := \{x \in [-R, R]^n | Ax \leqslant b + \varepsilon \mathbb{1}\}$$

. Dann gilt:

(a) $P = \emptyset \iff P_{R,\varepsilon = \emptyset}$

(b) Falls $P \neq \emptyset$, dann $\text{vol}(P_{R,\varepsilon}) \geqslant \left(\frac{2\varepsilon}{n^{2\,\text{size}(A)}}\right)^n$.

Beweis. (a) $P_{R,\varepsilon} = \emptyset \implies P_{R,0} = \emptyset$ ist trivial. Nach Theorem 94 gilt $P_{R,0} = \emptyset \implies P = \emptyset$.

Behauptung 1. $P = \emptyset \implies P_{R,\varepsilon} = \emptyset$.

Unterbeweis. Angenommen $P = \emptyset$. Nach dem Lemma von Farkas (Theorem 23) existiert dann ein $y \geqslant 0$, so dass $y^t A = 0$ und $y^t b = -1$. Nach Theorem 94 hat

$$\begin{aligned} \min \quad & \mathbb{1}^t y \\ \text{s.d.} \quad & A^t y = 0 \\ & b^t y = -1 \\ & y \geqslant 0 \end{aligned}$$

eine Optimallösung y^*, so dass $|y_i^*| \leqslant 2^{4n(\text{size}(A)+\text{size}(b))}$ für alle i. Folglich ist $y^t(b + \varepsilon\mathbb{1}) < -1 + (n+1)2^{4n(\text{size}(A)+\text{size}(b))}\varepsilon < 0$.

Nach dem Lemma von Farkas hat daher $Ax \leqslant b + \varepsilon\mathbb{1}$ keine zulässige Lösung. Insbesondere existiert keine Lösung in $[-R, R]^n$, folglich ist $P_{R,\varepsilon} = \emptyset$. ■

(b) Falls $P \neq \emptyset$, dann ist $P_{R-1,0} \neq \emptyset$. Für jedes $z \in P_{R-1,0}$ gilt

$$\{x \in \mathbb{R}^n | \|x - z\|_\infty < \frac{\varepsilon}{n \cdot 2^{\text{size}(A)}}\} \subseteq P_{R,\varepsilon}$$

Folglich ist $\text{vol}(P_{R,\varepsilon}) \geqslant \text{vol}\left(\{x \in \mathbb{R}^n | \|x - z\|_\infty \leqslant \frac{\varepsilon}{n 2^{\text{size}(A)}}\}\right) = \left(\frac{2\varepsilon}{n 2^{\varepsilon\,\text{size}(A)}}\right)$.

□

Theorem 110. Zu einem gegebenen Polyeder $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ mit $A \in \mathbb{Q}^{m \times n}$ und $b \in \mathbb{Q}^m$ kann man in polynomieller Zeit entscheiden, ob P leer ist.

Beweis. Setze $R' = 1 + 2^{4n(\text{size}(A)+\text{size}(b))}$, $\varepsilon = \left(2n \cdot 2^{4n(\text{size}(A)+\text{size}(b))}\right)^{-1}$ und $\varepsilon' = \left(\frac{2\varepsilon}{n 2^{\text{size}(A)}}\right)^n$. Wende die Ellipsoid-Methode mit $R = \lceil \sqrt{n} R' \rceil$ und ε' als Schranke an das Volumen an, um zu entscheiden, ob $K = P_{R',\varepsilon}$ leer ist. Damit wird auch überprüft, ob P leer ist

Die Anzahl der Iterationen ist polynomiell beschränkt. Es reicht den absoluten Rundungsfehler durch ein $\delta < \left(2^{6\left(N(R,\varepsilon')+1\right)} 16 n^3\right)^{-1}$ zu beschränken. Folglich handelt es sich um ein polynomielles Verfahren. □

Theorem 111. Es gibt einen polynomiellen Algorithmus, der zu einem gegebenen LP $\max\{c^t x | Ax \leqslant b\}$ mit $A \in \mathbb{Q}^{m \times n}$, $c \in \mathbb{Q}^m$, $b \in \mathbb{Q}^n$ eine Optimallösung findet, falls eine solche existiert.

Beweis. Seien $a_i^t x \leqslant b_i$ die Ungleichungen aus $Ax \leqslant b$.

Überprüfe zunächst ob $Ax \leqslant b$ zulässig ist. Falls nicht, sind wir fertig.

Sonst führe folgende Schritte durch:

Ersetze $a_i^t x \leqslant b$ durch $a_i^t = b$. Bleibt das System zulässig, so erhalte die Gleichung, anderenfalls entferne die Gleichung. Dies liefert ein zulässiges Gleichungssystem, dessen Lösungen auch Lösungen von $Ax \leqslant b$ sind. Dieses kann mit dem Gauß-Verfahren gelöst werden.

Durch die Kombination aus primalen und dualen Systems kann sichergestellt werden, dass es sich um eine Optimallösung handelt:

Betrachte das System

$$\begin{aligned} Ax &\leqslant b \\ y^t A &= c \\ y &\geqslant 0 \\ c^t x &\geqslant b^t y \end{aligned}$$

Jede Lösung des erweiterten Systems enthält einen Teilvektor x, der einer optimale primale Lösung ist. Daher genügt es, das Verfahren auf das erweiterte System anzuwenden. □

Remark 111.1. Das Verfahren berechnet stets eine Lösung aus einer minimalen Fläche, insbesondere also eine Ecke, falls es sich um ein spitzes Polyeder handelt (Theorem 44).

Die Laufzeit ist durch $\mathcal{O}((m+n)^7(\mathrm{size}(A)+\mathrm{size}(b)+\mathrm{size}(c))^4)$ beschränkt.

8.4 Separation und Optimierung

Problem. *In manchen Fällen ist ein LP durch exponentiell viele Nebenbedingungen gegeben.*

Example 112. Betrachte das Matching-Problem: Gegeben sei ein ungerichteter Graph G, gesucht ist eine möglichst große Menge $M \subseteq E(G)$ mit $|\delta_G(v) \cap M| \leqslant 1|$ für alle $v \in V(G)$.

ILP-Formulierung:

$$\begin{aligned} \max \sum_{e \in E(G)} x_e & & \\ \text{s.d.} \sum_{e \in \delta_G(v)} x_e &\leqslant 1 & \forall v \in V(G) \\ x_e &\in \{0,1\} & \forall e \in E(G) \end{aligned}$$

Die Relaxierung hiervon ist allerdings nicht sonderlich hilfreich. (Betrachte z.B. die fraktionale Lösung, die für ein Dreieck für jede Kante $x_e = \frac{1}{2}$ setzt.) Bei der LP-Relaxierung darf man folgende Nebenbedingungen einfügen:

$$\sum_{e \in E(G[U])} x_e \leqslant \frac{|U|-1}{2} \forall U \subseteq V(H), |U| \text{ungerade}$$

Hierdurch werden einige uninteressante Lösungen des relaxierten LP entfernt. Allerdings führt dies zu einer exponentiellen Anzahl von Nebenbedingungen. Hierfür kann jedoch ein Separationsorakel angegeben werden.

Betrachte ab jetzt abgeschlossene konvexe Mengen K, für die es Zahlen r und R it $0 < r < \frac{R}{2}$ gibt, so dass $rB^n \subseteq K \subseteq RB^n$. Solche Mengen heißen **r-R-sandwiched Mengen**.

Wir betrachten das **schwache Optimierungsproblem**: Zu einer gegebene Menge $K \subseteq \mathbb{R}^n$, $\varepsilon > 0$ und einem Vektor $c \in \mathbb{Q}^n$ suchen wir ein $x \in K$ mit $c^t x \geqslant \max\{c^t z | z \in K\} - \varepsilon$.

Lemma 113. Sei $K \subseteq \mathbb{R}^n$ eine r-R-sandwiched konvexe Menge, $c \in \mathbb{R}^n$, $\delta = \sup\{c^t x | x \in K\}$ und $0 < \varepsilon < \delta$. Außerdem sei $U = \{x \in K | c^t x \geqslant \delta - \varepsilon\}$. Dann gilt:

$$\operatorname{vol}(U) \geqslant \left(\frac{\varepsilon}{2\|c\|R}\right)^{n-1} r^{n-1} \frac{1}{n^n} \frac{\varepsilon}{2\|c\|} \frac{1}{n}$$

Beweis. Übung □

Theorem 114. Gegeben sei ein Separationsorakel für eine r-R-sandwiched konvexe Menge $K \subseteq \mathbb{R}^n$ mit Laufzeit polynomiell in $\operatorname{size}(R)$, $\operatorname{size}(r)$ und $\operatorname{size}(x)$ (wobei x der Eingabevektor für das Orakel sei), eine Zahl $\varepsilon > 0$ und einen Vektor c. Dann gibt es einen polynomiellen Algorithmus bezüglich $\operatorname{size}(R), \operatorname{size}(r), \operatorname{size}(c), \operatorname{size}(\varepsilon)$, der einen Vektor $v \in K$ mit $c^t v \geqslant \sup\{c^t x | x \in K\} - \varepsilon$ berechnet.

Beweis. Sei $\delta = \sup\{c^t x | x \in K\}$. Wende die Ellipsoid-Methode auf die Menge aller fast-optimalen Elemente von K an, also auf $\{x \in K | c^t x \geqslant \delta - \varepsilon\}$. Nach dem vorigen Lemma kann das Volumen dieser Menge nicht beliebig klein sein. (Solange wir Punkte in K finden, die nicht fast-optimal sind (erkennbar daran, dass das Volumen des Ellipsoids noch zu groß ist), kann c als Normalenvektor für die trennende Hyperebene verwendet werden). □

Definition 115 (Schwaches Separationsorakel)**.** Ein **schwaches Separationsorakel** für eine konvexe Menge $K \subseteq \mathbb{R}^n$ ist ein Algorithmus, der zu gegebenem $x \in \mathbb{R}^n$ und η mit $0 < \eta < \frac{1}{2}$ entweder "$x \in K$" ausgibt oder einen Vektor $v \in \mathbb{R}^n$ findet mit $v^t z \leqslant 1$ für alle $z \in K$ und $v^t x \geqslant 1 - \eta$.

Remark 115.1. In Theorem 114 genügt ein schwaches Separationsorakel an Stelle eines Separationsorakels.

Notation 115.2. Für $K \subseteq \mathbb{R}^n$ sei $K^* := \{y \in \mathbb{R}^n | y^t x \leqslant 1 \forall x \in K\}$.

Lemma 116. Sei $0 \in K \subseteq \mathbb{R}^n$ konvex und abgeschlossen. Dann ist $K^{**} = K$.

Beweis. "$\supseteq$" : Für $x \in K$ gilt $y^t x \leqslant 1$ für alle $y \in K^*$. Folglich gilt $x \in K^{**}$.

"$\subseteq$" : Sei $z \in \mathbb{R}^n \backslash K$. Sei $w \in K$ ein Vektor, sodass $\|z - w\|$ minimal unter allen Vektoren in K ist. Sei $u := z - w$. Dann gilt $u^t x \leqslant u^t w < u^t z$ für alle $y \in K$.

Wegen $0 \in K$ gilt $u^t w \geqslant 0$. Nach Skalierung können wir annehmen, dass $u^t z > 1$ und $u^t x \leqslant 1$ für ale $x \in K$. Es folgt $u \in K^*$ und $u^t z > 1$, somit ist $z \notin K^{**}$. □

Theorem 117. Wenn es einen Algorithmus mit Laufzeit polynomiell in size(r) und size(R) gibt, der lineare Zielfunktionen über einer abgeschlossenen konvexen r-R-sandwiched Menge $K \subseteq \mathbb{R}^n$ maximiert, dann gibt es ein schwaches Separationsorakel für K, dessen Laufzeit polynomiell in $\text{size}(r), \text{size}(R)$ und size(η) ist.

Beweis. Sei $x \in \mathbb{R}^n$ eine Instanz für das schwache Separationsorakel. Falls $x = 0$, dann ist $x \in K$. Also sei $x \neq 0$. Falls $\|x\| > R$, dann wähle $v = \frac{x}{\|x\|}$.

Sei nun $0 < \|x\| < R$.

Behauptung 1. *Wir können das (starke) Separationsproblem für K^* lösen.*

Unterbeweis. Sei $y \in \mathbb{R}^n$. Betrachte das Problem $\max\{y^t x | x \in K\}$. Sei x^* eine Optimallösung. Falls $y^t x^* \leqslant 1$, so ist $y \in K^*$. Anderenfalls ist $K^* \subseteq \{x \in \mathbb{R}^n | {x^*}^t x < {x^*}^t y\}$. ■

K^* ist eine abgeschlossene $\frac{1}{R} - \frac{1}{r}$-sandwiched Menge. Somit können wir das schwache Optimierungsproblem für K^* mit $c = \frac{x}{\|x\|}$ und $\varepsilon = \frac{\eta}{R}$ polynomiell lösen. Wir erhalten so einen Vektor $v_0 \in K^*$ mit $\frac{x^t}{\|x\|} v_0 \geqslant \max\{\frac{x^t}{\|x\|} v | v \in K^*\} - \frac{\eta}{R}$. Falls $\frac{x^t}{\|x\|} v_0 \geqslant \frac{1}{\|x\|} - \frac{\eta}{R}$, dann ist $v_0^t x \geqslant 1 - \frac{\eta}{R}\|x\| \geqslant 1 - \eta$ und $v_0^t z \leqslant 1$ für alle $z \in K$ (wegen $v_0 \in K^*$).

Anderenfalls gilt $\max\{\frac{x^t}{\|x\|} v | v \in K^*\} \leqslant \frac{1}{\|x\|}$. Folglich ist $\max\{x^t v | v \in K^*\} \leqslant 1$ und somit $x \in K^{**} \overset{\text{Lemma 116}}{=} K$. □

Theorem 118. Sei $n \in \mathbb{N}$ und $c \in \mathbb{Q}^n$. Sei $P \subseteq \mathbb{R}^n$ ein rationales Polytop und sei $x_0 \in P$ ein Vektor im Inneren von P. Sei $T \in \mathbb{N} \backslash \{0\}$ mit $\text{size}(x_0) \leqslant \log(T)$ und $\text{size}(x) \leqslant \log(T)$ für alle Ecken x von P.

Zu gegebenen n,c, x_0, T und einem polynomiellen Separationsorakel für P kann eine Ecke x^* von P, in der das Maximum $\max\{c^t x | x \in P\}$ angenommen wird, in einer Laufzeit gefunden werden, die polynomiell in n, $\log(T)$ und $\text{size}(c)$ ist.

Beweis. Siehe Korte und Vygen [2018] □

Es gilt auch die Umkehrung:

Theorem 119. Sei $n \in \mathbb{N}$ und $c \in \mathbb{Q}^n$. Sei $P \subseteq \mathbb{R}^n$ ein rationales Polytop und sei $x_0 \in P$ ein Vektor im Inneren von P. Sei $T \in \mathbb{N}\backslash\{0\}$, sodass $\text{size}(x_0) \leqslant \log(T)$ und $\text{size}(x) \leqslant \log(T)$ für alle Ecken x von P gilt.

Zu gegebenen n, y, x_0, T und einem Orakel, welches für ein gegebenes $c \in \mathbb{Q}^n$ eine Ecke $x^* \in \text{argmax}\{c^t x | x \in P\}$ von P zurückgibt, kann ein Separationsorakel für P und y mit einer Laufzeit, die polynomiell in n, $\log(T)$ und $\text{size}(y)$ ist, implementiert werden.

Falls $y \notin P$ können wir in dieser Laufzeit eine facettenbestimmende Ungleichung für P finden, die von y verletzt wird.

Beweis. Siehe Korte und Vygen [2018] □

9 Innere Punkte Methoden

Die Ellipsoid-Methode ist lediglich von theoretischem Interesse; die im Folgenden beschriebene Innere-Punkte-Methode ist hingegen in Theorie und Praxis relevant.

Es gibt verschiedene Verfahren, die als "Innere Punkte Methode" bezeichnet werden. Der folgende Teil orientiert sich an der Darstellung von Mehlhorn und Saxena (2015) eines Algorithmus von Karmakar (1984).

Wir betrachten ein LP $\max\{c^t x | Ax \leqslant b\}$. Um die Notation zu vereinfachen fügen wir Slack-Variablen s ein und es ergibt sich

$$\begin{aligned} \max \quad & c^t x && (14)\\ \text{s.d.} \quad & Ax + s = b && (15)\\ & s \geqslant 0 && (16) \end{aligned}$$

Das duale LP ist

$$\begin{aligned} \min \quad & b^t y && (17)\\ \text{s.d.} \quad & A^t y = y && (18)\\ & y \geqslant 0 && (19)\\ & && (20) \end{aligned}$$

Im Folgenden nehmen wir an, dass die Spalten von A linear unabhängig sind (anderenfalls gibt es redundante Ungleichungen) und dass die Zahl der Zeilen größer ist als die Zahl der Spalten (anderenfalls löse das Gleichungssystem und prüfe ob die Lösung positiv ist).[4]

Auf Grund des komplementären Schlupfes (Theorem 28) haben wir eine Optimallösung für beide Probleme gefunden, wenn wir Lösungen x, s und y finden, so dass $y^t s = 0$. Wir wollen also eine Lösung für das folgende Problem finden:

$$Ax + s = b \tag{21}$$
$$A^t y = c \tag{22}$$
$$y^t s = 0 \tag{23}$$
$$y \geqslant 0 \tag{24}$$
$$s \geqslant 0 \tag{25}$$
$$\tag{26}$$

$y^t s = 0$ ist dabei eine nichtlineare Bedingung. Ohne diese Bedingung ist $y^t s$ der Abstand zwischen dem (dualen) Wert der dualen Lösung und dem Wert der primalen Lösung, denn

$$b^t y - c^t x = x^t A^t y + s^t y - c^t x = x^t c + s^t y - c^t x = s^t y$$

(21) hat genau dann eine Lösung, wenn sowohl das primale als auch das duale LP zulässig und beschränkt sind. Dies werden wir zunächst annehmen und später sehen, wie diese Eigenschaft erzwungen werden kann.

Innere-Punkte-Methoden betrachten Vektoren im Inneren des Lösungsraumes. In (21) sind $y \geqslant 0$ und $s \geqslant 0$ die einzigen Ungleichungen, daher werden wir Lösungen x, s, y mit $y > 0, s > 0$ betrachten.

Die Bedingung $y^t s = 0$ ersetzen wir druch $\sigma^2 := \sum_{i=1}^m \left(\frac{y_i s_i}{\mu} - 1\right)^2 \leqslant \frac{1}{4}$ für ein $\mu > 0$. μ werden wir dabei nach und nach gegen 0 laufen lassen.

$$Ax + s = b \tag{27}$$
$$A^t y = c \tag{28}$$
$$\sum_{i=1}^m \left(\frac{y_i s_i}{\mu} - 1\right)^2 \leqslant \frac{1}{4} \tag{29}$$
$$y > 0 \tag{30}$$
$$s > 0 \tag{31}$$

Die grundlegende Strategie besteht aus drei Teilen

[4] Das sind die selben Annahmen wie für den Simplex-Algorithmus, allerdings für die transponierte Matrix.

(I) Berechne eine initiale Lösung. (Unterabschnitt 9.1)

(II) Reduziere μ um einen konstanten Faktor und passe x, y, s so an, dass zu diesem neuen μ passen. Wiederhole dies bis μ klein genug ist.

(III) Berechne eine Optimallösung.

9.1 Modifikation des LP und Berechnen einer initialen Lösung

Wir wollen (27) so ändern, dass wir einfach eine Startlösung berechnen können. Insbesondere werden hierbei das primale und das duale LP zulässig gemacht.

Zunächst sorgen wir dafür, dass das duale LP (17) beschränkt ist. Nach Theorem 94 existiert, falls (17) beschränkt ist, ein $W \in 2^{\Theta(m(\mathrm{size}(A)+\mathrm{size}(c)))}$, so dass eine Optimallösung $y \geqslant 0$ existiert mit $y \leqslant W \cdot \mathbb{1}$.

Äquivalent dazu können wir nach einem Vektor $y \geqslant 0$ mit $\mathbb{1}^t y \leqslant m$ und $A^t y = \frac{1}{W} c$ suchen (teile alles durch W). Indem wir $\mathbb{1}^t y \leqslant m$ durch $\mathbb{1}^t y \leqslant m+1$ ersetzen und eine Slack-Variable $y_{m+1} \geqslant 0$ hinzufügen, erhalten wir das folgende LP, welches zu (17) äquivalent ist, vorausgesetzt (17) ist beschränkt:

$$\begin{aligned}
\min \quad & b^t y && (32)\\
\text{s.d.} \quad & A^t y = \frac{1}{W} c && (33)\\
& \mathbb{1}^t y + y_{m+1} = m+1 && (34)\\
& y \geqslant 0 && (35)\\
& y_{m+1} \geqslant 0 && (36)
\end{aligned}$$

Im zweiten Schritt machen wir dieses LP zulässig: Füge eine neue Variable y_{m+2} ein, sodass das $\mathbb{1}$ eine zulässige Lösung wird. Sei H eine Konstante, deren Wert noch zu bestimmen ist.

$$\begin{aligned}
\min \quad & b^t y + H y_{m+2} && (37)\\
\text{s.d.} \quad & A^t y + \left(\frac{1}{W} c - A^t \mathbb{1}\right) y_{m+2} = \frac{1}{W} c && (38)\\
& \mathbb{1}^t y + y_{m+1} \geqslant 0 && (39)\\
& y \geqslant 0 && (40)\\
& y_{m+1} \geqslant 0 && (41)\\
& y_{m+2} \geqslant 0 && (42)
\end{aligned}$$

H soll dabei so groß gewählt werden, dass, falls eine Lösung mit $y_{m+2} = 0$ existiert, bereits $y_{m+2} = 0$ in jeder Optimallösung gelten muss.

Nach Corollary 95 wissen wir, dass eine Konstante l existiert, so dass, falls $y_{m+2} > 0$ in einer Optimallösung gilt, auch eine Optimallösung mit $y_{m+2} \geqslant 2^{-4ml(\text{size}(A)+\text{size}(c)+\text{size}(W))}$ existiert. Andererseits ist $b^t y \leqslant \|b\|_1(m+2)$, daher ist $H := (\|b\|_1(m+2)+1) \cdot 2^{4ml(\text{size}(A)+\text{size}(c)+\text{size}(W))}$ eine geeignete Wahl.

(37) ist offensichtlich beschränkt und zulässig. Ferner kann aus einer Optimallösung eine Optimallösung von (17) rekonstruiert werden: Sei $y_1, \ldots, y_{m+2}$ eine solche Optimallösung.

- $y_{m+2} > 0$: (17) ist nicht zulässig.
- $y_{m+2} = 0$: (17) ist zulässig.
  - $y_{m+1} > 0$: (17) ist beschränkt.
  - $y_{m+1} = 0$: (17) kann beschränkt oder unbeschränkt sein. Ersetze c durch den Nullvektor und löse das entstehende Problem. Nach dem Lemma von Farkas ist (17) genau dann beschränkt, wenn der Wert einer optimalen Lösung des neuen LPs nicht-negativ ist.

Dualisieren von (37) ergibt:

$$\begin{aligned}
\max \quad & \frac{1}{W}c^t x + (m+2)x_{n+1} && (43)\\
\text{s.d.} \quad & Ax + x_{n+1}\mathbb{1} + s = b && (44)\\
& (\frac{1}{W}c^t - \mathbb{1}^t A)x + x_{n+1} + s_{m+2} = H && (45)\\
& x_{n+1} + s_{m+1} = 0 && (46)\\
& s \geqslant 0 && (47)\\
& s_{m+1} \geqslant 0 && (48)\\
& s_{m+2} \geqslant 0 && (49)\\
& && (50)
\end{aligned}$$

Statt des primal-dualen Paares (14) (17) betrachten wir nun das Paar (37) (43). Offenbar sind beide LPs zulässig und beschränkt. Eine Startlösung finden wir durch $y = \mathbb{1}$ und $x = 0$. Ferner wählen wir $s_{m+1} = \frac{\mu}{y_{m+1}} = \mu$. Dann muss $x_{n+1} = -\mu, s_{m+2} = H + \mu$ und $s_i = b_i - x_{n+1} = b_i + \mu$ sein.

Es ergibt sich

$$\begin{aligned}
\frac{y_i s_i}{\mu} - 1 &= \frac{b_i}{y} \quad i = 1, \ldots, m\\
\frac{y_{m+1}s_{m+1}}{\mu} - 1 &= 0\\
\frac{y_{m+2}s_{m+2}}{\mu} - 1 &= \frac{H}{\mu}
\end{aligned}$$

Folglich ist

$$\sigma^2 = \sum_{i=1}^{m+2} \left(\frac{y_i s_i}{\mu} - 1 \right)^2 = \frac{1}{\mu^2} \left(H^2 + \sum_{i=1}^{m} b_i^2 \right)$$

Indem wir $\mu = 2\sqrt{H^2 + \sum_{i=1}^m b_i^2}$ wählt, wird somit $\sigma^2 \leqslant \frac{1}{4}$ sicher gestellt. Wegen $\mu > |b_i|$ folgt ferner $s_i = b_i + \mu > 0$ für $i = 1, \ldots, m$.

Im Folgenden schreiben wir (37) als

$$\begin{aligned} \min \quad & \tilde{b}^t y \\ \text{s.d.} \quad & \tilde{A}^t y = \tilde{c} \\ & y \geqslant 0 \end{aligned}$$

und (43) als

$$\begin{aligned} \max \quad & \tilde{c}^t x \\ \text{s.d.} \quad & \tilde{A}x + s = \tilde{b} \\ & s \geqslant 0 \end{aligned}$$

Wir haben somit eine Startlösung $\mu^{(0)}, x^{(0)}, y^{(0)}, s^{(0)}$ gefunden für

$$\tilde{A}x + s = \tilde{b} \tag{51}$$

$$\tilde{A}^t y = \tilde{c} \tag{52}$$

$$\sum_{i=1}^{m+2} \left(\frac{y_i s_i}{\mu} - 1 \right)^2 \leqslant \frac{1}{4} \tag{53}$$

$$y > 0 \tag{54}$$

$$s > 0 \tag{55}$$

$$\tag{56}$$

9.2 Reduktion von μ

Gegeben eine Lösung $\mu^{(k)}, x^{(k)}, y^{(k)}, s^{(k)}$ für (51) wollen wir nun eine Lösung $\mu^{(k+1)}$, $x^{(k+1)}$, $y^{(k+1)}$, $s^{(k+1)}$ finden, wobei $\mu^{(k+1)} = (1-\delta)\mu^{(k)}$ für eine δ, welches nicht von der Lösung abhängt.

Hierbei werden wir zunächst annehmen, dass wir exakt rechnen können und später zeigen, wie Zwischenergebnisse geeignet gerundet werden können.

Schreibe $x^{(k+1)} = x^{(k)} + f$, $y^{(k+1)} = y^{(k)} + g$ und $s^{(k+1)} = s^{(k)} + h$. f, g,h seien dabei relativ klein.

Für ein festes $\mu^{(k+1)}$ können f, g, h wie folgt bestimmt werden:

Aus (51) folgt $\tilde{A}f + h = 0$ und $\tilde{A}^t g = 0$. Ferner wollen wir f und h so wählen, dass $(y_i^{(k)} + g_i)(s_i^{(k)} + h_i)$ nache an $\mu^{(k+1)}$ liegt. Da $(y_i^{(k)} + g_i)(s_i^{(k)} + h_i) =$

$y_i^{(k)} s_i^{(k)} + g_i s_i^{(k)} + y_i^{(k)} h_i + g_i h_i$ und das Produkt $g_i h_i$ klein ist, fordern wir nur $y_i^{(k)} s_i^{(k)} + g_i s_i^{(k)} + y_i^{(k)} h_i = \mu^{(k+1)}$.

Wir suchen also f, g, h so dass

$$\begin{aligned} \tilde{A}^t g &= 0 \\ \tilde{A} f + h &= 0 \\ s_i^{(k)} g_i + y_i^{(k)} h_i = \mu^{(k+1)} - y_i^{(k)} s_i^{(k)} i &= 1, \dots, m+2 \end{aligned}$$

$y^{(k)}$ und $s^{(k)}$ sind in diesem Kontext konstant. $y^{(k+1)} > 0$ und $s^{(k+1)} > 0$ wurden nicht aufgenomen, wir werden allerdings sehen, dass wir trotzdem positive Werte erhalten.

Sei f, g, h eine Lösung von (57). Nach Konstruktion ist dann

$$(y^{(k)} + g)^t (s^{(k)} + h) = (m+2)\mu^{(k+1)} + g^t h$$

sowie

$$g^t h = -g^t \tilde{A} f = 0^t f = 0$$

Es folgt

$$\begin{aligned} \tilde{b}^t y^{(k+1)} &= \left(\tilde{A} \left(x^{(k)} + f \right) + (s^{(k)} + h) \right)^t (y^{(k)} + g) - \tilde{c}^t (x^{(k)} + f) \\ &= \left(\tilde{A}(x^{(k)} + f) \right)^t (y^{(k)} + g) + (m+2)\mu^{(k+1)} - \tilde{c}^t (x^{(k)} + f) \\ &= \left(x^{(k)} + f \right)^t \tilde{A}^t y^{(k)} + (m+2)\mu^{(k+1)} - \tilde{c}^t (x^{(k)} + f) \\ &= (m+2)\mu^{(k+1)} \end{aligned}$$

Lemma 120. (57) hat eine eindeutige Lösung.

Es ist $\sigma^{(k)} = \sqrt{\sum_{i=1}^{m+2} \left(\frac{y_i^{(k)} s_i^{(k)}}{\mu^{(k)}} - 1 \right)^2}$ und $\sigma^{(k+1)} = \sqrt{\sum_{i=1}^{m+2} \left(\frac{y_i^{(k+1)} s_i^{(k+1)}}{\mu^{(k+1)}} - 1 \right)^2} = \sqrt{\sum_{i=1}^{m+2} \left(\frac{g_i h_i}{\mu^{(k+1)}} \right)^2}$.

Zu zeigen ist noch, dass $y^{(k+1)} > 0, s^{(k+1)} > 0$ und $\sigma^{(k+1)} \leqslant \frac{1}{2}$.

Wir zeigen zunächst, dass $\sigma^{(k+1)} \leqslant \frac{1}{2}$ für eine geeignete Wahl von $\mu^{(k+1)}$.

Lemma 121. (a) Für $i = 1, \dots, m+2$ gilt $\frac{\mu^{(k)}}{y_i^{(k)} s_i^{(k)}} \leqslant \frac{1}{1-\sigma^{(k)}}$.

(b) $\sum_{i=1}^{m+2}\left|1-\frac{y_i^{(k)}s_i^{(k)}}{\mu^{(k)}}\right| \leqslant \sigma^{(k)}\sqrt{m+2}$.

Lemma 122. Falls $\delta = \frac{1}{8\sqrt{m+2}}$ (d.h. $\mu^{(k+1)} = (1-\frac{1}{8\sqrt{m+2}})\mu^{(k)}$), dann gilt $\sigma^{(k+1)} < \frac{1}{2}$.

Lemma 123. $y^{(k+1)} > 0$ und $s^{(k+1)} > 0$.

Beweis.

Behauptung 1. $y_i^{(k+1)}s_i^{(k+1)} > 0$ *für* $i = 1,\ldots,m+2$.

Unterbeweis. Angenommen $y_j^{(k+1)}s_j^{(k+1)} \leqslant 0$ für ein $j \in \{1,\ldots,m+2\}$. Dann ist

$$\left(\sigma^{(k+1)}\right)^1 = \sum_{i=1}^{m+2}\left(\frac{y_i^{(k+1)}s_i^{(k+1)}}{\mu_i^{(k+1)}}-1\right)^2 \geqslant \left(\frac{y_j^{(k+1)}s_j^{(k+1)}}{\mu_j^{(k+1)}}-1\right)^2 \geqslant 1$$

im Widerspruch zu Lemma 122 ■

Daher sind $y_i^{(k+1)}$ und $s_i^{(k+1)}$ entweder beide positiv oder beide nicht-positiv. Angenommen $y_i^{(k+1)} = y_i^{(k)} + g_i \leqslant 0, s_i^{(k+1)} = s_i^{(k)} + h_i \leqslant 0$.

Dann folgt wegen $s_i^{(k)} > 0$ und $y_i^{(k)} > 0$, dass

$$0 \geqslant s_i^{(k)}\left(y_i^{(k)}+g_i\right) + y_i^{(k)}\left(s_i^{(k)}+h_i\right) = s_i^{(k)}y_i^{(k)} + \mu^{(k+1)}$$

Jedoch sind $s_i^{(k)}, y_i^{(k)}$ und $\mu^{(k+1)}$ positiv $\lightning$ □

9.2.1 Runden von Zwischenergebnissen

Es muss verhindert werden, dass die Größe der Darstellungen von f, g und h in den Iterationen zu stark steigt. Daher speichern wir statt exakten Werten $y_i^{(k)}$ und $s_i^{(k)}$ gerundete Werte $\tilde{y}_i^{(k)}$ und $\tilde{s}_i^{(k)}$, sodass $|\frac{\tilde{y}_i^{(k)}\tilde{s}_i^{(k)}-y_i^{(k)}s_i^{(k)}}{\mu^{(k)}}| \leqslant \varepsilon$ für ein $0 < \varepsilon < \frac{1}{m+2}\frac{1}{300}$. Indem wir den Lösungraum auf ein Polytop einschränken, können wir annehmen, dass polynomieller Speicher genügt, um diese gerundeten Werte zu speichern.

Wir erhalten

$$\begin{aligned}\sum_{i=1}^{m+2}\left(1-\frac{\tilde{y}_i^{(k)}\tilde{s}_i^{(k)}}{\mu^{(k)}}\right)^2 &\leqslant \sum_{i=1}^{m+2}\left(1-\frac{y_i^{(k)}s_i^{(k)}}{\mu^{(k)}}\right)^2 + \sum_{i=1}^{m+2}2\varepsilon\left(1-\frac{y_i^{(k)}s_i^{(k)}}{\mu^{(k)}}\right) + (m+2)\varepsilon^2 \\ &\leqslant \sum_{i=1}^{m+2}\left(1-\frac{y_i^{(k)}s_i^{(k)}}{\mu^{(k)}}\right)^2 + \frac{1}{100}\end{aligned}$$

Daher genügt eine leicht verbeserte Schranke für $\sum_{i=1}^{m+2}\left(1-\frac{y_i^{(k)}s_i^{(k)}}{\mu^{(k)}}\right)^2$. Für die Startlösung kann das leicht sichergestellt werden (verinngere $\mu^{(0)}$ ein wenig). Beim Berechnen einer neuen Lösung stellt dies ebenfalls kein Problem dar, da da Ungleichung in Lemma 122 leicht verschärft werden kann.

9.3 Finden einer optimalen Lösung

Im Folgenden wird beschrieben, wie sich eine optimale Lösung zu (51) finden lässt. Sei dabei y^* eine Optimallösung von (51) und x^*, s^* eine Optimallösung von (51).

Nach Corollary 95 können wir annehmen, dass alle positiven Einträge von y^* und s^* mindestens Größe η für ein $\eta = 2^{-\Theta(\text{size}(\tilde{A})+\text{size}(\tilde{b})+\text{size}(\tilde{c}))}$ haben.

Lemma 124. Sei μ, x, y, s eine Lösung von (51), $i \in \{1, \dots, m+2\}$. Dann gilt

(a) $y_i < \frac{\eta}{4(m+2)} \implies y_i^* = 0$

(b) $s_i < \frac{\eta}{4(m+2)} \implies s_i^* = 0$

Es gibt mehrere Methoden, um einen inneren Punkt zu einer optimalen Lösung zu runden.

9.3.1 Einfache Methode

Sei k groß genug, dass $\mu^{(k)} < \frac{\eta^2}{32(m+2)^2}$. Dann ist für jedes i bereits $y_i^{(k)} < \frac{\eta}{4(m+2)}$ oder $s_i^{(k)} < \frac{\eta}{4(m+2)}$.

Sei $\overline{A}^t y = \overline{c}$ das Teilsystem von $\tilde{A}^t y = \tilde{c}$, welches aus allen Zeilen mit Indizes i besteht, für die $s_i^{(k)} < \frac{\eta}{4(m+2)}$, d.h. $s_i^* = 0$, gilt. Für alle anderen Zeilen wissen wir bereits, dass $y_i^* = 0$, und können diese daher ignorieren. Falls $\overline{A}^t y = \overline{c}$ genau eine Lösung hat, können wir diese einfach berechnen und erhalten eine Optimallösung.

Anderenfalls prüfen wir, ob $y_{i_0}^{(k)} < \frac{\eta}{4(m+2)}$ für ein $i_0 \in \{1, \dots, m\}$. Falls ein solches existiert, wissen wir, dass es eine Optimallösung mit $y_{i_0}^* = 0$ gibt, können also $\tilde{A}^t y = \tilde{c}$ weiter verkleinern.

Falls kein solches i_0 existiert, stelle sicher, dass $\tilde{A}x = \tilde{b}$ keine Lösung besitzt: Falls es eine Lösung gibt, ändere $\tilde{b}$ zu b^* ab, indem zu einem Eintrag ein $\varepsilon > 0$ addiert wird:

Wähle n linear unabhängige Zeilen von A. Ändere $\tilde{b}$ eine einer anderen Zeile um ε. ε sie dabei so klein gewählt, dass eine optimale Lösung des dualen LPs bzgl. b^* stets auch eine optimale Lösung des ursprünglichen dualen LP ist.

ε kann dabei so gewählt werden, dass size(ε) polynomiell beschränkt ist: Der Betrag der Differenz der Kosten zweier optimaler Basislösungen ist 0 oder $\geqslant 2^{-L}$ für ein L polynomiell in der Eingabegröße. Selbes gilt also auch für $|\tilde{b}^t u|$, wobei u die Differenz zweier Basislösungen sei.

Angenommen das initiale duale LP ist zulässig und beschränkt. Dann können wir Optimallösungen x^*, y^*, s^* der modifizierte LPs ((37), (43)) berechnen, indem wir Optimallösungen des initialen primalen und dualen LPs kanonisch erweitern. Insbesondere setzen wir dabei $x_{n+1} = 0$. Dann ist $Ax^* + s^* = b^*$, aber $Ax = b^*$ hat keine zulässige Lösung. Folgich muss ein $i_i \in \{1, \dots, m\}$ existieren, so dass $s^*_{i_0} > 0$, d.h. $y^{(k)}_{i_0} < \frac{\eta}{4(m+2)} \implies y^*_{i_0} = 0$. Hiermit konnte eine weitere duale Variable eliminiert werden.

Spätenstens wenn alle Variablen eliminiert wurden terminiert das Verfahren.

Dieser iterative Prozess lässt sich jedoch effizienter gestalten:

9.3.2 Effizientere Methode

Betrachte erneut die Probleme (51) und (51). Aus Theorem 32 folgt, dass wir die Indexmenge $\{1, \dots, m+2\}$ als $B \sqcup N$ partitionieren können, sodass für $i \in B$ eine duale Optimallösung y^* mit $y^*_i > 0$ und für $i \in N$ eine primale Optimallösung x^*, s^* mit $s^*_i > 0$ existiert. Jede Optimallösung kann als Konvexkombination von Basislösungen geschrieben werden. Daher kann in Lemma 124 für jedes i nur höchstens einer der Fälle $y_i < \frac{\eta}{4(m+2)}$ oder $s_i < \frac{\eta}{4(m+2)}$ eintreten.

Wenn wir k groß genug wählen, sodass $\mu^{(k)} < \frac{\eta^2}{32(m+2)^2\Delta}$ für ein noch zu bestimmendes $\Delta \geqslant 1$, so gilt für jedes i genau eine der Ungleichungen $u_{<} \frac{\eta}{4(m+2)\Delta}$ und $s_i < \frac{\eta}{4(m+2)\Delta}$. Hierdurch erhalten wir die Partitionierung $\{1, \dots, m+2\} = B \sqcup N$.

Insbesondere ist $y_i \geqslant \frac{\eta}{4(m+2)}$ für alle $i \in B$ und $y_i < \frac{\eta}{4(m+2)\Delta}$ für alle $i \in N$.

Seien A_B und A_N die Teilmatrizen von $\tilde{A}$ mit Zeilen mit Index aus B bzw. aus N.

Im Folgenden sei $\|\cdot\|$ die Euklidische Norm, bzw. die davon induzierte Spektralnorm.

Theorem 125. Sei $\Delta = \max\{\sqrt{m+2} \cdot \|A_B(A_B^t A_B)^{-1} A_N^t\|, 1\}$ und k groß genug, so dass $\mu^{(k)} < \frac{\eta^2}{32(m+2)^2\Delta}$.

Sei Y_B eine Diagonalmatrix, deren Zeilen und Spalten über B indiziert werden, so dass $Y_{B\,i,i} = y^{(k)}_i$ und

$$d_y := Y_B A_B (A_B^t (Y_B)^2 A_B)^{-1} A_N^t y_N^{(k)}$$

und $\tilde{y}_B = Y_B d_y + y_B^{(k)}$

Dann gilt:

(a) $A_B^t \tilde{y}_B = \tilde{c}$

(b) $\|d_y\| < 1$

(c) $\tilde{y} \in \mathbb{R}^{m+2}$ ($\tilde{y}_B$ mit zusätzlichen Einträgen 0 für Indizes aus N) ist eine Optimallösung.

Beweis. (a) Es ist $A_B^t(Y_B d_y + y_B^{(k)}) = A_N^t y_N^{(k)} + A_B^t y_B^{(k)} = \tilde{c}$.

(b)

$$\begin{aligned}
\|d_y\| &= \|Y_B A_B (A_B^t (Y_B)^2 A_B)^{-1} A_N^t y_N^{(k)}\| \\
&= \|Y_B A_B (A_B^t (Y_B)^2 A_B)^{-1} \underbrace{A_B^t Y_B Y_B^{-1} A_B (A_B^t A_B)^{-1}}_{=I_n} A_N^t y_N^{(k)}\| \\
&= \underbrace{\|Y_B A_B (A_B^t (Y_B)^2 A_B)^{-1} A_B^t Y_B\|}_{=1} \cdot \|Y_B^{-1} A_B (A_B^t A_B)^{-1} A_N^t y_N^{(k)}\| \\
&\leqslant \underbrace{\|Y_B^{-1}\|}_{\leqslant \frac{4(m+2)}{\eta}} \cdot \underbrace{\|A_B (A_B^t A_B)^{-1} A_N^t\|}_{\leqslant \frac{\Delta}{\sqrt{m+2}}} \cdot \underbrace{\|y_N^{(k)}\|}_{< \frac{\eta\sqrt{m+2}}{4(m+2)\Delta}} \\
&\leqslant 1
\end{aligned}$$

(c) Nach (a) gilt $\tilde{A}^t \tilde{y} = \tilde{c}$ und nach (b) wissen wir, dass $\tilde{y}_B > 0$, folglich ist $\tilde{y} \geqslant 0$. Daher ist $\tilde{y}$ zulässig. Ferner wissen wir, dass eine zulässige primale Lösung existiert, in der $s_i = 0$ für alle $i \in B$. Nach Theorem 28 ist $\tilde{y}$ daher eine optimale Lösung des dualen LP.

□

Theorem 126. Gegeben ein zulässiges und beschränktes lineares Programm $\min\{b^t y | y^t A = c^t, y \geqslant 0\}$ mit $A \in \mathbb{Q}^{m\times n}$, $b \in \mathbb{Q}^m$ und $c \in \mathbb{Q}^n$ berechnet die Innere-Punkte-Methode eine optimale Lösung in polynomieller Zeit. Ferner entscheidet der Algoritmus korrekt, ob ein LP beschränkt und zulässig ist.

10 Ganzzahlige lineare Programmierung

Durch Ganzzahligkeitsbedingungen können weitere Bedingungen modelliert werden, die mit linearer Programmierung nicht ausgedrückt werden können. Beispielsweise reicht es, für einige Variablen Bedingungen der Form $x \in \{0, 1\}$ zuzulassen (**Binäre Lineare Programmierung**) um folgendes auszudrücken:

- $(x \geqslant a \vee y \geqslant b) \wedge x, y \geqslant 0$

- $x \in \{s_1, \ldots, s_k\}$

Wir haben bereits gesehen, dass NP-schwere Optimierungsprobleme als (gemischt) ganzzahlige lineare Programme geschrieben werden können, somit können wir nicht hoffen, einen linearen Algorithmus für dieses Problem zu finden.

10.1 Ganzzahlige Polyeder

Definition 127. Sei $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ ein Polyeder. Wir nennen $P_I := \operatorname{conv}\{x \in \mathbb{Z}^n | Ax \leqslant b\}$ die **ganzzahlige Hülle** (integer hull) von P.

Definition 128. Ein Polyeder P heißt **ganzzahlig**, wenn $P = P_I$.

Definition 129. Ein Polyeder $P \subseteq \mathbb{R}^n$ heißt **rational**, wenn es $A \in \mathbb{Q}^{m \times n}$ und $b \in \mathbb{Q}^m$ mit $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ gibt.

Warning. *Ein rationales Polyeder enthält i.A. nicht nur rationale Vektoren. Insbesondere ist der Begriff auch nicht analog zum Begriff eines ganzzahligen Polyeder. Der Begriff ist somit irgendwie irreführend und man sollte ihn vielleicht lieber nicht verwenden. Das gelingt uns hier allerdings nicht.*

Observe.
- *Für rationale polyedrische Kegel ist $C_I = C$, denn ein polyedrischer Kegel ist genau dann rational, wenn er von einer endlichen Anzahl ganzzahliger Vektoren aufgespannt wird.*
- *P_I ist nicht notwendigerweise ein Polyeder.*
- *Falls P ein Polytop ist, so ist P_I ein Polyeder, da dann $P \cap \mathbb{Z}^n$ endlich ist.*

Lemma 130. Seien P, Q Polyeder. Dann ist $P_I + Q_I \subseteq (P + Q)_I$.

Beweis. Übung □

Theorem 131. Es sei $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ ein rationales Polyeder, d.h. $A \in \mathbb{Q}^{m \times n}, b \in \mathbb{Q}^m$. Dann ist P_I ein Polyeder.

Beweis. Sei $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ rational. Wir können P schreiben als $P = \operatorname{conv}(V) + \operatorname{cone}(E)$, wobei $V, E \subseteq \mathbb{R}^n$ endlich sind (Theorem 55). Aus dem Beweis wird auch klar, dass für rationale Polyeder $V, E \subseteq \mathbb{Q}^n$ gewählt werden können.

Somit können wir nach Skalierung sogar annehmen, dass $E = \{y_1, \ldots, y_s\} \subseteq \mathbb{Z}^n$. Definiere

$$B := \left\{ \sum_{i=1}^{s} \lambda_i y_i | 0 \leqslant \lambda_i \leqslant 1 \right\}$$

.

Behauptung 1. $P_I = (\mathrm{conv}(V) + B)_I + \mathrm{cone}(E)$.

Unterbeweis. "$P_I \subseteq (\mathrm{conv}(V) + B)_I + \mathrm{cone}(E)$".

Sei $p \subseteq P \cap \mathbb{Z}^n$. Dann ist $p = q + c$ für ein $q \in \mathrm{conv}(V)$ und ein $c = \sum_{i=1}^{s} \mu_i y_i \in \mathrm{cone}(E)$, wobei $\mu_i \geqslant 0$..

Es gilt

$$c = \sum_{i=1}^{s} \mu_i y_i = \underbrace{\sum_{i=1}^{s} (\mu_i - \lfloor \mu_i \rfloor) y_i}_{\in B} + \underbrace{\sum_{i=1}^{s} \lfloor \mu_i \rfloor y_i}_{\in \mathrm{cone}(E) \cap \mathbb{Z}^n}$$

Folglich kann $c = b + c'$ für $b \in B, c' \in \mathrm{cone}(E) \cap \mathbb{Z}^n$ geschrieben werden. Somit ist $p = (q + b) + c'$. Es ist $q + b \in \mathrm{conv}(V) + B$ und $q + b = p - c' \in \mathbb{Z}^n$

"$P_I \supseteq (\mathrm{conv}(V) + B)_I + \mathrm{cone}(E)$ ".

Es ist

$$\begin{aligned}
(\mathrm{conv}(V) + B)_I + \mathrm{cone}(E) &\subseteq P_I + \mathrm{cone}(E) \\
&= P_I + (\mathrm{cone}(E))_I \\
&\subseteq (P + \mathrm{cone}(E))_I \\
&= P_I
\end{aligned}$$

■

Hieraus folgt das Theorem, da $\mathrm{conv}(V) + B$ ein Polytop und daher $(\mathrm{conv}(V) + B)_I$ polyedrisch ist.

□

Remark 131.1. In diesem Fall kann man ganzzahlige Optimierung lösen, indem man Ecklösungen von P_I betrachtet. Das Problem dabei ist jedoch, dass P_I u.U. eine wesentlich höhere Komplexität als P besitzt.

Theorem 132. Sei $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ mit $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$, sodass $P_I \neq \varnothing$. Sei $c \in \mathbb{R}^n$.

Dann ist $\max\{c^t x | x \in P\}$ genau dann beschränkt, wenn $\max\{c^t x | x \in P_I\}$ beschränkt ist.

Beweis. " $\Longrightarrow$ " klar.

" $\Longleftarrow$ " Angenommen $\max\{c^t x | x \in P\}$ ist unbeschränkt.

Dann ist das duale LP $\min\{b^t y | y^t A = c, y \geqslant 0\}$ unzulässig. Nach dem Lemma von Farkas (Theorem 23) existiert ein Vektor z mit $c^t z < 0, Az \geqslant 0$.

Das LP $\min\{c^t x : Ax \geqslant 0, -\mathbb{1}_n \leqslant x \leqslant \mathbb{1}_n\}$ ist daher zulässig und hat eine Optimallösung mit negativem Wert. Sei x^* eine rationale Optimallösung. Nach Multiplikation von x^* mit dem Hauptnenner erhalten wir einen Vektor $w \in \mathbb{Z}^n$ mit $Aw \geqslant 0$, $c^t w < 0$.

Für jedes $v \in P_I$ und $k \in \mathbb{N}$ gilt $v - k \cdot w \in P_I$. Somit ist $\max\{c^t x | x \in P_I\}$ unbeschränkt. □

10.2 Ganzzahlige Lösungen von Gleichungssystemen

Goal. *Finde ein Zertifikat dafür, dass ein Gleichungssystem keine ganzzahlige Lösung hat.*

Definition 133. Eine $m \times n$-Matrix A ist in **Hermitescher Normalform**, wenn sie in der Form $A = [B\ 0]$ geschrieben werden kann, wobei B eine reguläre nicht-negative untere Dreiecksmatrix ist, sodass in jeder Zeile von B der Diagonaleintrag der größte Eintrag ist.

Die folgenden Modifikationen von Matrizen heißen **elementare unimodulare Spaltenoperationen**:

- Vertausche zwei Spalten
- Multipliziere eine Spalte mit -1
- Addiere ein ganzzahliges Vielfaches einer Spalte zu einer anderen Spalte.

Theorem 134. Jede Matrix $A \in \mathbb{Q}^{m \times n}$ mit Rang m kann durch eine Folge von unimodularen Spaltenoperationen in eine Matrix in Hermitescher Normalform gebracht werden.

Beweis. O.B.d.A. sei A ganzzahlig. Wir nehmen an, dass A bereits in eine Matrix $\begin{pmatrix} F & 0 \\ G & H \end{pmatrix}$ transformiert wurde, wobei F eine linke untere Dreiecksmatrix mit positiver Diagonale ist.

Die erste Zeile von H habe die Einträge $h_{11}, \ldots, h_{1k}$. Wende elementare unimodulare Spaltenoperationen an, sodass alle h_{1j} nicht-negativ sind und $\sum_{j=1}^{k} h_{1j}$ so klein wie möglich ist.

Außerdem sei $h_{11} \geqslant h_{12} \geqslant \ldots \geqslant h_{1k}$. Dann ist $h_{11} > 0$ (da A Rang m hat) und $h_{12} = \ldots = h_{1k} = 0$. Am Ende erhalten wir eine Matrix $[B\ 0]$, wobei B eine linke untere Dreiecksmatrix mit positiver Diagonale ist.

Bezeichne die Einträge von B mit $b_{ij}, i, j \in \{1, \ldots, n\}$. Führe zum Schluss für $i = 2, \ldots, m$ die folgenden Schritte durch:

Für $j = 1, \ldots, i-1$ addiere ein ganzzahliges Vielfaches der i-ten Spalte zur j-ten Spalte von B, sodass b_{ij} nicht-negativ und kleiner als b_{ii} ist. □

Wir bekommen für ganzzahlige Lösbarkeit das folgende Analogon zu Farkas:

Corollary 135. Sei $A \in \mathbb{Q}^{m\times n}$ und $b \in \mathbb{Q}^m$. Dann hat $Ax = b$ genau dann eine ganzzahlige Lösung x, wenn $b^t y \in \mathbb{Z}$ für jedes $y \in \mathbb{Q}^m$ mit $A^t y \in \mathbb{Z}^n$.

Ein Zertifikat dafür, dass $Ax = b$ nicht ganzzahlig lösbar ist, ist somit ein Vektor $y \in \mathbb{Q}^m$, sodass $A^t y \in \mathbb{Z}^n, b^t y \notin \mathbb{Z}$.

Beweis. „ $\Longrightarrow$ " Wenn $x \in \mathbb{Z}^n$ mit $Ax = b$ und $y^t A \in \mathbb{Z}^n$ ist, dann ist $\underbrace{y^t A}_{\in\mathbb{Z}} \underbrace{x}_{\in\mathbb{Z}} = y^t b \in \mathbb{Z}$.

„ $\Longleftarrow$ " Es sei $b^t y$ ganzzahlig für jedes $y \in \mathbb{Q}^m$, für das $A^t y$ ganzzahlig ist.

Dann muss $Ax = b$ zulässig sein, denn sonst gäbe es nach Farkas einen Vektor $y \in \mathbb{Q}^m$ mit $y^t A = 0$ und $y^t b = -\frac{1}{2}$.

Wir können annehmen, dass A Rang m hat, indem wir redundante Gleichungen entfernen.

Die Aussage des Korollars gilt genau dann für A, wenn sie für eine Matrix $\tilde{A}$ gilt, die aus A durch Anwendung von elementaren unimodularen Spaltenoperationen hervorgeht. (Ay ist ganzzahlig, genau dann wenn $\tilde{A}y$ ganzzahlig ist).

O.B.d.A. sei A in Hermitescher Normalform $A = [B\ 0]$. Es ist $B^{-1}A = B^{-1}[B\ 0] = [I_m\ 0]$ eine ganzzahlige Matrix. Nach Voraussetzung (angewandt auf die Zeilen von B^{-1}) ist daher $B^{-1}b$ ganzzahlig.

Wegen $[B\ 0]\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} = b$ ist der Vektor $x := \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$ eine ganzzahlige Lösung von $[B\ 0]x = b$. □

10.3 TDI-Systeme

Theorem 136. Sei $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ wobei $A \in \mathbb{Q}^{m\times n}$ und $b \in \mathbb{Q}^m$. Dann sind äquivalent:

(a) P ist ganzzahlig.

(b) Jede Fläche von P enthält mindestens einen ganzzahligen Vektor.

(c) Jede minimale Fläche von P enthlält mindestens einen ganzzahligen Vektor.

(d) Jede Stützhyperebene von P enthält mindestens einen ganzzahligen Vektor.

(e) Jede rationale Stützhyperebene von P enthält mindestens einen ganzzahligen Vektor.

(f) $\max\{c^t x | x \in P\}$ wird für jeden ganzzahligen Vektor c, für den das Maximum endlich ist, von einem ganzzahligen Vektor angenommen.

(g) $\max\{c^t x | x \in P\}$ ist für jeden ganzzahligen Vektor c, für den das Maximum endlich ist, ganzzahlig.

Beweis.

- $(b) \iff (c)$, $(b) \implies (d)$, $(d) \implies (e)$, $(f) \implies (g)$ klar.
- $(a) \implies (b)$ Sei P ganzzahlig. Sei $F = P \cap H$ eine Fläche von P, wobei $H = \{x \in \mathbb{R}^n | c^t x = \delta\}$ eine Stützhyperebene von P sei. Dabei sei $\delta = \max\{c^t x | x \in P\}$. Jedes $z \in F$ ist Konvexkombination von ganzzahligen Vektoren $v_1, \dots, v_k \in P$. Falls $v_i \in P \backslash V$, so gilt $c^t v_i < \delta$. Somit muss mindestens eines der v_i bereits in F liegen.
- $(c) \implies (f)$: Die Menge der Optimallösungen enthält immer eine minimale Fläche.
- $(f) \implies (a)$: Angenommen es gelte (f), aber $P \neq P_I$. Dann gibt es ein $x^* \in P \backslash P_I$. Da P rational ist, ist P_I ein Polyeder. Es gibt daher eine Ungleichung $a^t x \leqslant \beta$, die von den Vektoren in P_I, aber nicht von x^* erfüllt wird. $\lightning$ zu (f), da $\max\{a^t x | x \in P\}$ endlich ist (Theorem 132)
- $(e) \implies (c)$: Wir können annehmen, dass A, b ganzzahlig sind. Sei $F = \{x \in \mathbb{R}^n | A' x = b'\}$ eine minimale Fläche von P, wobei $A'x \leqslant b'$ ein Teilsystem von $Ax \leqslant b$ sei. Wenn es keinen ganzzahligen Vektor x mit $A'x = b'$ gibt, dann gibt es nach Corollary 135 einen rationalen Vektor y, sodass $c := (A')^t y$ ganzzahlig ist, während $\delta := y^t b'$ nicht ganzzahlig ist. Wir können annehmen, dass alle Einträge von y positiv sind (sonst addiere einen geeigneten ganzzahligen Vektor zu y).

  Die Menge $H := \{x \in \mathbb{R}^n | c^t x = \delta\}$ enthält daher keinen ganzzahligen Vektor.

  Behauptung 1. $H \cap P = F$

  Unterbeweis. Nach Konstruktion gilt $F \subseteq H$. Sei $x \in H \cap P$. Dann ist $y^t A' x = c^t x = \delta = y^t b'$. Es folgt $y^t (A'x - b') = 0$. Weil alle Einträge von y positiv sind, folgt $A'x = b' \implies x \in F$. ■

- $(g) \implies (e)$: Sei $H = \{x \in \mathbb{R}^n | c^t x = \delta\}$ eine rationale Stützhyperebene von P mit $\max\{c^t x | x \in P\} = \delta$.

  Angenommen H enthält keinen ganzzahligen Vektor. Nach Corollary 135 existiert ein $\gamma > 0$ mit $\gamma \cdot c \in \mathbb{Z}^n, \gamma \cdot \delta \notin \mathbb{Z}$.

  Dann ist $\max\{(\gamma c)^t x | x \in P\} = \gamma \cdot \max\{c^t x | x \in P\} = \gamma \cdot \delta \notin \mathbb{Z}$ im Widerspruch zu (g).

□

Definition 137. Ein Ungleichungssystem $Ax \leqslant b$ heißt ~~vollständig dual ganzzahlig~~[a] **totally dual integral** (**TDI**), wenn das LP $\min\{b^t y | A^t y = c, y \geqslant 0\}$ für jeden ganzzahligen Vektor c, für den das LP zulässig ist und beschränkt ist, eine ganzzahlige Optimallösung hat.

[a] Dieser Begriff wird nie verwendet

Warning. *TDI zu sein ist eine Eigenschaft eines Ungleichungssystems, nicht eines Polyeders.*

Theorem 138. Seien $A \in \mathbb{Q}^{m \times n}$ und $b \in \mathbb{Z}^n$, sodass $Ax \leqslant b$ TDI ist. Dann ist das Polyeder $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ ganzzahlig.

Beweis. Wenn $Ax \leqslant b$ TDI ist, dann ist $\min\{b^t y | A^t y = c, y \geqslant 0\}$ für jeden Vektor $c \in \mathbb{Z}^n$, für den das Minimum endlich ist, eine ganze Zahl (da b ganzzahlig ist). [5]

Auf Grund der Dualität ist $\max\{c^t x | Ax \leqslant b\}$ für jeden Vektor $c \in \mathbb{Z}^n$, für den das Maximum endlich ist, eine ganze Zahl. Mit der Implikation $(g) \implies (a)$ aus Theorem 136 folgt die Aussage. □

Theorem 139 (Hinzufügen redundanter Ungleichung)**.** Wenn $Ax \leqslant b$ TDI ist und $a^t x \leqslant \beta$ für jedes $x \in \mathbb{R}^n$ mit $Ax \leqslant b$ erfüllt ist, dann ist das System $Ax \leqslant b, a^t x \leqslant \beta$ auch TDI.

Beweis. Sei $Ax \leqslant b$ TDI und $a^t x \leqslant \beta$ wie im Satz.

Sei c ein ganzzahliger Vektor, für den das LP $\min\{b^t y + \beta\gamma | A^t y + \gamma a = c, y \geqslant 0, \gamma \geqslant 0\}$ zulässig und beschränkt ist.

Dann ist

$$\begin{aligned} & \min\{b^t y + \beta\gamma | A^t y + \gamma a = c, y \geqslant 0, \gamma \geqslant 0\} \\ \overset{\text{Dualität}}{=} & \max\{c^t x | Ax \leqslant b, a^t x \leqslant \beta\} \\ = & \max\{c^t x | Ax \leqslant b\} \\ \overset{\text{Dualität}}{=} & \min\{b^t y | A^t y = c, y \geqslant 0\} \end{aligned}$$

Das letzte LP hat eine ganzzahlige Optimallösung y^*. Zusammen mit $\gamma^* = 0$ liefert das eine ganzzahlige Optimallösung des ersten LPs. □

[5] Achtung: Aus TDI folgt nur, dass der Lösungsvektor ganzzahlig ist. Um zu zeigen, dass der Lösungswert ganzzahlig ist, wird die Ganzzahligkeit von b benötigt.

Theorem 140 (Ungleichung durch Gleichung ersetzen)**.** Wenn $Ax \leqslant b$, $a^t x \leqslant \beta$ TDI ist und a ganzzahlig, dann ist $Ax \leqslant b, a^t x = \beta$ TDI.

Beweis. Sei c ein ganzzahliger Vektor, für den

$$\max\{c^t x | Ax \leqslant b, a^t x = \beta\} \tag{57}$$
$$= \min\{b^t y + \beta(\lambda - \mu) | y \geqslant 0, \lambda, \mu \geqslant 0, A^t y + (\lambda - \mu)a = c\} \tag{58}$$

endlich ist.[6]

Seien x^* und y^*, λ^*, μ^* eine primale und eine duale Optimallösung. Sei $\tilde{c} := c + \lceil \mu^* \rceil a$. Dann ist

$$\max\{\tilde{c}^t x | Ax \leqslant b, a^t x \leqslant \beta\} \tag{59}$$
$$= \min\{b^t y + \beta\lambda | y \geqslant 0, \lambda \geqslant 0, A^t y + \lambda a = \tilde{c}\} \tag{60}$$

endlich, weil x^* eine zulässige Lösung des Maximierungsproblems und y^* und $\lambda^* + \lceil \mu^* \rceil - \mu^*$ eine zulässige Lösung des Minimierungsproblems ist.

Da $Ax \leqslant b, a^t x \leqslant \beta$ TDI ist, hat das letzte Minimierungsproblem eine ganzzahlige Optimallösung $\tilde{y}, \tilde{\lambda}$.

Behauptung 1. *$y := \tilde{y}$, $\lambda := \tilde{\lambda}$, $\mu := \lceil \mu^* \rceil$ ist eine ganzzahlige Optimallösung für das Minimierungsproblem in* (58)*.*

Unterbeweis. Die Lösung ist offenbar zulässig, denn

$$A^t y + \lambda a - \mu a = \underbrace{A^t \tilde{y} + \tilde{\lambda} a}_{\tilde{c}} - \lceil \mu^* \rceil a = c$$

Die Kosten sind

$$\begin{aligned} b^t \tilde{y} + \beta(\tilde{\lambda} - \lceil \mu^* \rceil) &= b^t \tilde{y} + \beta \tilde{\lambda} - \beta \lceil \mu^* \rceil \\ &\leqslant b^t y^* + \beta\left(\lambda^* + \lceil \mu^* \rceil - \mu^*\right) - \beta \lceil \mu^* \rceil \\ &= b^t y^* + \beta(\lambda^* - \mu^*) \end{aligned}$$

wobei die Ungleichung gilt, da $y^*, \lambda^* + \lceil \mu^* \rceil$ eine zulässige Lösung des Minimierungsproblemes in (60) ist. ■

Insbesondere hat das Minimierungsproblem eine ganzzahlige Optimallösung, daher ist $Ax \leqslant b, a^t x = \beta$ TDI.

□

[6] Hier wird vor dem Dualisieren die Gleichung durch zwei Ungleichungen ersetzt.

Definition 141. Eine endliche Menge $\{v_1, \dots, v_t\}$ von Vektoren heißt **Hilbert-Basis**, wenn jeder ganzzahlige Vektor in $\operatorname{cone}(\{v_1, \dots, v_t\})$ als nicht-negative ganzzahlige Linearkombination von $v_1, \dots, v_t$ geschrieben werden kann.

Theorem 142. Jeder rationale polyedrische Kegel wird durch eine ganzzahlige Hilbert-Basis erzeugt.

Beweis. Sei C ein rationaler polyedrischer Kegel. Dann wird C von rationalen Vektoren $b_1, \dots, b_k$ erzeugt.

O.B.d.A. seien $b_1, \dots, b_k$ ganzzahlig. Definiere

$$H := \mathbb{Z}^n \cap \left\{ \sum_{i=1}^{k} \lambda_i b_i | \forall i: \; 0 \leqslant \lambda_i \leqslant 1 \right\}$$

H ist offenbar endlich.[7]

Behauptung 1. *H ist eine Hilbert-Basis, die C erzeugt.*

Unterbeweis. Wegen $\{b_1, \dots, b_k\} \subseteq H \subseteq C$ gilt $C = \operatorname{cone}(H)$.

Sei b ein ganzzahliger Vektor in C. Dann gibt es nichtnegative Zahlen $\mu_1, \dots, \mu_k$ mit $b = \sum_{i=1}^{k} \mu_i b_i$, also

$$b = \left(\sum_{i=1}^{k} \lfloor \mu_i \rfloor b_i \right) + \sum_{i=1}^{k} (\mu_i - \lfloor \mu_i \rfloor) b_i$$

Der Vektor

$$b - \left(\sum_{i=1}^{k} \lfloor \mu_i \rfloor b_i \right) = \sum_{i=1}^{k} (\mu_i - \lfloor \mu_i \rfloor) b_i$$

ist ganzzahlig und ein Element von P, wegen $0 \leqslant \mu_i - \lfloor \mu_i \rfloor \leqslant 1$ also ein Element von H. Folglich kann b als nicht-negative ganzzahlige Linearkombination von Elementen aus H geschrieben werden.

■

□

Notation 142.1. Für ein Ungleichungssystem $Ax \leqslant b$ und eine Fläche F von $\{x \in \mathbb{R}^n | Ax \leqslant b\}$ heißt eine Zeile von A **aktiv** (in F), wenn die zugehörige Ungleichung in $Ax \leqslant b$ von allen Vektoren $x \in F$ mit Gleichheit erfüllt wird.

[7] Vergleiche den Beweis von Theorem 131.

Theorem 143. Ein zulässiges, rationales Ungleichungssystem $Ax \leqslant b$ ist genau dann TDI, wenn für jede minimale Fläche F von $P := \{x \in \mathbb{R}^n | Ax \leqslant b\}$ die Zeile von A, die in F aktiv sind, eine Hilbert-Basis bilden.

Beweis. „ $\Longrightarrow$ " Sei $Ax \leqslant b$ TDI. Sei F eine minimale Fläche von P und seien $a_1, \dots, a_t$ die Zeilen von A, die für F aktiv sind.

Zu zeigen ist, dass $\{a_1, \dots, a_t\}$ eine Hilbert-Basis bildet.

Sei c ein ganzzahliger Vektor in $\text{cone}\{a_1, \dots, a_t\})$. Das Maximum der Gleichung

$$\max\{c^t x | Ax \leqslant b\} = \min\{b^t y | A^t y = c, y \geqslant 0\} \tag{61}$$

wird von jedem $x \in F$ angenommen.

Da $Ax \leqslant b$ TDI ist, hat das duale LP eine ganzzahlige Optimallösung y. Aus dem komplementäre Schlupf (Theorem 28) folgt, dass die Einträge von y, die zu Zeilen von A gehören, welche für F nicht aktiv sind, 0 sind. c ist daher eine ganzzahlige nicht-negative Linearkombination von $a_1, \dots, a_t$. Folglich ist $a_1, \dots, a_t$ eine Hilbert-Basis.

„ $\Longleftarrow$ " Angenommen es bilden die aktiven Zeilen von A für jede minimale Fläche F eine Hilbert-Basis.

Sei $c \in \mathbb{Z}^n$ ein Vektor, für den die Optima in (61) endlich sind.

Zu zeigen ist, dass das Minimum von einem ganzzahligen Vektor angenommen wird.

Sei F eine minimale Fläche von P, sodass von jedem Vektor in F das Maximum in (61) angenommen wird.

Seien $a_1, \dots, a_t$ die in F aktiven Zeilen von A. Mit dem komplementären Schlupf folgt, dass $c \in \text{cone}\{a_1, \dots, a_t\}$.

Nach Annahme ist $\{a_1, \dots, a_t\}$ eine Hilbert-Basis. Daher gibt es $\lambda_1, \dots, \lambda_t \in \mathbb{Z}_{\geqslant 0}$ mit $c = \sum_{i=1}^t \lambda_i a_i$. $(\lambda_1, \dots, \lambda_t)$ kann mit Nullen zu einem Vektor $y \in \mathbb{Z}^m$ mit $y \geqslant 0, A^t y = c$ und $b^t y = x^t A^t y^t$ für alle $x \in F$ erweitert werden.

Dieses y ist somit eine ganzzahlige Optimallösung. $\square$

Theorem 144. Ein rationales Ungleichungssystem $Ax \leqslant 0$ ist genau dann TDI, wenn die Zeilen von A eine Hilbert-Basis bilden.

Beweis. In der Minimalfläche von $\{x \in \mathbb{R}^n | Ax \leqslant 0\}$ ist die 0 enthalten und alle Zeilen sind aktiv. Mit Theorem 143 folgt die Aussage. $\square$

Theorem 145 (Giles und Pullyblank)**.** Für jedes rationale Polyeder $P \subseteq \mathbb{R}^n$ gibt es ein rationales TDI-System $Ax \leqslant b$ mit $A \in \mathbb{Z}^{m \times n}$ und $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$. Der Vektor b kann genau dann ganzzahlig gewählt werden, wenn P ganzzahlig ist.

Beweis. Es sei o.B.d.A. $P \neq \emptyset$. Sei F eine minimale Fläche von P. Für eine Darstellung $P = \{\tilde{A}x \leqslant \tilde{b}\}$ sei C_F der von den in F aktiven Zeilen von $\tilde{A}$ erzeuge Kegel (C_F ist unabhängig von der Wahl von $\tilde{A}$, denn $C_F = \{c \in \mathbb{R}^n | \forall z \in F\ c^t z = \max\{c^t x | x \in P\}\}$).

Sei $a_1, \dots, a_t$ eine ganzzahlige Hilbert-Basis, die C_F erzeugt. Wähle $x_0 \in F$ und definiere $\beta_i := a_i^t x_0$ für $i = 1, \dots, t$. Es ist $\beta_i = \max\{a_i^t x | x \in P\}$ $(i = 1, \dots, t)$.

Sei $\mathcal{S}_F$ das Ungleichungssystem $a_1^t x \leqslant \beta_1, \dots, a_t^t x \leqslant \beta_t$. Alle Ungleichungen in $\mathcal{S}_F$ werden von jedem Element aus P erfüllt.

Sei $Ax \leqslant b$ die Vereinigung aller Systeme $\mathcal{S}_F$ über alle minimalen Flächen F von P.

Es ist $P \subseteq \{x \in \mathbb{R}^n | Ax \leqslant b\}$. Falls $x^* \in \mathbb{R}^n \backslash P$, dann gibt es eine Stützhyperebene von P, die x^* von P trennt, und diese Stützhyperebene berührt P in einer minimalen Fläche F. In dem zugehörigen Ungleichungssystem $\mathcal{S}_F$ wird mindestens eine Ungleichung verletzt ↯ Folglich ist $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$. Nach Konstruktion ist $Ax \leqslant b$ TDI nach Konstruktion, da die $\mathcal{S}_F$ stets eine Hilbert-Basis bilden.

Falls P ganzzahlig ist, so können alle β_i ganzzahlig gewählt werden, indem x_0 ganzzahlig gewählt wird.

Wegen Theorem 138 gilt andererseits $b \in \mathbb{Z}^m \implies P = P_I$. □

Für das primal-duale Paar $\max\{c^t x | Ax \leqslant b\} = \min\{b^t y | A^t y = c, y \geqslant 0\}$ wissen wir, dass wenn beide Optima endlich sind, das Minimierungsproblem eine Optimallösung y mit höchstens $\operatorname{rank}(A)$ von 0 verschiedenen Einträgen besitzt. Wenn wir ganzzahlige Lösungen mit $Ax \leqslant b$ TDI und b ganzzahlig suchen, so ist das nicht Notwendigerweise der Fall: Betrachte z.B. $A = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$, $b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ und $c = (1)$. Für volldimensionale Lösungsräume erhalten wir dennoch eine Schranke für die Anzahl an von 0 verschiedenen Einträgen:

Theorem 146. Sei $Ax \leqslant b$ ein TDI-System mit $A \in \mathbb{Z}^{m \times n}$ so dass $\dim(\{x \in \mathbb{R}^n | Ax \leqslant b\}) = n$. Sei $c \in \mathbb{Z}^n$, so dass

$$\max\{c^t x | Ax \leqslant b\} = \min\{b^t y | A^t y = c, y \geqslant 0\}$$

endlich sind. Dann hat das Minimierungsproblem eine ganzzahlige Optimallösung y mit höchstens $2r - 1$ positiven Einträgen, wobei $r := \operatorname{rank}(A)$.

Beweis.

Behauptung 1. *Sei* $\{a_1, \ldots, a_t\} \subseteq \mathbb{Z}^n$ *eine Hilbert-Basis, so dass* $C := \operatorname{cone}(\{a_1, \ldots, a_t\})$ *ein spitzer* k*-dimensionaler Kegel ist. Dann ist jeder ganzzahlige Vektor* $c \in C$ *nicht-negative ganzzahlige Kombination von höchstens* $2k - 1$ *Vektoren aus* $a_1, \ldots, a_t$.

Unterbeweis. Seien $\lambda_1, \ldots, \lambda_t$ so gewählt, dass

$$\max \left\{ \sum_{i=1}^{t} \lambda_i | \lambda_1, \ldots, \lambda_t \geqslant 0; c = \sum_{i=1}^{t} \lambda_i a_i \right\}$$

angenommen wird. Da C spitz ist, ist das Maximum endlich (das duale LP $\min\{c^t y | y^t a_i \geqslant 1 \ \forall i \in \{1, \ldots, t\}\}$. Wir können annehmen, dass höchstens k der λ_i von 0 verschieden sind.

Definiere

$$c' := c - \sum_{i=1}^{t} \lfloor \lambda_i \rfloor a_i = \sum_{i=1}^{t} (\lambda_i - \lfloor \lambda_i \rfloor) a_i$$

Dann ist c' ein ganzzahliger Vektor in C und kann folglich als $c' = \sum_{i=1}^{t} \mu_i a_i$ für ganzzahlige $\mu_i \geqslant 0$ geschrieben werden.

Da $\lambda_1, \ldots, \lambda_t$ eine Optimallösung war, und $\mu_1 + \lfloor \lambda_1 \rfloor, \ldots, \mu_t + \lfloor \lambda_i t \rfloor$ eine zulässige Lösung ist, ist $\sum_{i=1}^{t} \mu_i + \lfloor \lambda_i \rfloor \leqslant \sum_{i=1}^{t} \lambda_i$ und daher

$$\sum_{i=1}^{t} \mu_i \leqslant \sum_{i=1}^{t} \lambda_i - \sum_{i=1}^{t} \lfloor \lambda_i \rfloor < k$$

denn höchstens k der λ_i sind von 0 verschieden. Daher sind höchstens $k - 1$ der μ_i von 0 verschieden und mit

$$c = \sum_{i=1}^{t} (\lfloor \lambda_i \rfloor + \mu_i) a_i$$

folgt die Aussage. ■

Behauptung 2. *Falls nun* $P := \{x \in \mathbb{R}^n | Ax \leqslant b\}$ *volldimensional ist, dann muss ein von den für eine minimale Fläche* F *von* P *aktiven Zeilen von* A *erzeugter Kegel spitz sein.*

Unterbeweis. Anderenfalls wäre ein Paar $v, -v$ im Kegel enthalten. Folglich würden sich Ungleichungen $v^t x \leqslant \beta_1$ und $-v^t x \leqslant \beta_2$ als nicht-negative Linearkombinationen von Ungleichungen aus $Ax \leqslant b$, welche zu aktiven Zeilen gehören, schreiben lassen. Für $x \in F$ gilt dann $v^t x = \beta_1$ und $-v^t x = \beta_2$, daher $\beta = -\beta_2$ und somit $P \subseteq \{x \in \mathbb{R}^n | v^t x = \beta_1\}$ im Widerspruch zur Annahme, dass P volldimensional ist. ■

Nach Theorem 143 folgt, dass die aktiven Zeilen für eine minimale Fläche $F \subseteq \operatorname{argmax}\{c^t x | Ax \leqslant b\}$ eine Hilbert-Basis bilden. □

10.4 Vollständig unimodulare Matrizen

Definition 147. Eine $m \times n$ -Matrix A mit Rang m heißt **unimodular**, wenn $A \in \mathbb{Z}^{m\times n}$ und $\det(B) \in \{-1, 1\}$ für alle regulären $m \times m$-Teilmatrizen B von A.

Observe. *Eine quadratische unimodulare Matrix hat eine ganzzahlige Inverse (Carmersche Regel).*

Definition 148. Eine Matrix A heißt **vollständig unimodular** (totally unimodular, TU), wenn jede Unterdeterminante von A (d.h. jede Determinante von quadratische Untermatrizen von A), 0, -1 oder 1 ist.

(Insbesondere sind alle Einträge 0, -1 oder 1).

Observe. *A ist genau dann vollständig unimodular, wenn $[I_m\ A]$ unimodular ist.*

Theorem 149. Sei A vollständig unimodular und sei b ein ganzzahliger Vektor. Dann ist das Polyeder $P = \{x \in \mathbb{R}^n | Ax \leqslant b\}$ ganzzahlig.

Beweis. Es genügt zu zeigen, dass jede minimale Fläche F von P einen ganzzahligen Vektor enthält.

Jede minimale Fläche von P kann als $F = \{x \in \mathbb{R}^n | A'x = b'\}$ geschrieben werden, wobei $A'x \leqslant b'$ ein Teilsystem von $Ax \leqslant b$ ist.

O.B.d.A. hat A' vollen Zeilenrang. Nach Vertauschen von Spalten können wir $A' = [U\ V]$ für eine Matrix U mit $\det(U) \in \{-1, 1\}$ annehmen.

Es ist $\begin{pmatrix} U^{-1}b' \\ 0 \end{pmatrix}$ ein ganzzahliger Vektor in F. □

Theorem 150. Sei $A \in \mathbb{Z}^{m\times n}$ eine Matrix mit Rang m. Dann ist A genau dann unimodular, wenn für jeden ganzzahligen Vektor b das Polyeder $\{x \in \mathbb{R}^n | Ax = b, x \geqslant 0\}$ ganzzahlig ist.

Beweis. “ $\Longrightarrow$ ” Sei A unimodular und b ein ganzzahliger Vektor.

Sei x' eine Ecke von $\{x \in \mathbb{R}^n | Ax = b, x \geqslant 0\}$. (Polyeder in Standart-Gleichungsform sind spitz). Dann gibt es n linear unabhängige Ungleichungen in $Ax \leqslant b$, $-Ax \leqslant -b$, $-I_n x \leqslant 0$, die von x' mit Gleichheit erfüllt werden. Die Spalten von A, die zu Nicht-Null-Einträgen von x' gehören, sind linear unabhängig. Diese Spaltenmenge kann zu einer regulären $m \times m$ Untermatrix B von A erweitert werden.

Die Einschränkung von x' auf Koordinaten, die zu B gehören ist $B^{-1}b$. Wegen $\det(B) \in \{-1, 1\}$ ist $B^{-1}b$ ganzzahlig. Alle anderen Einträge von x' sind 0, folglich ist x' ganzzahlig.

" $\Longleftarrow$ " Angenommen $\{x \in \mathbb{R}^n | Ax = b, x \geqslant 0\}$ ist für jeden ganzzahligen Vektor b ganzzahlig.

Sei B eine reguläre $m \times m$-Untermatrix von A.

Zu zeigen: $\det(B) \in \{-1, 1\}$.

Nach der Cramer'schen Regel genügt es zu zeigen, dass $B^{-1}u$ für jeden ganzzahligen Vektor u ganzzahlig ist.

Sei u ein ganzzahliger Vektor. Sei y ein ganzzahliger Vektor mit $z := y + B^{-1}u \geqslant 0$. Dann ist $b := Bz$ ganzzahlig.

Ergänze z mit Nullen zu einem Vektor z' mit $Az' = Bz = b$.

z' ist ein zulässige Basislösung von $Ax = b$. Folglich ist es eine Ecke von $\{x \in \mathbb{R}^n | Ax = b, x \geqslant 0\}$. Daher ist z' ganzzahlig und somit auch z. Insbesondere ist $B^{-1}u = z - y$ ganzzahlig. □

Theorem 151 (Hoffman und Kruskal)**.** Eine ganzzahlige Matrix A ist genau dann TU, wenn für jeden ganzzahligen Vektor b das Polyeder $\{x \in \mathbb{R}^n | Ax \leqslant b, x \geqslant 0\}$ ganzzahlig ist.

Beweis. A ist genau dann TU, wenn $[I_m\ A]$ unimodular ist.

Sei b ein ganzzahliger Vektor. Die Ecken von $\{x \in \mathbb{R}^n | Ax \leqslant b, x \geqslant 0\}$ sind genau dann ganzzahlig, wenn die Ecken von $\{z \in \mathbb{R}^{m+n} | [I_m\ A]z = b, z \geqslant 0\}$ ganzzahlig sind.

Die Aussage folgt daher aus Theorem 150. □

Theorem 152. Eine ganzzahlige Matrix A ist genau dann vollständig unimodular, wenn für alle ganzzahligen Vektoren b und c die Optima für beide Seiten der Dualitäts-Gleichung

$$\max\{c^t x | Ax \leqslant b, x \geqslant 0\} = \min\{b^t \mid A^t y \geqslant c, y \geqslant 0\}$$

von ganzzahligen Vektoren angenommen wird (wenn beide Optima existieren).

Beweis. Folgt direkt aus Theorem 151, denn A ist genau dann TU, wenn A^t TU ist. □

Corollary 153. Eine ganzzahlige Matrix A ist genau dann TU, wenn das Ungleichungssystem $Ax \leqslant b, x \geqslant 0$ für jeden Vektor b TDI ist.

Beweis. „ $\Longrightarrow$ " Wenn A TU ist, dann ist A^t TU. Aus Theorem 151 folgt, dass $\{y \in \mathbb{R}^m | A^t y \geqslant c, y \geqslant 0\}$ für jeden ganzzahligen Vektor c ganzzahlig ist, d.h. $\min\{b^t y | A^t y \geqslant c, y \geqslant 0\}$ für jeden Vektor b und jeden ganzzahligen Vektor c, für die das Minimum endlich ist, von einem ganzzahligen Vektor angenommen wird.

Folglich ist $Ax \leqslant b, x \geqslant 0$ für jeden Vektor b TDI.

„ $\Longleftarrow$ " Es sei $Ax \leqslant b, x \geqslant 0$ für jeden Vektor b TDI. Dann ist das Polyeder $\{x \in \mathbb{R}^n | Ax \leqslant b, x \geqslant 0\}$ für jeden ganzzahligen Vektor b ganzzahlig. Mit Theorem 151 folgt, dass A TU ist. □

Theorem 154 (Ghouila-Houri)**.** Eine Matrix $A = (a_{ij})_{\substack{i=1,\ldots,m\\ j=1,\ldots,n}} \in \mathbb{Z}^{m \times n}$ ist genau dann TU, wenn es für jede Menge $R \subseteq \{1, \ldots, n\}$ Mengen R_1, R_2 mit $R = R_1 \sqcup R_2$ gibt, sodass für jedes $i \in \{1, \ldots, m\}$ gilt:

$$\sum_{j \in R_1} a_{ij} - \sum_{j \in R_2} a_{ij} \in \{-1, 0, 1\}$$

Beweis. „ $\Longrightarrow$ " Sei A vollständig unimodular und $R \subseteq \{1, \ldots, n\}$. Definiere $d \in \{0,1\}^n$ durch $d_i = [i \in R]$. Da A TU ist, ist auch $\begin{pmatrix} A \\ -A \\ I_n \end{pmatrix}$ TU. Wegen Theorem 151 ist das Polytop

$$P := \{x \in \mathbb{R}^n | Ax \leqslant \left\lceil \frac{1}{2} Ad \right\rceil, Ax \geqslant \left\lfloor \frac{1}{2} Ad \right\rfloor, x \leqslant d, x \geqslant 0\}$$

ganzzahlig. Wegen $\frac{1}{2} d \in P$ ist $P \neq \varnothing$. Sei z eine ganzzahlige Ecke von P. Dann gilt für jedes $i \in \{1, \ldots, m\}$:

$$\sum_{j=1}^{n} a_{ij} z_j \leqslant \left\lceil \frac{1}{2} \sum_{j=1}^{n} a_{ij} d_j \right\rceil \leqslant \frac{1}{2} + \frac{1}{2} \sum_{j=1}^{n} a_{ij} d_j$$

sowie

$$\sum_{j=1}^{n} a_{ij} z_j \geqslant \left\lfloor \frac{1}{2} \sum_{j=1}^{n} a_{ij} d_j \right\rfloor \geqslant -\frac{1}{2} + \frac{1}{2} \sum_{j=1}^{n} a_{ij} d_i$$

Folglich ist

$$-1 \leqslant \sum_{j=1}^{n} a_{ij} (d_j - 2 z_j) \leqslant 1$$

Definiere $R_1 := \{j | d_j - 2z_j = 1\}$, $R_2 := \{j | d_j - 2z_j = -1\}$.

" $\Longleftarrow$ "

Angenommen für jedes $R \subseteq \{1, \ldots, n\}$ gibt es R_1, R_2 mit den geforderten Eigenschaften.

Behauptung 1. *Jede $k \times k$-Untermatrix von A hat Deterimiante* $\in \{-1, 0, 1\}$

Unterbeweis. Induktion nach k. $k = 1$ ist klar.

Sei $k > 1$ und die Aussage für alle $k' < k$ bereits gezeigt. Sei $B = (b_{ij})_{i,j \in \{1,\ldots,k\}}$ eine Untermatrix von A. (Eigentlich müssten das Doppelindices sein, der Einfachheit halber gehen wir davon aus,, dass ...

O.B.d.A. sei B regulär. Nach der Cramer'schen Regel ist jeder Eintrag von B^{-1} in der Form $\frac{\det(B')}{\det(B)}$, wobei B' aus B entsteht, indem eine Spalte durch einen Einheitsvektor ersetzt wird.

$\det(B')$ ist die Determinante einer $(k-1) \times (k-1)$ Untermatrix von A (Entwickle nach der Spalte, die durch den Einheitsvektor ersetzt wurde). Insbesondere ist nach Induktionsannahme $\det(B') \in \{-1, 0, 1\}$.

Insbesondere sind alle Einträge von $B^* := (\det(B))B^{-1}$ in $\{-1, 0, 1\}$. Sei b^* die erste Spalte von B^*. Es ist

$$Bb^* = \det(B)e_1$$

Definiere $R := \{j \in \{1, \ldots, k\} | b_j^* \neq 0\}$. Für $i \in \{2, \ldots, k\}$ ist $0 = (Bb^*)_i = \sum_{j \in R} b_{ij} b_j^*$. Folglich ist $|\{j \in R | b_{ij} \neq 0\}|$ gerade.

Sei $R = R_1 \sqcup R_2$ sodass $\sum_{j \in R_1} b_{ij} - \sum_{j \in R_2} b_{ij} \in \{-1, 0, 1\}$ für alle $i = 1, \ldots, k$.

Für $i = 2, \ldots, k$ ist die Anzahl der von 0 verschiedenen Summanden gerade, d.h. $\sum_{j \in R_1} b_{ij} - \sum_{j \in R_2} b_{ij} = 0$.

Da B regulär ist, ist $\sum_{j \in R_1} b_{1j} - \sum_{j \in R_2} b_{ij} \neq 0$, d.h. Element von $\{-1, 1\}$. (Anderefalls wären die Spalten linear abhängig.)

Für $x \in \{-1, 0, 1\}^k$ mit $x_j = [j \in R_1] - [j \in R_2]$ gilt $Bx \in \{e_1, -e_1\}$. Folglich ist $b^* = \det(B)B^-1e_1 \in \{\det(B)x, -\det(B)x\}$. b^* und x sind vom Nullvektor verschiedene Vektoren mit Einträgen $-1, 0, 1$. Folglich ist $\det(B) = \pm 1$. ■

□

10.5 Anwendung: Inzidenzmatrizen

Definition 155. Die **Inzidenzmatrix** eines Graphen G ist die Matrix $A_G = (a_{v,e})_{v \in V(G), e \in E(G)}$, wobei $a_{v,e} = [v \in e]$ im ungerichten Fall, bzw. $a_{v,(x,y)} = [v = x] - [v = y]$ im gerichteten Fall.

Beweis. klar ∎

Theorem 156. Die Inzidenzmatrix eines ungerichteten Graphen G ist genau dann TU, wenn G bipartit ist.

Beweis. trivial ∎

Corollary 157. Max-Flow und Min-Cost-Flow Probleme bei ganzzahligen b-Werten und Kapazitäten haben immer eine ganzzahlige Lösung.

Corollary 158 (König)**.** In einem bipartiten Graphen ist die Größe einer minimalen Knotenüberdeckung gleich der Größe eines maximalen Matching ist.

Beweis. Das Problem, ein maximales Matching zu finden, kann wie folgt als LP formuliert werden:

$$\begin{aligned} \max \ & \mathbb{1}_{E(G)}^t x \\ \text{s.d.} \quad & x \in \{0,1\}^{E(G)} \\ & A_G^t x \leqslant \mathbb{1}_{V(G)} \end{aligned}$$

Das Duale des relaxierten LP ist:

$$\begin{aligned} \min \ & \mathbb{1}_{V(G)}^t y + \mathbb{1}_{E(G)}^t z \\ \text{s.d} \quad & A_G y + z \geqslant \mathbb{1}_{E(G)} \\ & y \geqslant 0 \\ & z \geqslant 0 \end{aligned}$$

Die ist eine Relaxierung der LP-Formulierung des Problemes, eine minimale Knotenüberdeckung zu finden. Da G bipartit ist, ist A_G TU (Theorem 156) und mit Theorem 152 folgt die Aussage.

∎

Stichwortverzeichnis