Mathematische Optimierung I

Dieter Rautenbach
Wintersemester 2004/2005[1]

[1]Dieses vorläufige Skript enstand parallel zur Vorlesung und enthält fast keine Beweise. Diese finden sich in der zitierten Literatur und werden in einer späteren Fassung des Skriptes ausgearbeitet.

Inhaltsverzeichnis

0 Formalitäten

Vorlesungen
Jeweils Dienstags und Freitags 14:15 Uhr bis 15:45 Uhr.

Übungen
Jeweils Freitags 16:15 Uhr bis 17:45 Uhr.
Die Übungen leitet Jens Maßberg (massberg@or.uni-bonn.de).

Sprechstunden
Nach Vereinbarung.

Scheinkriterien
Jeweils 50% der erreichbaren Punkte in den Übungen vor bzw. nach den Weihnachtsferien.
Die Übungen können in Kleingruppen von bis zu 3 Personen bearbeitet werden.
Aktive Mitarbeit in den Übungen.
Ggf. Programmieraufgaben.

1 Einleitung

In beiden Teilen der Vorlesung geht es um Optimierungsprobleme der Form

$$\inf_{x \in X} f(x) = \inf\{f(x) \mid x \in X\}. \tag{1}$$

Typischerweise ist f hierbei eine konvexe Funktion und X eine konvexe Menge und wir schreiben oft $\min\{f(x) \mid x \in X\}$ statt $\inf\{f(x) \mid x \in X\}$.

Im ersten Teil behandeln wir hauptsächlich sog. lineare Programme, d.h. Probleme des Typs (1), für die f eine lineare Funktion ist und X durch ein System linearer Ungleichungen beschrieben wird.

Beispiel 1.1 *2-dimensionale lineare Optimierungsprobleme lassen sich noch konstruktiv geometrisch lösen (→ Schule).*

Für 3-dimensionale lineare Programme wie z.B.

$$\begin{array}{lll} \max & x+y+2z & \\ s.th. & x+z & \leq \ 6 \\ & y+z & \leq \ 6 \\ & 3y-z & \leq \ 6 \\ & x & \leq \ 3 \\ & x,y,z & \geq \ 0 \end{array}$$

ist die nicht mehr möglich.

Die Menge der zulässigen Punkte ist die "konvexe Hülle" der Punkte

$$E = \{(0,0,0),(3,0,0),(3,2,0),(0,2,0),(3,0,3),(3,3,3),(0,3,3),(0,0,6)\}$$

Alle Punkte für die $f(x,y,z) = x+y+2z$ konstant ist, liegen auf Ebenen mit Normalenvektor $(1,1,2)$. An einer Skizze erkennt man leicht, daß $f(3,3,3) = 12$ der optimale Wert ist und $(3,3,3)$ der einzige Punkt ist, in dem dieser Wert angenommen wird.

Multiplizieren wir die ersten 4 Ungleichungen mit nicht-negativen Koeffizienten $a,b,c,d \geq 0$ und bilden die Summe der entstehenden Ungleichungen, so ergibt sich

$$(a+d)x + (b+3c)y + (a+b-c)z \leq 6(a+b+c) + 3d.$$

Können wir nun $a,b,c,d \geq 0$ so wählen, daß

$$\begin{array}{ll} a+d & \geq \ 1 \\ b+3c & \geq \ 1 \\ a+b-c & \geq \ 2 \end{array}$$

gilt, so erhalten wir $6(a+b+c)+3d$ als obere Schranke für den optimalen Wert von $f(x,y,z)$.

Die beste obere Schranke, die wir auf diese Weise erzeugen können ist offenbar die optimale Lösung von folgendem linearen Programm.

$$
\begin{array}{lll}
\min & 6(a+b+c)+3d & \\
s.th. & a+d & \geq 1 \\
 & b+3c & \geq 1 \\
 & a+b-c & \geq 2 \\
 & a,b,c,d & \geq 0.
\end{array}
$$

Dieses neue lineare Programm nennt man das zu obigem LP "duale lineare Programm" und das obige LP nennt man das "primale lineare Programm". In unserem Beispiel ergibt $a = b = 1$ *und* $c = d = 0$ *sofort* $f(x, y, z) \leq 12$*, so daß 12 tatsächlich der optimale Wert ist, d.h. die optimalen Werte von primalem und dualem LP sind identisch.*

Aus obigem Beispiel ergibt sich, welche Themen wir genauer behandeln sollten, um solche Probleme allgemein zu lösen:

- Konvexe Mengen
- Systeme linearer Ungleichungen
- Allgemeine Aussagen über lineare Programme
- Polyeder

Nach diesen Themen wenden wir uns dann Verfahren zu, die lineare Programme lösen

- Simplex-Verfahren (Dantzig 1947)
- Ellipsoid-Methode (Khachiyan 1979)
- Innere Punkt Methoden (Kamarkar 1984)

1.1 Beispiele linearer Programme

Beispiel 1.2 (**Diätproblem**) Eine Einheit des Nahrungsmittels i (Brot, Butter, Bananen...) enthalte $a_{i,j}$ Einheiten vom Stoff j (Kalorien, Kohlenhydrate, Koffein,...) und koste c_i (Dollar, Dinar, Dublonen,...). Wie muss man sich ernähren, um vom Stoff j jeweils mindestens b_j Einheiten zu bekommen und möglichst wenig dafür auszugeben? Nimmt man vom Nahrungsmittels i jeweils x_i Einheiten zu sich, so erhält man folgendes LP.

$$
\begin{array}{lll}
\min & c^T x & \\
\text{s.th.} & \sum_i a_{i,j} x_i & \geq b_j \; \forall j \\
 & x_i & \geq 0 \; \forall i.
\end{array}
$$

Beispiel 1.3 (**Chebyshev Zentrum eines Polyeders**) Wir wollen die größte n-dimensionale euklidische Kugel

$$
B = \{x \in \mathbb{R}^n \mid ||x - x_c|| \leq r\} = \{x_c + u \mid ||u|| \leq r\}
$$

bestimmen, die ganz in der Menge

$$P = \{x \in \mathbb{R}^n \mid a_i^T x \leq \beta_i, i = 1, 2, ..., m\}$$

($a_i \in \mathbb{R}^n$, $\beta_i \in \mathbb{R}$) enthalten ist.

Für alle $1 \leq i \leq m$ muss $a_i^T(x_c + u) \leq \beta_i$ für alle $u \in \mathbb{R}^n$ mit $||u|| \leq r$ gelten. Da

$$\sup\{a_i^T u \mid ||u|| \leq r\} = a_i^T r \frac{a_i}{||a_i||} = r||a_i||$$

gilt, erhalten wir

$$\begin{array}{lrcl} \max & r & & \\ \text{s.th.} & a_i^T x_c & \leq & \beta_i - r||a_i|| \; \forall i. \end{array}$$

Beispiel 1.4 (**Wirtschaftsplan**) Für $j \in \{1, 2, ..., n\}$ (Aktivität) und $t \in \{1, 2, ..., N\}$ (Zeitperiode, Tag) messe $x_j(t) \geq 0$ die Intensität der Aktivität j in der Zeitperiode t.

Eine Einheit der Aktivität j erzeugt eine Menge $a_{i,j}$ des Gutes i und verbraucht eine Menge $b_{i,j}$ des Gutes i. In der Zeitperiode t werden daher die Menge $\sum_j a_{i,j} x_j(t) = (Ax(t))_i$ an Gut i erzeut und die Menge $\sum_j b_{i,j} x_j(t) = (Bx(t))_i$ an Gut i verbraucht. Zu Beginn stehe eine Menge $(g_0)_i$ an Gut i zur Verfügung.

Es sollte möglichst gelten, daß

$$\begin{array}{rcl} s(0) & = & g_0 - Bx(1) \geq 0 \\ s(t) & = & Ax(t) - Bx(t+1) \geq 0, t = 1, 2, ..., N-1 \\ s(N) & = & Ax(N) \geq 0 \end{array}$$

Maximieren möchte man den gewichteten Überschuss

$$c^T s(0) + \gamma c^T s(1) + ... + \gamma^N c^T s(N).$$

Insgesamt erhält man

$$\begin{array}{llcl} \max & c^T s(0) + \gamma c^T s(1) + ... + \gamma^N c^T s(N) & & \\ \text{s.th.} & x(t) & \geq & 0 \forall t = 1, ..., N \\ & s(t) & \geq & 0 \forall t = 0, ..., N \\ & s(0) & = & g_0 - Bx(1) \\ & s(t) & = & Ax(t) - Bx(t+1) \forall t = 1, ..., N-1 \\ & s(N) & = & Ax(N). \end{array}$$

Beispiel 1.5 (**Minimierung einer stückweise linearen, konvexen Funktion**)

$$f(x) = \max_{1 \leq i \leq n} (a_i^T x + \beta_i),\; x \in \mathbb{R}^n \rightsquigarrow \begin{array}{lrcl} \min & t & & \\ \text{s.th.} & a_i^T x + \beta_i & \leq & t \forall i. \end{array}$$

2 Konvexe Mengen

Definition 2.1 *(i) Eine Menge $M \subseteq \mathbb{R}^n$ ist genau dann konvex, wenn*

$$\lambda x_1 + (1-\lambda)x_2 \in M$$

für alle $x_1, x_2 \in M$ und $\lambda \in [0,1]$ gilt.

(ii) Eine Menge $M \subseteq \mathbb{R}^n$ ist genau dann ein Kegel, wenn

$$\lambda x \in M$$

für alle $x \in M$ und $\lambda \in \mathbb{R}_{\geq 0}$ gilt.

(iii) Eine Menge $M \subseteq \mathbb{R}^n$ ist genau dann ein konvexer Kegel, wenn

$$\lambda_1 x_1 + \lambda_2 x_2 \in M$$

für alle $x_1, x_2 \in M$ und $\lambda_1, \lambda_1 \in \mathbb{R}_{\geq 0}$ gilt.

Beispiel 2.2 *(i) Für eine Indexmenge I, $a_i \in \mathbb{R}^n$ und $\beta_i \in \mathbb{R}$ für $i \in I$ ist $M = \{x \mid a_i^T x \leq \beta_i, i \in I\}$ konvex.*

(ii) Für eine beliebige Norm $||\cdot||_?$ auf $\mathbb{R}^n$ ist $M = \{x \mid ||x||_? \leq 1\}$ konvex.

(iii) Für eine symmetrische, positiv definite Matrix $Q \in \mathbb{R}^{n\times n}$, $a \in \mathbb{R}^n$ und $r > 0$ ist $M = \{x \mid (x-a)^T Q(x-a) \leq r^2\}$ konvex. ($||x||_Q = \sqrt{x^T Q x}$ definiert eine Norm.)

(iv) Ist $M \subseteq \mathbb{R}^n$, $r > 0$ und $||\cdot||_?$ eine beliebige Norm auf $\mathbb{R}^n$, so ist

$$M_r = \{y \mid \inf_{x\in M} ||x-y||_? \leq r\}$$

konvex.

(v) Sind $X, Y \subseteq \mathbb{R}^n$ konvex, $a \in \mathbb{R}^n$ und $\mu \in \mathbb{R}$, so sind auch

$$\begin{aligned} X + a &= \{x + a \mid x \in X\} \\ \mu X &= \{\mu x \mid x \in X\} \\ X + Y &= \{x + y \mid x \in X, y \in Y\} \end{aligned}$$

konvex.

Bemerkung 2.3 *Ist $M \subseteq \mathbb{R}^n$ konvex, $x_1, x_2, ..., x_t \in M$ und $\lambda_1, \lambda_2, ..., \lambda_t \geq 0$ mit $\sum_{i=1}^t \lambda_i = 1$, dann gilt $\sum_{i=1}^t \lambda_i x_i \in M$.*

Definition 2.4 *Für eine Menge $X \subseteq \mathbb{R}^n$ sind die lineare Hülle* $\mathrm{lin}(X)$*, die affine Hülle* $\mathrm{aff}(X)$*, die konvexe Hülle* $\mathrm{conv}(X)$ *und der durch X erzeugte konvexe Kegel cone(X) wie* *folgt definiert*

$$
\begin{aligned}
\operatorname{lin}(X) &= \left\{ \sum_{i=1}^{t} \lambda_i x_i \mid t \in \mathbb{Z}_{\geq 0} \textit{ und } x_i \in X, t_i \in \mathbb{R} \textit{ für } 1 \leq i \leq t \right\} \\
\operatorname{aff}(X) &= \left\{ \sum_{i=1}^{t} \lambda_i x_i \mid t \in \mathbb{Z}_{\geq 0} \textit{ und } x_i \in X, t_i \in \mathbb{R} \textit{ für } 1 \leq i \leq t \textit{ mit } \sum_{i=1}^{t} \lambda_i = 1 \right\} \\
\operatorname{conv}(X) &= \left\{ \sum_{i=1}^{t} \lambda_i x_i \mid t \in \mathbb{Z}_{\geq 0} \textit{ und } x_i \in X, t_i \in \mathbb{R}_{\geq 0} \textit{ für } 1 \leq i \leq t \textit{ mit } \sum_{i=1}^{t} \lambda_i = 1 \right\} \\
\operatorname{cone}(X) &= \left\{ \sum_{i=1}^{t} \lambda_i x_i \mid t \in \mathbb{Z}_{\geq 0} \textit{ und } x_i \in X, t_i \in \mathbb{R}_{\geq 0} \textit{ für } 1 \leq i \leq t \right\}.
\end{aligned}
$$

Satz 2.5 *Für eine Menge $X \subseteq \mathbb{R}^n$ ist $\operatorname{conv}(X)$ die kleinste konvexe Menge, die X enthält, d.h. $Y = \operatorname{conv}(X)$ gilt genau dann, wenn folgende drei Eigenschaften erfüllt sind.*

(i) Y ist konvex.

(ii) $X \subseteq Y$.

(iii) $Y \subseteq Z$ für alle konvexen Mengen Z mit $X \subseteq Y$.

Satz 2.6 *(i) Der Durchschnitt beliebig vieler konvexer Mengen ist konvex.*

(ii) Für $X \subseteq \mathbb{R}^n$ gilt

$$\operatorname{conv}(X) = \bigcap_{X \subseteq Y, Y \textit{ konvex}} Y.$$

Definition 2.7 *Seien $x \in \mathbb{R}^n$ und $M \subseteq \mathbb{R}^n$.*

(i) Die Euklidische Norm $||x||$ von $x = (x_1, ..., x_n)$ ist definiert durch

$$||x|| = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

(ii) Für $\epsilon > 0$ ist die (offene) ϵ-Umbegung $U_\epsilon(x)$ von x definiert durch

$$U_\epsilon(x) = \{y \in \mathbb{R}^n \mid ||x - y|| < \epsilon\}.$$

(iii) x ist genau dann ein innerer Punkt von M, falls $U_\epsilon(x) \subseteq M$ für ein $\epsilon > 0$ gilt.

(iv) x ist genau dann ein Häufungspunkt von M, falls $(U_\epsilon(x) \setminus \{x\}) \cap M \neq \emptyset$ für alle $\epsilon > 0$ gilt.

(v) Der Abschluss $\overline{M}$ von M ist die Vereinigung von M mit allen Häufungspunkten von M.

(vi) Das Innere $\operatorname{Inn}(M)$ von M ist die Menge aller inneren Punkte von M.

(vii) Der Rand $\delta(M)$ von M ist die Menge $\overline{M} \setminus \operatorname{Inn}(M)$.

Satz 2.8 *Ist $M \subseteq \mathbb{R}^n$ konvex, so sind auch $\overline{M}$ und $\operatorname{Inn}(M)$ konvex.*

2.1 Trennsätze

Definition 2.9 *Für $a \in \mathbb{R}^n$ und $\beta \in \mathbb{R}$ nennt man*

$$H = \{x \in \mathbb{R}^n \mid a^T x = \beta\}$$

eine Hyperebene und

$$H^+ = \{x \in \mathbb{R}^n \mid a^T x \geq \beta\}$$

bzw.

$$\{x \in \mathbb{R}^n \mid a^T x > \beta\}$$

einen abgeschlossenen bzw. offenen Halbraum.

Satz 2.10 *Ist $M \subseteq \mathbb{R}^n$ konvex, abgeschlossen und nicht-leer und $0 \notin M$, so existiert ein eindeutiges $x_0 \in M$ mit $||x_0|| = \min\{||x|| \mid x \in M\}$, d.h. die Projektion von 0 auf M existiert und ist eindeutig bestimmt.*

Satz 2.11 *Ist $M \subseteq \mathbb{R}^n$ konvex und $0 \notin \overline{M}$, so existiert ein $a \in \bar{M}$ mit $a^T x \geq a^T a > 0$ für alle $x \in M$.*

Definition 2.12 *Der Abstand $d(X, Y)$ zweier nicht-leerer Mengen $X, Y \subseteq \mathbb{R}^n$ ist definiert durch*

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

Satz 2.13 *Sind $X, Y \subseteq \mathbb{R}^n$ konvex und nicht-leer und gilt $d(X, Y) > 0$, so existieren $a \in \mathbb{R}^n$ und $\beta_1, \beta_2 \in \mathbb{R}$ mit*

$$a^T x > \beta_1 > \beta_2 > a^T y$$

für alle $x \in X$ und $y \in Y$.

Satz 2.14 *Ist $M \subseteq \mathbb{R}^n$ konvex und $0 \in \delta(M)$, so existiert ein $a \in \mathbb{R}^n$ mit $a^T x \geq 0$ für alle $x \in \overline{M}$.*

Satz 2.15 *Sind $X, Y \subseteq \mathbb{R}^n$ konvex und nicht-leer und gilt $0 \in \delta(X - Y)$, so existieren $a \in \mathbb{R}^n$ und $\beta \in \mathbb{R}$ mit $a^T x \geq \beta$ für alle $x \in \overline{X}$ und $a^T y \leq \beta$ für alle $y \in \overline{Y}$.*

Folgerung 2.16 *Ist $M \subseteq \mathbb{R}^n$ konvex und $b \in \delta(M)$, so existieren $a \in \mathbb{R}^n$ mit $a^T x \geq a^T b$ für alle $x \in \overline{M}$.*

Definition 2.17 *Eine Hyperebene $H = \{x \in \mathbb{R}^n \mid a^T x = \beta\}$ heißt Stützhyperebene von $M \subseteq \mathbb{R}^n$ genau dann, wenn $a^T x \geq \beta$ für alle $x \in \overline{M}$ gilt und ein $x_0 \in \overline{M}$ mit $a^T x = \beta$ existiert.*

Satz 2.18 *Ist $M \subseteq \mathbb{R}^n$ konvex, so gilt*

$$\overline{M} = \bigcap_{H \text{ ist Stützhyperebene von } M} H^+.$$

2.2 Ecken

Definition 2.19 *Ist $M \subseteq \mathbb{R}^n$ konvex und $x \in M$, so nennt man x eine Ecke von M, falls aus $x = \lambda y + (1-\lambda)z$ für $0 < \lambda < 1$ und $y, z \in M$ folgt, daß $x = y = z$ gilt.*

Bemerkung 2.20 *Ist $M \subseteq \mathbb{R}^n$ konvex und $x \in M$, so ist x genau dann eine Ecke von M, wenn*

(i) aus $x = \lambda y + (1-\lambda)z$ für $0 \leq \lambda \leq 1$ und $y, z \in M$ mit $y \neq z$ folgt, daß $\lambda \in \{0, 1\}$ gilt.

(ii) aus $x = \frac{1}{2}(y+z)$ für $y, z \in M$ folgt, daß $x = y = z$ gilt.

Satz 2.21 *Sei $M \subseteq \mathbb{R}^n$ konvex und $x \in M$. x genau dann eine Ecke von M, wenn $M \setminus \{x\}$ konvex ist.*

Satz 2.22 *Jede nicht-leere, konvexe und kompakte Menge besitzt mindestens eine Ecke.*

Satz 2.23 *Jede Stützhyperebene einer nicht-leeren, konvexen und kompakten Menge M enthält mindestens eine Ecke von M.*

Satz 2.24 (Krein und Milman) *Jede nicht-leere, konvexe und kompakte Menge ist Abschluss der konvexen Hülle der Menge ihrer Ecken.*

Gilt auch ohne Abschluss (vgl. Corollary 18.5.1 in [16]).

Satz 2.25 *Ist $M \subseteq \mathbb{R}^n$ und x eine Ecke von* $\mathrm{conv}(M)$, *so gilt $x \in M$.*

Definition 2.26 *Die Vektoren $x_1, x_2, ..., x_l \in \mathbb{R}^n$ heißen affin unabhängig, falls keine Zahlen $\lambda_1, ..., \lambda_l \in \mathbb{R}$ existieren, die nicht alle gleich 0 sind und*

(i) $0 = \sum_{i=1}^{l} \lambda_i x_i$ und

(ii) $0 = \sum_{i=1}^{l} \lambda_i$ gilt.

Bemerkung 2.27 *Die Vektoren $x_1, x_2, ..., x_l \in \mathbb{R}^n$ sind genau dann affin unabhängig, wenn die Vektoren $x_2 - x_1, x_3 - x_1, ..., x_l - x_1$ linear unabhängig sind.*

Satz 2.28 (Carathéodory 1911) *Ist $M \subseteq \mathbb{R}^n$, so läßt sich jedes $x \in$* $\mathrm{conv}(M)$ *als Konvexkombination von affin unabhängigen Vektoren aus M schreiben, d.h. es existieren affin unabhängige $x_1, ..., x_m \in M$ und $\lambda_1, ..., \lambda_m \in \mathbb{R}_{\geq 0}$ mit*

$$x = \sum_{i=1}^{m} \lambda_i x_i$$

und

$$1 = \sum_{i=1}^{m} \lambda_i.$$

Folgerung 2.29 *Ist $M \subseteq \mathbb{R}^n$, so läßt sich jedes $x \in \operatorname{conv}(M)$ als Konvexkombination von höchstens $n+1$ Vektoren aus M schreiben.*

Satz 2.30 (Carathéodory 1911) *Ist $M \subseteq \mathbb{R}^n$, so läßt sich jedes $x \in \operatorname{cone}(M)$ als nicht-negative Linearkombination von linear unabhängigen Vektoren aus M schreiben, d.h. es existieren linear unabhängige $x_1, ..., x_m \in M$ und $\lambda_1, ..., \lambda_m \in \mathbb{R}_{\geq 0}$ mit*

$$x = \sum_{i=1}^{m} \lambda_i x_i.$$

Folgerung 2.31 *Ist $M \subseteq \mathbb{R}^n$, so läßt sich jedes $x \in \operatorname{cone}(M)$ als positive Linearkombination von höchstens n Vektoren aus M schreiben.*

Satz 2.32 *Ist $M \subseteq \mathbb{R}^n$, so läßt sich jedes $x \in \delta(\operatorname{conv}(M))$ als Konvexkombination von höchstens n Vektoren aus M schreiben.*

Satz 2.33 (Radon) *Jede Menge X von mindestens $n+2$ Vektoren im $\mathbb{R}^n$ kann so in zwei Teilmengen X_1, X_2 partitioniert werden ($X = X_1 \cup X_2$, $X_1 \cap X_2 = \emptyset$), daß $\operatorname{conv}(X_1) \cap \operatorname{conv}(X_2) \neq \emptyset$ gilt.*

Satz 2.34 (Helly) *Sind $M_1, M_2, ..., M_m \subseteq \mathbb{R}^n$ konvex und haben jeweils $n+1$ der Mengen einen nicht-leeren Schnitt, so gilt $M_1 \cap M_2 \cap ... \cap M_m \neq \emptyset$.*

3 Systeme linearer Ungleichungen

Satz 3.1 (Lemma von Farkas (1894, 1899)) *Sind $a_1, a_2, ..., a_m, b \in \mathbb{R}^n$ so gilt genau eine der folgenden beiden Alternativen*

(i) Es existiert ein $x \in \mathbb{R}^n$ mit $a_i^T x \leq 0$ für $1 \leq i \leq m$ und $b^T x > 0$.

(ii) $b \in \operatorname{cone}(\{a_1, a_2, ..., a_m\})$.

Satz 3.2 (Motzkin) *Für $m_a \in \mathbb{N}$ und $m_b, m_c \in \mathbb{Z}_{\geq 0}$ seien $a_i, b_j, c_k \in \mathbb{R}^n$ mit $1 \leq i \leq m_a$, $1 \leq j \leq m_b$ und $1 \leq k \leq m_c$. Sei (P) das folgende System linearer Ungleichungen*

$$\begin{array}{rcl} a_i^T x & < & 0,\ 1 \leq i \leq m_a \\ b_j^T x & \leq & 0,\ 1 \leq j \leq m_b \\ c_k^T x & = & 0,\ 1 \leq k \leq m_c. \end{array}$$

Es gilt genau eine der folgenden zwei Alternativen

(i) (P) has eine Lösung.

(ii) Es existieren Zahlen $u_1, ..., u_{m_a} \in \mathbb{R}_{\geq 0}$ mit $\max\{u_1, ..., u_{m_a}\} > 0$, $v_1, ..., v_{m_b} \in \mathbb{R}_{\geq 0}$ und $w_1, ..., w_{m_c} \in \mathbb{R}$ mit

$$0 = \sum_{i=1}^{m_a} u_i a_i + \sum_{j=1}^{m_b} v_j b_j + \sum_{k=1}^{m_c} w_k c_k.$$

Folgerung 3.3 (Lemma von Farkas, inhomogene Variante) *Seien $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}^m$ und $\beta \in \mathbb{R}$. Es gilt genau eine der folgenden zwei Alternativen*

(i) Es existiert ein $x \in \mathbb{R}^n$ mit $Ax \leq c$ und $b^T x > \beta$.

(ii) $\{y \in \mathbb{R}^m \mid y^T A = b^T, y^T c \leq \beta, y \geq 0\} \cup \{y \in \mathbb{R}^m \mid y^T A = 0, y^T c < 0, y \geq 0\} \neq \emptyset$.

Folgerung 3.4 (Lemma von Farkas, Variante) *Sind $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^n$ so gilt genau eine der folgenden beiden Alternativen*

(i) $\{x \in \mathbb{R}^n \mid Ax \leq b\} \neq \emptyset$.

(i) $\{y \in \mathbb{R}^m \mid y^T A = 0, y \geq 0, y^T b < 0\} \neq \emptyset$.

Folgerung 3.5 (Lemma von Farkas, Variante) *Sind $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^n$ so gilt genau eine der folgenden beiden Alternativen*

(i) $\{x \in \mathbb{R}^n \mid Ax \leq b, x \geq 0\} \neq \emptyset$.

(i) $\{y \in \mathbb{R}^m \mid y^T A \geq 0, y \geq 0, y^T b < 0\} \neq \emptyset$.

Folgerung 3.6 (Lemma von Farkas, Variante) *Ist $A \in \mathbb{R}^{m \times n}$ so gilt genau eine der folgenden beiden Alternativen*

(i) $\{x \in \mathbb{R}^n \mid Ax \geq 0, Ax \neq 0\} \neq \emptyset$.

(i) $\{y \in \mathbb{R}^m \mid y^T A = 0, y > 0\} \neq \emptyset$.

3.1 Fourier-Motzkin Elimination

(Fourier 1827, Dines 1918-19, Motzkin 1936) Zu einer gegebenen Matrix $A \in \mathbb{R}^{m\times n}$ und einem Vektor $b \in \mathbb{R}^m$ sei (P) folgendes System von Ungleichungen

$$(P) : Ax \leq b.$$

Multipliziert man die m Ungleichungen von (P) mit passenden Zahlen ≥ 0 ergibt sich ein äquivalente System (P') mit

$$\begin{array}{rcl} & (P') & \\ x_1 + a_i^T x' & \leq & \beta_i, 1 \leq i \leq m' \\ -x_1 + a_i^T x' & \leq & \beta_i, m'+1 \leq i \leq m'' \\ a_i^T x' & \leq & \beta_i, m''+1 \leq i \leq m \end{array}$$

wobei die a_i passende Vektoren aus $\mathbb{R}^{n-1}$ sind und $x = (x_1, x')$ gilt.

Da die ersten m'' Ungleichungen von (P') zu

$$\max_{m'+1\leq i\leq m''}(a_i^T x' - \beta_i) \leq x_1 \leq \min_{1\leq i\leq m'}(\beta_i - a_i^T x')$$

äquivalent sind, kann man x_1 aus (P') eliminieren.

$$\begin{array}{rcl} (P'') & & \\ a_j^T x' - \beta_j & \leq & \beta_i - a_i^T x', 1 \leq i \leq m', m'+1 \leq j \leq m'' \\ a_i^T x' & \leq & \beta_i, m''+1 \leq i \leq m. \end{array}$$

(P'') ist wiederum äquivalent zu (P''') mit

$$\begin{array}{rcl} (P''') & & \\ (a_j + a_i)^T x' & \leq & \beta_j + \beta_i, 1 \leq i \leq m', m'+1 \leq j \leq m'' \\ a_i^T x' & \leq & \beta_i, m''+1 \leq i \leq m. \end{array}$$

Das System (P''') besteht aus $m'(m''-m') + m - m''$ vielen Ungleichung für $n-1$ viele Variablen. Die Eliminierung von x_1 entspricht der Projektion von $\{x \mid Ax \leq b\}$ entlang der x_1-Achse.

Mit Hilfe der Fourier-Motzkin Elimination ist ein alternativer Beweis von Folgerung 3.4 möglich (vgl. [17], Seite 156).

Bemerkung 3.7 *Die Fourier-Motzkin Elimination ist kein polynomielles Verfahren (vgl. [17], Seite 156).*

3.2 Relaxierungs-Methode

(Agmon 1954, Motzkin und Schoenberg 1954) Zu einer gegebenen Matrix $A \in \mathbb{R}^{m \times n}$ und einem Vektor $b \in \mathbb{R}^m$ sei (P) wieder folgendes System von Ungleichungen

$$(P) : Ax \leq b.$$

Die Methode bestimmt ausgehend von einem beliebigen Startpunkt $z^0 \in \mathbb{R}^n$ und mittels eines festen Parameters $\lambda > 0$ eine Folge $z^1, z^2, ...$ von Punkten und terminiert genau dann mit z^t, falls z^t der erste Punkt ist, der alle Ungleichungen von (P) erfüllt.

Ist z^i für ein $i \geq 0$ bereits bestimmt und erfüllt z^i nicht alle Ungleichungen von (P), so sei $a^T x \leq \beta$ eine verletzte Ungleichung und

$$z^{i+1} = z^i + \lambda \frac{1}{||a||^2} (\beta - a^T z^i) a.$$

Für $\lambda = 1$ ist z^{i+1} die Projektion von z^i auf die Hyperebene $H = \{x \mid a^T x = \beta\}$ und für $\lambda = 2$ ist z^{i+1} die Spiegelung von z^i an H.

Satz 3.8 (Motzkin und Schoenberg 1954) *Hat $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ ein nicht-leeres Inneres und wählt man zur Bestimmung von z^{i+1} bei der Relaxierungs-Methode $\lambda = 2$ und jeweils eine verletzte Ungleichung, für die $\frac{a^T z^i - \beta}{||a||}$ maximal ist, so terminiert die Relaxierungs-Methode für einen beliebigen Startpunkt z^0.*

Satz 3.9 (Hoffman 1952, Goffin 1980, Jeroslow 1979) *Ist $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ nicht-leer, $0 < \lambda < 2$, z^0 beliebig und wählt man zur Bestimmung von z^{i+1} bei der Relaxierungs-Methode jeweils eine verletzte Ungleichung, für die $\frac{a^T z^i - \beta}{||a||}$ maximal ist, so terminiert die Relaxierungs-Methode oder die Folge $z^1, z^2, ...$ konvergiert gegen einen Punkt z^* in $\delta(P)$ mit $||z^i - z^*|| \leq \mu \theta^i$ für $\mu \geq 0$ und $0 < \theta < 1$ und alle $i \in \mathbb{N}$.*

Bemerkung 3.10 *Wie schon die Fourier-Motzkin Elimination so ist auch die Relaxierungs-Methode kein polynomielles Verfahren. Dies sieht man an folgendem Beispiel von Goffin (1982) und Todd (1979).*

4 Die Hauptsätze der linearen Programmierung

Zu einer Matrix $A \in \mathbb{R}^{m\times n}$ und Vektoren $b \in \mathbb{R}^m$ und $c \in \mathbb{R}^n$ sei

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

und (P) das folgende Optimierungsproblem

$$(P) : \max\{c^T x \mid x \in P\} = \max\{c^T x \mid Ax \leq b\}.$$

Zu (P) betrachten wir das sogenannte “duale” Problem (D)

$$(D) : \min\{b^T y \mid y \in D\} = \min\{b^T y \mid y^T A = c^T, y \geq 0\}.$$

mit

$$D = \{y \in \mathbb{R}^m \mid y^T A = c^T, y \geq 0\}.$$

Elemente von P bzw. D nennt man zulässige Lösungen für (P) bzw. (D). Elemente $x' \in P$ bzw. $y' \in D$ mit

$$c^T x' = \max\{c^T x \mid x \in P\}$$

bzw.

$$b^T y' = \min\{b^T y \mid y \in D\}$$

nennt man (optimale) Lösungen.

Satz 4.1 (Dualitätssatz der Linearen Optimierung, von Neumann 1947, Gale, Kuhn und Tucker 1951) *Sind P und D wie oben und beide Mengen nicht leer, so gilt*

$$\min\{b^T y \mid y \in D\} = \max\{c^T x \mid x \in P\}.$$

Satz 4.2 *Sind (P) und (D) wie oben so gilt genau eine der folgenden Möglichkeiten.*

(i) (P) und (D) besitzen beide Lösungen (mit gleichem Wert nach Satz 4.1).

(ii) $P = D = \emptyset$.

(iii) $P = \emptyset$ und $\inf\{b^T y \mid y \in D\} = -\infty$.

(iv) $D = \emptyset$ und $\sup\{c^T x \mid x \in P\} = \infty$.

Satz 4.3 (Charaktersierungssatz der Linearen Optimierung) *Sei (P) wie oben. Ein Vektor $x \in \mathbb{R}^n$ mit $Ax \leq b$ ist genau dann optimale Lösung von (P), wenn ein $y \in \mathbb{R}^m$ existiert mit $c^T = y^T A$, $y \geq 0$ und $y^T(Ax - b) = 0$.*

Folgerung 4.4 (Satz von komplementären Schlupf) *Besitzt eines der Probleme (P) und (D) wie oben eine optimale Lösung so auch das andere und es gilt für $x \in P$ und $y \in D$, daß x und y genau dann optimale Lösungen von (P) und (D) sind, wenn $y^T(b - Ax) = 0$ gilt.*

Satz 4.5 (Satz von (starken) komplementären Schlupf) *Es seien (P) und (D) wie oben.*

Existieren für (P) und (D) zulässige Lösungen, so existiert ein Paar x, y optimaler Lösungen, so daß $y_i > 0 \Leftrightarrow a_i^T x = b_i$ für $1 \leq i \leq m$ gilt.

Folgerung 4.6 (Affine Form des Lemma von Farkas, Haar 1918, Weyl 1935) *Es sei P wie oben und $P \neq \emptyset$. Erfüllen alle Elemente von P die Ungleichung $c^T x \leq \delta$, so existiert ein $\delta' \leq \delta$ für das die Ungleichung $c^T x \leq \delta$ nicht-negative Linearkombination der Ungleichungen $Ax \leq b$ ist.*

Weitere Paare von dualen linearen Programmen für die obige Sätze analog gelten sind

$$(P_1) : \max\{c^T x \mid Ax \leq b, x \geq 0\}$$

und

$$(D_1) : \min\{b^T y \mid y^T A \geq c^T, y \geq 0\}.$$

oder

$$(P_1) : \min\{c^T x \mid Ax \geq b, x \geq 0\}$$

und

$$(D_1) : \max\{b^T y \mid y^T A \leq c^T, y \geq 0\}.$$

5 Polyeder

Die Menge der zulässigen Lösungen für die linearen Programme (P) und (D) aus dem letzten Abschnitt nennt man Polyeder.

Definition 5.1 *(i) Ist $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$, so nennt man die Menge $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ einen Polyeder.*

(ii) Die konvexe Hülle endliche vieler Punkte des $\mathbb{R}^n$ nennt man Polytop.

Definition 5.2 *(i) Ist $A \in \mathbb{R}^{m \times n}$, so nennt man die Menge $P = \{x \in \mathbb{R}^n \mid Ax \leq 0\}$ einen polyhedralen Kegel.*

(ii) Einen durch eine endliche Menge erzeugten konvexen Kegel nennt man endlich erzeugten (konvexen) Kegel.

Satz 5.3 (Farkas 1898, 1902, Minkowski 1896, Weyl 1935) *Ein konvexer Kegel ist genau dann polyhedral, wenn er endlich erzeugt ist.*

Satz 5.4 (Fundamentalsatz über lineare Ungleichungen) *Sind $a_1, a_2, ..., a_m, b \in \mathbb{R}^n$, so gilt genau eine der beiden folgenden Alternativen.*

(i) b ist nicht-negative Linearkombination von linear unabhängigen Vektoren in

$$\{a_1, a_2, ..., a_m\}.$$

(ii) Es existiert eine Hyperebene $\{x \mid c^T x = 0\}$, die $t-1$ linear unabhängige Vektoren aus $\{a_1, a_2, ..., a_m\}$ enthält, so daß $c^T b < 0$ und $c^T a_i \geq 0$ für $1 \leq i \leq m$ gilt. Dabei ist t der Rang der Matrix $[a_1, a_2, ...a_m, b]$.

Satz 5.5 (Motzkin 1936) *Eine Menge $P \subseteq \mathbb{R}^n$ ist genau dann ein Polyeder, wenn $P = Q + C$ für ein Polytop Q und einen polyhedralen Kegel C gilt.*

Definition 5.6 *Es sei $P = \{x \mid Ax \leq b\}$ ein nicht-leerer Polyeder mit $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$.*

(i) Eine Ungleichung $a^T x \leq \beta$ in $Ax \leq b$ nennt man eine implizite Gleichung, wenn $a^T x = \beta$ für alle $x \in P$ gilt.

(ii) Der charakteristische Kegel (engl.: characteristic cone, recession cone) $\text{char.cone}(P)$ von P ist definiert durch

$$\text{char.cone}(P) = \{y \mid x + y \in P \textit{ für alle } x \in P\}.$$

(iii) Der Linearitätsraum (engl.: linearity space) $\text{lin.space}(P)$ von P ist definiert durch

$$\text{lin.space}(P) = \text{char.cone}(P) \cap (-\text{char.cone}(P))$$

$\text{lin.space}(P)$ *ist ein Vektorraum. Gilt* $\dim(\text{lin.space}(P)) = 0$, *so nennt man P spitz (engl.: pointed).*

(iv) Die Dimension von P ist die Dimension der affinen Hülle von P. Ist die Dimension von P gleich n, so nennt man P volldimensional (engl.: full-dimensional).

(v) Eine Seitenfläche (engl.: face) von P ist der Schnitt von P mit einer Stützhyperebene. Eine inklusions-maximale Seitenfläche von P, die von P verschieden ist, nennt man Facette (engl.: facet).

(vi) Man sagt, daß eine Ungleichung $c^T x \leq \delta$ eine Facette von P definiert, falls $c^T x \leq \delta$ für alle $x \in P$ gilt und $\{x \in P \mid c^T x = \delta\}$ eine Facette von P ist.

Satz 5.7 *Es sei $P = \{x \mid Ax \leq b\}$ ein nicht-leerer Polyeder mit $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$. Für $F \subseteq P$ sind folgende Aussagen äquivalent.*

(i) F ist eine Seitenfläche von P.

(ii) Es existiert ein $c \in \mathbb{R}^n$, so daß $\delta = \max\{c^T x \mid x \in P\}$ endlich ist und $F = \{x \in P \mid c^T x = \delta\}$.

(ii) $F = \{x \in P \mid A'x = b'\}$ für ein Teilsystem $A'x \leq b'$ von $Ax \leq b$.

Folgerung 5.8 *(i) Die Menge der optimalen Lösungen von*

$$(P) : \max\{c^T x \mid x \in P\}$$

für einen nicht-leeren, beschränkten Polyeder P und ein $c \in \mathbb{R}^n$ ist eine Seitenfläche von P.
(ii) Eine Seitenfläche F eines Polyeders P ist ein Polyeder. Eine Teilmenge von F ist genau dann eine Seitenfläche von P, wenn sie eine Seitenfläche von F ist.

Satz 5.9 *Es sei $P \subseteq \{x \mid Ax = b\}$ ein nicht-leerer Polyeder der Dimension $n - \operatorname{rang}(A)$. Sei $A'x \leq b'$ ein minimales Ungleichungssystem, so daß $P = \{x \mid Ax = b, A'x \leq b'\}$.*

Dann folgt, daß jede Ungleichung von $A'x \leq b'$ eine Facette von P definiert und jede Facette von P von einer Ungleichung von $A'x \leq b'$ definiert wird.

Satz 5.10 (Hoffman und Kruskal 1956) *Es sei $P = \{x \mid Ax \leq b\}$ ein Polyeder. Eine nicht-leere Menge $F \subseteq P$ ist genau dann eine (inklusions-)minimale Seitenfläche von P, wenn $F = \{x \mid A'x = b'\}$ für ein Teilsystem $A'x \leq b'$ von $Ax \leq b$ gilt.*

Folgerung 5.11 *Ist $P = \{x \mid Ax \leq b\}$ ein Polyeder so haben alle minimalen Seitenflächen von P die Dimension $n - \operatorname{rang}(A)$. Alle minimalen Seitenflächen eines Polytops sind Ecken.*

Satz 5.12 (Minkowski 1896) *Ist $C = \{x \mid Ax \leq 0\}$ ein polyhedraler Kegel mit $A \in \mathbb{R}^{m \times n}$, so wird C von einer Teilmenge der Lösungen der Systeme $My = b'$ erzeugt, wobei M aus n linear unabhängigen Zeilen von $[A^T \mid I^T]^T$ besteht und $b' = \pm e_j$ für einen Einheitsvektor e_j gilt.*

6 Das Simplexverfahren

Wir betrachten nun lineare Programme in der folgenden kanonischen Form.

$$(P)\quad \begin{array}{lll} \min & c^T x & \\ s.th. & Ax & = b \\ & x & \geq 0 \end{array}$$

für eine Matrix $A \in \mathbb{R}^{m \times n}$ und Vektoren $b \in \mathbb{R}^m$ und $c \in \mathbb{R}^n$. O.B.d.A. können wir annehmen, daß $\mathrm{rang}(A) = m \leq n$ gilt, da anderenfalls einige der Gleichungen in $Ax = b$ redundant sind und entfernt werden können.

Bemerkung 6.1 *Alle linearen Programme lassen sich in kanonischer Form schreiben, denn $a^T x \leq \beta$ ist äquivalent zu $a^T(x_+ - x_-) + y = \beta$ mit $x_+, x_-, y \geq 0$.*

Die Spalten von A bezeichnen wir mit $a^1, a^2, ..., a^n$, d.h.

$$A = [a^1 \mid a^2 \mid ... \mid a^n].$$

Die Zeilen von A bezeichnen wir mit $a_1, a_2, ..., a_m$, d.h.

$$A = [a_1^T \mid a_2^T \mid ... \mid a_m^T]^T.$$

Sei

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

der Polyeder der zulässigen Lösungen.

Zu einem $x \in P$ sei $Z \subseteq \{1, 2, ..., n\}$ die Menge der Indizes i mit $x_i > 0$. Man nennt $\{a^i \mid i \in Z\}$ die zu x gehörende Spaltenmenge. Es gilt natürlich $\sum_{i \in Z} x_i a^i = b$.

Folgerung 6.2 *Ist $P \neq \emptyset$, so besitzt P Ecken, d.h. insbesondere, daß alle minimalen Seitenflächen Ecken sind.*

Folgerung 6.3 *Ein Element $x \in P$ ist genau dann Ecke von P, wenn die zu x gehörende Spaltenmenge linear unabhängig ist.*

Folgerung 6.4 *Alle Ecken von P haben höchstens m positive Koordinaten.*

Definition 6.5 *Eine Ecke von P heißt nichtentartet, falls sie genau m positive Koordinaten besitzt. Anderenfalls heißt sie entartet.*

Zu einer Ecke x nennt man jede m-elementige Menge von linear unabhängigen Spalten von A, die die zu x gehörende Spaltenmenge enthält, eine Basis. Die Menge der Spaltenindizes nennt man Basisindexmenge.

Folgerung 6.6 *Die Menge der optimalen Lösungen von (P) ist ein Polyeder P_{opt}. Die Ecken von P_{opt} sind genau die Ecken von P, die in P_{opt} liegen.*

Folgerung 6.7 *Besitzt* (P) *eine optimale Lösung (ist also nicht unbeschränkt), so existiert eine optimale Lösung, die eine Ecke von* P *ist.*

Bemerkung 6.8 *Die obigen Sätze liefern schon ein Verfahren zur Lösung von* (P)*, falls* (P) *beschränkt ist: Man betrachtet alle* $\binom{n}{m}$ *Untermatrizen* A' *von* A*, die aus* m *Spalten von* A *bestehen. Besitzt* $A'x = b$ *eine Lösung* x*, so ist* x *Ecke von* P*. Da auf diese Weise alle Ecken entstehen und mindestens eine optimale Lösung unter diesen ist,...*

Diese Verfahren ist natürlich hoffnungslos ineffizient. Das Simplexverfahren modifiziert diese Methode, indem es sicherstellt, daß die Zielfunktionswerte der nacheinander betrachteten Ecken niemals abnehmen.

6.1 Ecken-Basis-Austausch

Wir beschreiben nun den wichtigsten Schritt des Simplexverfahrens, den Übergang von der Basis einer Ecke von P zu einer neuen Basis einer möglicherweise anderen Ecke von P, deren Zielfunktionswert nicht kleiner als der der bisherigen Ecke ist.

Sei $\bar{x} \in P$ eine Ecke und Z eine Basisindexmenge zu $\bar{x}$. Sei $NZ = \{1, 2, ..., m\} \setminus Z$. Durch umnummerieren der Indizes können wir A, $x, \bar{x} \in A$ und c wie folgt zerlegen:

$$A = [A_Z \mid A_{NZ}],\ x = \begin{pmatrix} x_Z \\ x_{NZ} \end{pmatrix},\ \bar{x} = \begin{pmatrix} \bar{x}_Z \\ 0 \end{pmatrix},\ c = \begin{pmatrix} c_Z \\ c_{NZ} \end{pmatrix}.$$

Da $A_Z \in \mathbb{R}^{m \times m}$ invertierbar ist, gilt weiter.

$$Ax = [A_Z \mid A_{NZ}] \begin{pmatrix} x_Z \\ x_{NZ} \end{pmatrix} = A_Z x_Z + A_{NZ} x_{NZ} = b$$

und daher

$$\begin{aligned} x_Z &= A_Z^{-1}(b - A_{NZ} x_{NZ}), \\ \bar{x}_Z &= A_Z^{-1} b \\ c^T x &= c_Z^T x_Z + c_{NZ}^T x_{NZ} \\ &= c_Z^T A_Z^{-1} b + (c_{NZ}^T - c_Z^T A_Z^{-1} A_{NZ}) x_{NZ} \\ &= c_Z^T \bar{x}_Z + (c_{NZ}^T - c_Z^T A_Z^{-1} A_{NZ}) x_{NZ} \end{aligned}$$

Wir setzen

$$Q = (q_{i,j})_{i \in Z, j \in NZ} = A_Z^{-1} A_{NZ}$$

und

$$p = (c_{NZ}^T - c_Z^T A_Z^{-1} A_{NZ})^T.$$

Bemerkung 6.9 *Ist $p \geq 0$, so ist $\bar{x}$ eine optimale Lösung von (P).*

(Es gilt sogar folgende Aussage: Eine Ecke $\bar{x}$ von P ist genau dann optimale Lösung von (P), falls eine Basisindexmenge Z zu $\bar{x}$ existiert, für die $p \geq 0$ ist. Dieser stärkere Satz wird später aus dem Satz über die Indexstrategie von Bland folgen.)

Wir beschreiben nun den Ecken-Basis-Austausch.

Ist $\bar{x}$ noch keine optimale Lösung von (P), so existiert ein $j \in NZ$ mit $p_j < 0$. Zu $\delta \geq 0$ betrachten wir $x(\delta)$ mit

$$\begin{aligned} x_j(\delta) &= \delta \\ x_k(\delta) &= 0,\ (k \in NZ \setminus \{j\}) \\ x_i(\delta) &= \bar{x}_i - \delta q_{i,j},\ (i \in Z). \end{aligned}$$

Es folgt

$$\begin{aligned}
Ax(\delta) &= A_Z x_Z + A_{NZ} x_{NZ} \\
&= A_Z(\bar{x}_Z - \delta q^j) + A_{NZ}\delta e_j \\
&= A_Z \bar{x}_Z - \delta A_Z A_Z^{-1} A_{NZ} e_j + A_{NZ}\delta e_j \\
&= b.
\end{aligned}$$

- Ist $q^j \leq 0$, so folgt $x(\delta) \in P$ für alle $\delta \geq 0$ und

  $$\lim_{\delta\to\infty} c^T x(\delta) = \lim_{\delta\to\infty} c^T \bar{x} + \delta p_j = -\infty,$$

  d.h. (P) ist unbeschränkt und besitzt also keine optimale Lösung.

- Nun sei $q_{i,j} > 0$ für mindestens ein $i \in Z$. Sei

  $$\delta_{\max} = \min_{i\in Z: q_{i,j}>0} \frac{\bar{x}_i}{q_{i,j}}$$

  und $\hat{x} = x(\delta_{\max})$. Sei $l \in Z$ mit $q_{l,j} > 0$ und $\delta_{\max} = \frac{\bar{x}_l}{q_{l,j}}$ und sei

  $$\hat{Z} = (Z \setminus \{l\}) \cup \{j\}.$$

  Es folgt $\hat{x}_l = \bar{x}_l - \frac{\bar{x}_l}{q_{l,j}} q_{l,j} = 0$ und

  $$c^T \hat{x} = c^T \bar{x} + \delta_{\max} p_j \geq c^T \bar{x}.$$

  Behauptung Die Vektoren $\{a^i \mid i \in \hat{Z}\}$ sind linear unabhängig.

  Beweis: Wir nehmen an, daß

  $$a^j + \sum_{k\in Z\setminus\{l\}} \beta_k a^k = 0$$

  gilt. Wegen $q^j = Qe_j = A_Z^{-1} A_{NZ} e_j = A_Z^{-1} a^j$, gilt

  $$a^j = A_Z q^j = \sum_{k\in Z} q_{k,j} a^k.$$

  Wir erhalten

  $$\sum_{k\in Z} q_{k,j} a^k + \sum_{k\in Z\setminus\{l\}} \beta_k a^k = q_{l,j} a^l + \sum_{k\in Z\setminus\{l\}} (q_{k,j} + \beta_k) a^k.$$

  Da Z eine Basisindexmenge ist, folgt der Widerspruch $q_{l,j} = 0$. $\square$

Wir erhalten damit, daß $\hat{Z}$ eine Basisindexmenge zur Ecke $\hat{x}$ von P ist. Ist $\delta_{\max} > 0$, so gilt $\hat{x} \neq \bar{x}$ und der Zielfunktionswert ist gestiegen. Ist $\delta_{\max} = 0$, so gilt $\hat{x} = \bar{x}$ und wir haben nur eine neue Basisindexmenge zu derselben Ecke von P gefunden. Um zu verhindern, daß der Simplexalgorithmus in einen Zykel gerät, kann man die Indizes j und k z.B. immer kleinstmöglich wählen (Indexstrategie von Bland, vgl. Beweis von Satz 5.4).

6.2 Simplexverfahrens in Tableauschreibweise

Wir betrachten das Starttableaux

$$\begin{array}{c|c|c} c_Z^T & c_{NZ}^T & 0 \\ \hline A_Z & A_{NZ} & b. \end{array}$$

Da A_Z invertierbar ist, existiert λ mit $c_Z^T = \lambda^T A_Z$. Es folgt

$$c_{NZ}^T - \lambda^T A_{NZ} = c_{NZ}^T - c_Z^T A_Z^{-1} A_{NZ} = p$$

und

$$c_Z^T \bar{x}_Z = \lambda^T A_Z A_Z^{-1} b = \lambda^T b.$$

Daher erhalten wir durch elementare Zeilenumformungen das folgende Tableau

$$\begin{array}{c|c|c} 0^T & p^T & -c_Z^T \bar{x}_Z \\ \hline I & Q & \bar{x}_Z = A_Z^{-1} b. \end{array}$$

Beispiel 6.10 *(Zwei Rechenbeispiele.)*

Verfahren zur Bestimmung einer Startecke

Überführt man $Ax \leq b$, $x \geq 0$ in die kanonische Form, $[A \mid I]\binom{x}{y} = b$, $\binom{x}{y} \geq 0$, so ist $\binom{\bar{x}}{\bar{y}} = \binom{0}{b}$ eine Ecke.

Im Allgemeinen betrachten wir zu $Ax = b$, $x \geq 0$, $\operatorname{rang}(A) = m$ und O.B.d.A. (!) $b \geq 0$ folgendes Hilfsproblem

$$(HP) \quad \begin{array}{lllll} \min & \sum_{i=1}^m y_i & & \\ s.th. & Ax + y & = & b \\ & x & \geq & 0 \\ & y & \geq & 0. \end{array}$$

Für HP ist $\binom{0}{b}$ eine Startecke. Sei nun $\binom{\bar{x}}{\bar{y}}$ eine optimale Lösung von (HP) und eine Ecke von

$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix} \mid [A|I] \begin{pmatrix} x \\ y \end{pmatrix} = b, \begin{pmatrix} x \\ y \end{pmatrix} \geq 0 \right\}.$$

Ist $\bar{y} \neq 0$, so gilt $P = \emptyset$. Ist $\bar{y} = 0$, so ist $\bar{x}$ eine Ecke von P (vgl. Folgerung 6.3).

6.3 Indexstrategie von Bland

Indexstrategie von Bland: Wähle beim Simplexalgorithmus, den Index $j \in NZ$, der zur Basisindexmenge hinzugefügt werden soll minimal, und wähle auch den Index $l \in Z$, der die Basisindexmenge verlassen soll minimal.

Satz 6.11 *Bei Anwendung der Indexstrategie von Bland wird beim Simplexalgorithmus keine Basis zweimal betrachtet.*

Beweis: Wir nehmen an, daß der Simplexalgorithmus bei Anwendung der Indexstrategie von Bland eine Basis zweimal betrachtet. Dann läuft er in einen Zykel $Z_1, Z_2, ..., Z_l, Z_1, Z_2, ...$ von Basen zu einer festen Ecke $\bar{x}$.

Sei t der größte Index, der während des Zykels neu in die Basisindexmenge aufgenommen wird. Sei

$$\begin{array}{c|c} p'^T & v' \\ \hline A' & b' \end{array}$$

das vollständige Tableau zu Beginn der Iteration des Zykels, in der t in die aktuelle Basisindexmenge Z' aufgenommen wird. Hierbei seien die Spalten nicht so umgeordnet, daß die Spalten, die Basisindizes entsprechen vorne stehen.

Es gilt für entsprechende definierte Vektoren (s.o.) und alle $x \geq 0$ mit $Ax = b$, daß

$$c^T x = c_{Z'}^T \bar{x}_{Z'} + (c_{NZ'}^T - c_{Z'}^T A_{Z'}^{-1} A_{NZ'}) x_{NZ'}.$$

Wählt man $\lambda'^T = c_{Z'}^T A_{Z'}^{-1}$ so folgt

$$c^T x = \lambda'^T b + (c_{NZ'}^T - \lambda'^T A_{NZ'}) x_{NZ'}.$$

Daher gilt für das obige Tableau

$$\begin{aligned} A' &= A_{Z'}^{-1} A \\ b' &= A_{Z'}^{-1} b \\ p'^T &= c^T - \lambda'^T A \\ v &= -\lambda'^T b = -c^T \bar{x}. \end{aligned}$$

(Beachte hierbei, daß die Einträge von p'^T, die Basisindizes entsprechen, durch diese Setzung gleich 0 sind.)

Sei weiter

$$\begin{array}{c|c} p''^T & v'' \\ \hline A'' & b'' \end{array}$$

das vollständige Tableau zu Beginn der Iteration des Zykels, in der t die aktuelle Basisindexmenge Z'' wieder verläßt und dafür der Index k neu aufgenommen wird. Wegen der Wahl von t gilt $k < t$.

Sei $a''^k = (q_1, q_2, ..., q_n)^T$ die k-te Spalte von A''. Setze man $y \in \mathbb{R}^n$ mit $y_k = 1$; $y_j = -q_i$, falls $j \in Z''$ und $a''^j = e_i$ und $y_j = 0$, sonst, so folgt

$$Ay = A_{Z''}(A_{Z''}^{-1} A y) = A_{Z''} A'' y = 0$$

und daher $A(\bar{x}+y)=b$ und

$$c^T(\bar{x}+y)=(p'^T+\lambda'^T A)(\bar{x}+y)=p'^T\bar{x}+p'^T y+\lambda'^T b=p'^T y-v.$$

Es gilt $p'_k y_k = p'_k \geq 0$ aufgrund der Wahl von t.

Ist $j \in Z''$, $j<t$ und $\bar{x}_j=0$, so folgt $y_j \geq 0$, sonst würde nach der Blandschen Regel ein kleinerer Index als t aus Z'' entfernt (es wäre $q_i>0$ wobei $a''^j=e_i$ und bei der Definition von $\delta_{\max}$ ginge der Quotient $\frac{\bar{x}_j}{q_i}=0$ ein). Weiter folgt wegen $j<t$, daß $p'_j \geq 0$.

Ist $j \in Z''$ und $j>t$ oder $j \in Z''$ und $j<t$ und $\bar{x}_j>0$, so ist $j \in Z'$ und daher $p'_j=0$.

Es ergibt sich einerseits

$$p'^T y=\sum_{j\in Z''} p'_j y_j+p'_k y_k=\sum_{j\in Z'',j<t} p'_j y_j+p'_t y_t+\sum_{j\in Z'',j>t} p'_j y_j+p'_k y_k \geq p'_t y_t>0.$$

Andererseits gilt für passende Vektoren

$$\begin{aligned} c^T(\bar{x}+y) &= (p''^T+\lambda''^T A)(\bar{x}+y) \\ &= \lambda''^T A\bar{x}+\lambda''^T Ay+p''^T\bar{x}+p''^T y \\ &= \lambda''^T b+p''^T y=p''^T y-v \end{aligned}$$

und somit

$$p''^T y=\sum_{j\in Z''} p''_j y_j+p''_k y_k=p''_k y_k<0.$$

Zusammen ergibt sich der Widerspruch $-v<c^T(\bar{x}+y)<-v$ und der Beweis ist vollständig.
$\square$

6.4 Die Laufzeit des Simplexverfahrens

(**Klee-Minty Beispiel**) Für $0 < \epsilon < \frac{1}{2}$ betrachten wir folgendes lineare Programm.

$$(P) \qquad \begin{array}{lllll} \min & -x_d & & & \\ s.th. & x_1 - x_{d+1} & = & \epsilon & \\ & x_1 + x_{2d+1} & = & 1 & \\ & x_j - \epsilon x_{j-1} - x_{d+j} & = & 0 & j = 2,3,...,d \\ & x_j + \epsilon x_{j-1} + x_{2d+j} & = & 1 & j = 2,3,...,d \\ & x_j & \geq & 0 & j = 1,2,...,3d \end{array}$$

($A \in \mathbb{R}^{2d\times 3d}$ mit rang$(A) = 2d$ - die Variabeln x_j für $d+1 \leq j \leq 3d$ kann man frei wählen.)

Lemma 6.12 *Zu jeder Menge $Z \subseteq \{1,2,...,3d\}$ mit $\{1,2,...,d\} \subseteq Z$ und $|\{d+j, 2d+j\} \cap Z| = 1$ für $1 \leq j \leq d$ existiert eine nicht entartete Ecke $\bar{x}_Z$ von $P = \{x \mid Ax = b, x \geq 0\}$, deren Basisindexmenge Z ist.*

Beweis: Setze $\bar{x}_{Z,j} = 0$ für alle $j \notin Z$. Ist $d+1 \notin Z$, so folgt $\bar{x}_{Z,1} = \epsilon$. Ist $2d+1 \notin Z$, so folgt $\bar{x}_{Z,1} = 1$.

Es gelte nun für ein $1 \leq j \leq d$, daß $0 < x_l \leq 1$ für alle $1 \leq l \leq j$. Ist $d+j \notin Z$, so folgt

$$0 < \bar{x}_{Z,j} = \epsilon \bar{x}_{Z,j-1} \leq 1$$

und

$$\bar{x}_{Z,2d+j} = 1 - (\bar{x}_{Z,j} + \epsilon \bar{x}_{Z,j-1}) = 1 - 2\epsilon \bar{x}_{Z,j-1} \geq 1 - 2\epsilon > 0.$$

Ist $2d+j \notin Z$, so folgt

$$0 < \bar{x}_{Z,j} = 1 - \epsilon \bar{x}_{Z,j-1} \leq 1$$

und

$$\bar{x}_{Z,d+j} = \bar{x}_{Z,j} - \epsilon \bar{x}_{Z,j-1} = 1 - 2\epsilon \bar{x}_{Z,j-1} \geq 1 - 2\epsilon > 0.$$

Es ergibt sich also $\bar{x}_j > 0$ genau dann wenn, $j \notin Z$. Da alle Koordinaten von $\bar{x}$ eindeutig bestimmt sind, sind die Spalten $\{a^j \mid j \in Z\}$ linear unabhängig und somit ist $\bar{x}$ eine Ecke und Z ihre Basisindexmenge. $\square$

Lemma 6.13 *Sind Z und Z' zwei Mengen wie in Lemma 6.12 und $2d \in Z \setminus Z'$, so folgt $\bar{x}_{Z,d} > \bar{x}_{Z',d}$.*

Gilt $Z' = (Z \setminus \{2d\}) \cup \{3d\}$, so folgt $\bar{x}_{Z',d} = 1 - \bar{x}_{Z,d}$.

Beweis: $2d \in Z \Rightarrow \bar{x}_{Z,3d} = 0 \Rightarrow \bar{x}_{Z,d} = 1 - \epsilon \bar{x}_{Z,d-1} > \frac{1}{2}$.

$3d \in Z' \Rightarrow \bar{x}_{Z',2d} = 0 \Rightarrow \bar{x}_{Z,d} = \epsilon \bar{x}_{Z,d-1} < \frac{1}{2}$.

$Z' = (Z \setminus \{2d\}) \cup \{3d\} \Rightarrow \bar{x}_{Z',d-1} = \bar{x}_{Z,d-1}$. $\square$

Lemma 6.14 *Die Teilmengen von $\{1,2,...,3d\}$ können so angeordnet werden*

$$Z_1, Z_2, ..., Z_{2^d},$$

daß für $1 \leq i \leq 2^d$ die Menge Z_{i+1} durch einen Ecken-Basis-Austausch aus Z_i entsteht und $\bar{x}_{Z_i,d} < \bar{x}_{Z_{i+1},d}$ gilt.

Beweis: (Induktion nach d.)

$d = 1$: $(\epsilon, 0, 1-\epsilon), (1, 1-\epsilon, 0)$.

$d \to (d+1)$: Seien $Z_1, ..., Z_{2^d}$ entsprechend geordnet. Sei

$$\begin{aligned} Z_i' &= \{1, 2, ..., d+1\} \\ &\cup \{j+1 \mid d+1 \leq j \leq 2d, j \in Z_i\} \\ &\cup \{j+2 \mid 2d+1 \leq j \leq 3d, j \in Z_i\} \\ &\cup \{3(d+1)\} \end{aligned}$$

und

$$\begin{aligned} Z_i'' &= \{1, 2, ..., d+1\} \\ &\cup \{j+1 \mid d+1 \leq j \leq 2d, j \in Z_i\} \\ &\cup \{j+2 \mid 2d+1 \leq j \leq 3d, j \in Z_i\} \\ &\cup \{2(d+1)\}. \end{aligned}$$

Es gilt

$$\bar{x}_{Z_i', d+1} = \epsilon \bar{x}_{Z_i', d} = \epsilon \bar{x}_{Z_i, d} < \epsilon$$

und

$$\bar{x}_{Z_i'', d+1} = 1 - \bar{x}_{Z_i', d+1} > \epsilon.$$

Somie folgt

$$\bar{x}_{Z_1', d+1} < ... < \bar{x}_{Z_{2^d}', d+1} < \bar{x}_{Z_{2^d}'', d+1} < ... < \bar{x}_{Z_1'', d+1}$$

und der Beweis ist vollständig. □

Satz 6.15 *Für jedes $d \in \mathbb{N}$ existiert ein kanonisches lineares Programm mit $2d$ Gleichungen, $3d$ Unbekannten, ganzzahligen Koeffizienten und einem optimalen Wert < 4, für das das Simplexverfahren ggf. $2^d - 1$ Austauschschritte benötigt, um eine optimale Lösung zu finden.*

Beweis: Wähle $\epsilon = 1/4$ und multipliziere alle Gleichungen mit 4. □

6.5 Probabilistische Analyse des Simplexverfahrens

Wir haben gesehen, daß die Laufzeit des Simplexalgorithmus exponentiell in der Dimension des gegebenen Problems sein kann. Warum aber funktioniert der Simplexalgorithmus in der Praxis so gut? Eine Antwort auf diese Frage liefert eine "Probabilistische Analyse" seiner Laufzeit.

Wir betrachten LP's der Form

$$\max\{c^T x \mid Ax \leq b\}$$

mit $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ und $c \in \mathbb{R}^n$ und wenden eine spezielle Form des Simplexalgorithmus, den sog. "*Schatteneckenalgorithmus*", an.

Dazu benötigen wir neben dem LP noch eine Startecke x_0 und einen Vektor $\bar{c} \in \mathbb{R}^n$ mit

$$\bar{c}^T x_0 = \max\{\bar{c}^T x \mid Ax \leq b\}.$$

Weiter nehmen wir an, daß A, b, c und $\bar{c}$ zufällig gewählt werden und daß das entsprechende Wahrscheinlichkeitsmaß folgende Eigenschaften besitzt.

P_1 Dreht man für ein festes $\bar{c}$ einige der Ungleichungen in $Ax \leq b$ um oder ersetzt c durch $-c$ so, daß der Wert des LP's endlich bleibt, ändert sich nicht die Wahrscheinlichkeit der Wahl von A, b, c und $\bar{c}$.

P_2 Die Wahrscheinlichkeit, A, b, c und $\bar{c}$ so zu wählen, daß entweder $\leq n$ linear abhängige Zeilen in $\begin{bmatrix} c^T \\ \bar{c}^T \\ A \end{bmatrix}$ existieren oder $\leq (n+1)$ linear abhängige Zeilen in $[A \mid b]$ existieren ist gleich 0.

Um den Schatteneckenalgorithmus zu beschreiben, müssen wir nur seine Pivotregel festlegen.

Pivotregel des Schatteneckenalgorithmus
Ist für ein $\lambda \geq 0$ eine Ecke x_k gegeben, die

$$(\lambda c + \bar{c})^T x_k = \max\{(\lambda c + \bar{c})^T x \mid Ax \leq b\}$$

und

$$c^T x_k < \max\{c^T x \mid Ax \leq b\}$$

erfüllt, so wähle entweder x_{k+1} als zu x_k adjazente Ecke (x_{k+1} entsteht aus x_k durch einen Simplexschritt) mit

$$(\lambda' c + \bar{c})^T x_{k+1} = \max\{(\lambda' c + \bar{c})^T x \mid Ax \leq b\}$$

für ein $\lambda' > \lambda$ oder finde ein $y \in \mathbb{R}^n$ mit $Ay \leq 0$ und

$$(\lambda' c + \bar{c})^T y > 0$$

für ein $\lambda' > \lambda$.

Der Algorithmus startet mit der Ecke $x_k = x_0$ und $\lambda = 0$. Es ist möglich, obige Pivotregel als Schritt des Simplexalgorithmus zu interpretieren.

Bemerkung 6.16 *Mit Wahrscheinlichkeit* 1 *gilt*

$$\begin{array}{rcl} c^T x_{k+1} & > & c^T x_k \\ \bar{c}^T x_{k+1} & < & \bar{c}^T x_k \end{array}$$

bzw.

$$\begin{array}{rcl} c^T y & > & 0 \\ \bar{c}^T y & < & 0. \end{array}$$

Bemerkung 6.17 *Geometrisch entspricht der Schatteneckenalgorithmus der Projektion von* $P = \{x \mid Ax \le b\}$ *auf den* 2*-dimensionalen Raum, der von* c *und* $\bar{c}$ *aufgespannt wird.*

Der Algorithmus betrachtet nur die Ecken von P*, die Ecken des projizierten Polyeders* Q *entsprechen. Solche Ecken nennt man Schattenecken.*

Satz 6.18 (Haimovich 1983) *Zur Lösung von entsprechend* P_1 *und* P_2 *gewählten linearen Programmen benötigt der Schatteneckenalgorithmus im Durchschnitt* $\frac{n}{2}$ *viele (Pivot)Schritte.*

7 Polynomielle Algorithmen

Für $f, g : D \to \mathbb{R}_{\geq 0}$ schreibt man $f = O(g)$ falls $f(x) \leq \alpha g(x) + \beta$ für geeignete $\alpha, \beta > 0$ und alle $x \in D$ gilt. Ist $f = O(g)$ und $g = O(f)$, so schreibt man $f = \Theta(g)$.

Um die Laufzeit eines **Algorithmus** abzuschätzen, zählen wir die "**elementaren Operationen**", die der Algorithmus ausführt. Beispiele für elementare Operationen sind einfache arithmetische Operationen, Vergleiche, Setzungen, Zugriff,....

Die verschiedenen Probleme, die ein und derselbe Algorithmus bearbeiten kann, nennt man **Instanzen**. Die **Inputgröße** (**Kodierungslänge,...**) einer Instanz, die eine Liste rationaler Zahlen ist, ist die Anzahl von **Bits**, die benötigt werden, um alle Zahlen der Liste in binärer Darstellung zu kodieren.

Man sagt, daß ein Algorithmus, dessen Instanzen Listen rationaler Zahlen sind, eine **polynomielle Laufzeit** besitzt, wenn ein $k \in \mathbb{N}$ existiert, so daß er für eine Instanz mit Inputgröße n insgesamt $O(n^k)$ elementare Operationen ausführt und alle Zahlen, die er als Zwischenergebnisse erzeugt mit $O(n^k)$ vielen Bits bespeichert werden können.

Man sagt, daß ein Algorithmus, dessen Instanzen Listen rationaler Zahlen sind, eine **streng polynomielle Laufzeit** besitzt, wenn er polynomielle Laufzeit besitzt und ein $k \in \mathbb{N}$ existiert, so daß er für eine Instanz, die eine Liste von n rationalen Zahlen ist insgesamt $O(n^k)$ elementare Operationen ausführt.

Definition 7.1 *Wir setzen* $\text{size}(n) = 1+\lceil \log_2(|n|+1)\rceil$ *für* $n \in \mathbb{Z}$, $\text{size}(r) = \text{size}(p)+\text{size}(q)$ *für* $r = \frac{p}{q} \in \mathbb{Q}$ *mit* $\text{ggt}(p,q) = 1$, $\text{size}(x) = n + \text{size}(x_1) + ... + \text{size}(x_n)$ *für* $x = (x_1, ..., x_n) \in \mathbb{Q}^n$ *und* $\text{size}(A) = mn + \sum_{i,j} \text{size}(a_{i,j})$ *für* $A = (a_{i,j}) \in \mathbb{Q}^{m\times n}$.

Lemma 7.2 *Sind* $r_1, r_2, ..., r_n \in \mathbb{Q}$, *so gilt*

$$\text{size}(r_1 r_2 ... r_n) \leq \text{size}(r_1) + ... + \text{size}(r_n)$$

und

$$\text{size}(r_1 + r_2 + ... + r_n) \leq 2(\text{size}(r_1) + ... + \text{size}(r_n)).$$

Lemma 7.3 *Sind* $x, y \in \mathbb{Q}^n$, *so gilt*

$$\text{size}(x + y) \leq 2(\text{size}(x) + \text{size}(y))$$

und

$$\text{size}(x^T y) \leq 2(\text{size}(x) + \text{size}(y)).$$

Lemma 7.4 *Ist* $A \in \mathbb{Q}^{n\times n}$, *so gilt* $\text{size}(\det(A)) \leq 2\text{size}(A)$.

Satz 7.5 *Besitzt das lineare Programm*

$$(P) \min\{c^T x \mid Ax \leq b\}$$

mit $A \in \mathbb{Q}^{m\times n}$ $b \in \mathbb{Q}^m$ $c \in \mathbb{Q}^n$ *— solch ein lineare Programm nennt man auch rationales lineares Programm — eine optimale Lösung, so existiert eine optimale Lösung* $x = (x_1, ..., x_n) \in \mathbb{Q}^n$ *mit*

$$\text{size}(x_i) \leq 4(\text{size}(A) + \text{size}(b))$$

und somit

$$\text{size}(x) \leq 4n(\text{size}(A) + \text{size}(b)).$$

Ist b *entweder* e_j *oder* $-e_j$ *für einen Einheitsvektor* e_j*, so existiert eine reguläre Teilmatrix* A' *von* A *und eine optimale Lösung* $x = (x_1, ..., x_n) \in \mathbb{Q}^n$ *mit*

$$\text{size}(x_i) \leq 4\text{size}(A')$$

und somit

$$\text{size}(x) \leq 4n\text{size}(A').$$

Bemerkung 7.6 *Gegeben sei ein System* $Ax \leq b$ *rationaler Ungleichungen.*

Ist $Ax \leq b$ *lößbar, so existiert eine Lösung, deren Kodierungslänge polynomiell beschränkt in der Kodierungslänge des Systems ist.*

Ist $Ax \leq b$ *nicht lößbar, so existiert nach dem Lemma von Farkas ein Vektor* $y \geq 0$ *mit* $y^T A = 0$ *aber* $y^T b = -1$*. Nach Satz 7.5 kann* y *wieder so gewählt werden, dass seine Kodierungslänge polynomiell beschränkt in der Kodierungslänge des Systems ist.*

Es existiert also sowohl für die Existenz als auch für die Nichtexistenz einer Lösung von $Ax \leq b$ *ein "kurzer" Beweis. In einer solchen Situation sagt man, daß das Entscheidungsproblem "Hat* $Ax \leq b$ *eine Lösung?" eine gute Charakterisierung besitzt.*

Folgerung 7.7 *Ist* $P \subseteq \mathbb{R}^n$ *ein rationales Polytop (=die konvexe Hülle endliche vieler rationaler Vektoren) und gilt* $\text{size}(x) \leq T$ *für alle Ecken* x *von* P *und ein* $T > 0$*, so gilt* $P = \{x \mid Ax \leq b\}$ *für ein System* $Ax \leq b$ *von Ungleichungen, wobei für*

$$\text{size}(a) + \text{size}(\beta) \leq 75n^2T$$

jede Ungleichung $a^T x \leq \beta$ *von* $Ax \leq b$ *gilt.*

7.1 Beispiele polynomieller Algorithmen

Der Euklidische Algorithmus

Instanz: Zwei natürliche Zahlen $p, q \in \mathbb{N}$

Ausgabe: Der größte gemeinsame Teiler $\text{ggt}(p, q)$ von p und q.

1. Solange $p, q > 0$ gilt, setze $q = q - \lfloor \frac{q}{p} \rfloor p$ falls $p < q$ gilt, sonst setze $p = p - \lfloor \frac{p}{q} \rfloor q$.

2. Gebe $\max\{p, q\}$ aus.

Satz 7.8 *Der Euklidische Algorithmus arbeitet korrekt und benötigt höchstens* $\text{size}(p) + \text{size}(q)$ *viele Wiederholungen von Schritt 1 (= Iterationen).*

Kettenbruchentwicklung

Instanz: Eine rationale Zahl $x = \frac{p}{q} \in \mathbb{Q}$

Ausgabe: Eine Folge $(x_i = \frac{p_i}{q_i})_{i \geq 0}$ mit $x_0 = x$ und $x_{i+1} = \frac{1}{x_i - \lfloor x_i \rfloor}$ für $i \geq 0$.

1. Setze $i = 0$, $p_0 = p$ und $q_0 = q$.
   Setze $g_{-2} = 0$, $g_{-1} = 1$, $h_{-2} = 1$ und $h_{-1} = 0$.

2. Solange $q_i \neq 0$ setze
   $a_i = \lfloor \frac{p_i}{q_i} \rfloor$.
   $g_i = a_i g_{i-1} + g_{i-2}$.
   $h_i = a_i h_{i-1} + h_{i-2}$.
   $q_{i+1} = p_i - a_i q_i$.
   $p_{i+1} = q_i$.
   $i = i + 1$.

Lemma 7.9 *Für alle $i \geq 0$ gilt für die Kettenbruchentwicklung*

(i) $a_i \geq 1$ für $i \geq 1$.

(ii) $h_i \geq h_{i-1}$.

(iii) $g_{i-1}h_i - g_i h_{i-1} = (-1)^i$

(iv) $x = \frac{p_i g_{i-1} + q_i g_{i-2}}{p_i h_{i-1} + q_i h_{i-2}}$

(v) $x \geq \frac{g_i}{h_i}$ für gerade i und $x \leq \frac{g_i}{h_i}$ für ungerade i.

Satz 7.10 (Khintchine 1956) *Es existiert ein Algorithmus mit polynomieller Laufzeit, der zu $\alpha \in \mathbb{Q}$ und $n \in \mathbb{N}$ eine Zahl $\beta \in \mathbb{Q}$ bestimmt, deren Nenner höchstens n ist und die unter dieser Bedingung $|\alpha - \beta|$ minimiert.*

(Die Laufzeit ist polynomiell in $\mathrm{size}(\alpha) + \mathrm{size}(n)$.)

Satz 7.11 (Edmonds 1967) *Die* **Gaußsche Elimination***, die zu einer gegeben rationalen Matrix $A = (a_{i,j}) \in \mathbb{Q}^{m \times n}$ den Rang r von A, eine maximale reguläre Teilmatrix $A' = (a_{row(i),col(j)})_{1 \leq i,j \leq r}$ von A und die Inverse von A' erzeugt, benötigt $O(mnr)$ elementare Operationen und alle Zahlen, die als Zwischenergebnisse entstehen, können mit $O(m(m+n)\mathrm{size}(A))$ vielen Bits gespeichert werden.*

Wir betrachten die folgenden Probleme.

(P1) Entscheide zu $A \in \mathbb{Q}^{m \times n}$ und $b \in \mathbb{Q}^m$, ob $Ax \leq b$ eine Lösung besitzt.

(P2) Entscheide zu $A \in \mathbb{Q}^{m \times n}$ und $b \in \mathbb{Q}^m$, ob $Ax \leq b$ eine Lösung besitzt und wenn ja bestimme eine Lösung.

(P3) Entscheide zu $A \in \mathbb{Q}^{m\times n}$ und $b \in \mathbb{Q}^m$, ob $Ax = b$ eine nicht-begative Lösung besitzt und wenn ja bestimme eine Lösung.

(P4) Gegeben $A \in \mathbb{Q}^{m\times n}$, $b \in \mathbb{Q}^m$ und $c \in \mathbb{Q}^n$, entscheide ob $(P) \max\{c^T x \mid Ab \leq b\}$ entweder keine zulässige Lösung besitzt oder unbeschränkt ist oder eine optimale Lösung besitzt.

Falls (P) eine optimale Lösung besitzt, finde eine.

Falls (P) unbeschränkt ist, finde einen Vektor $z \in \mathbb{Q}^n$ mit $Az \leq 0$ und $c^T z > 0$.

Satz 7.12 *Ist eines der Probleme (P1), (P2), (P3) oder (P4) in polynomieller Zeit lößbar, so sind alle anderen Probleme auch in polynomieller Zeit lößbar.*

Satz 7.13 *Seien $A \in \mathbb{R}^{m\times n}$ mit $\operatorname{rang}(A) = m$, $b', b'' \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ und $\Delta > 0$ so, daß $(P') \max\{c^T x \mid Ax \leq b'\}$ und $(P'') \max\{c^T x \mid Ax \leq b''\}$ endlich sind und die Beträge aller Einträge der Inversen jeder regulären Teilmatrix von A durch Δ beschränkt sind.*

Es existiert zu jeder optimalen Lösung x' von (P') eine optimale Lösung x'' von (P'') mit

$$||x' - x''||_\infty \leq n\Delta ||b' - b''||_\infty.$$

Bemerkung 7.14 *Der Faktor $n\Delta$ in Satz 7.13 ist best-möglich.*

8 Die Ellipsoidmethode

Die Ellipsoidmethode geht ursprünglich auf Yudin und Nemirovskii (1976) und Shor (1977) zurück und wurde zur Lösung konvexer 'feasibility'-Probleme entwickelt, d.h. Fragen der Form "Ist eine geg. konvexe Menge leer?" 1979 beobachtete Khachiyan, daß man mit ihr lineare Programme in polynomieller Zeit lösen kann.

Definition 8.1 *Eine Menge der Form $E(A, x) = \{z \in \mathbb{R}^n \mid (z - x)^T A^{-1}(z - x) \leq 1\}$ für eine symmetrische positiv definite Matrix $A \in \mathbb{R}^{n\times n}$ und ein $x \in \mathbb{R}^n$ nennt man Ellipsoid.*

Bemerkung 8.2 *(***Idee der Ellipsoidmethode***) Man möchte entscheiden, ob ein geg. Ungleichungssystem $A'z \leq b'$ eine Lösung besitzt und betrachtet dazu eine Folge von Ellipsoiden $E(A, x)$, die jeweils alle zulässigen Lösungen enthalten.*

Man wählt zunächst eine Kugel, die groß genug ist, um alle relevanten Punkte zu enthalten. Iterativ testet man dann jeweils den Mittelpunkt x eines Ellipsoiden $E(A, x)$.

Ist x Lösung von $A'z \leq b'$, so ist $A'z \leq b'$ lösbar und man hat sogar eine Lösung gefunden.

Ist x keine Lösung von $A'z \leq b'$, so geht man zu einem kleineren Ellipsoiden über.

Das Volumen eines Ellipsoiden $E(A, x)$ ist gegeben durch ([3], 163)

$$\operatorname{volume}(E(A, x)) = \sqrt{\det(A)} \operatorname{volume}(B(0, 1))$$

Zu einem gegebenen Ellipsoiden $E(A,x)$ und einer Hyperebene $\{z \mid a^T z = a^T x\}$ durch das Zentrum von $E(A,x)$ gibt es einen eindeutigen kleinsten Ellipsoiden $E(A',x')$, der

$$\{z \in E(A,x) \mid a^T z \geq a^T x\}$$

enthält. Hierbei gilt

$$A' = \frac{n^2}{n^2-1}\left(A - \frac{2}{n+1} bb^T\right),$$
$$x' = x + \frac{1}{n+1} b$$

und

$$b = \frac{1}{\sqrt{a^T A a}} Aa$$

Algorithmus 8.3 Die Ellipsoid Methode
Input: Ein $n \in \mathbb{N}$, $n \geq 2$. Eine Zahl $N \in \mathbb{N}$. $x_0 \in \mathbb{Q}^n$ und $R \in \mathbb{Q}_{\geq 0}$, $R \geq 2$.
Ausgabe: Eine Folge von Ellipsoiden $(E_i = E(A_i, x_i))_{i=1,2,\ldots,N}$.
1) Setze $p := \lceil 6N + \log(9n^3) \rceil$, $A_0 := R^2 I$ und $k := 0$.
2) Wähle ein $a_k \in \mathbb{Q}^n \setminus \{0\}$ (später wird a_k der Vektor sein, der eine im Mittelpunkt des aktuellen Ellipsoiden verletzte Ungleichung beschreibt.)
3) Setze

$$b_k := \frac{1}{\sqrt{a_k^T A_k a_k}} A_k a_k$$
$$x_{k+1} :\sim x_{k+1} := x_k + \frac{1}{n+1} b_k$$
$$A_{k+1} :\sim A_{k+1}^* := \frac{2n^2+3}{2n^2}\left(A_k - \frac{2}{n+1} b_k b_k^T\right)$$

wobei ":$\sim$" Rundung auf die p-te Nachkommastelle in binärer Darstellung der Zahlen entspricht. Es werde so gerundet, daß A_{k+1} wieder symmetrisch ist.
4) Setze $k := k+1$. Ist $k < N$, so gehe zu 2).

Wie bisher sei $||x||$ die Euklidische Norm von $x \in \mathbb{R}^n$. Weiter sei

$$||A|| = \max\{||Ax|| \mid ||x|| = 1\}.$$

Für $A^T = A$ ist $||A||$ der maximale Betrag eines Eigenwertes von A und

$$||A|| = \max\{x^T A x \mid ||x|| = 1\}.$$

Lemma 8.4 (Grötschel, Lovász, Schrijver 1981) *Für $0 \leq k \leq N$ ist die Matrix A_k positiv definit,*

$$||x_k|| \leq ||x_0|| + R2^k,$$
$$||A_k|| \leq R^2 2^k$$

und

$$||A_k^{-1}|| \leq R^{-2} 4^k.$$

Bemerkung 8.5 *Für die späteren Abschätzungen der Laufzeiten sind folgende Beobachtungen nützlich.*

$$p = O(N + \log n)$$

Da $|n| \leq \frac{1}{2}2^{\text{size}(n)} - 1$ *sind die Beträge aller Komponenten der Vektoren* x_k *durch* $O(2^{\text{size}(x_0)} + R2^N)$ *beschränkt und daher*

$$\text{size}(x_k) = O(n(\text{size}(x_0) + N \log R + p)).$$

Für die Einträge der Matrizen $A_k = (a_{i,j})$ *erhalten wir*

$$|a_{i,j}| = |e_i^T A_k e_j| \leq ||A_k||$$

und somit

$$\text{size}(A_k) = O(n^2(N \log R + p)).$$

Lemma 8.6 *Für* $0 \leq k \leq N-1$ *gilt*

$$\{x \in E_k \mid a_k^T x \geq a_k^T x_k\} \cap E_0 \subseteq E_{k+1}.$$

($E_0 = \{x \mid ||x - x_0|| \leq R\}$*)*

Bemerkung 8.7 *Die Einschränkung "*$\cap E_0$*" in obigem Lemma wird im Wesentlichen nur durch die Rundung :∼ nötig. Ohne Rundung könnte man sie weglassen.*

Anschaulich bedeutet diese Bedingung, daß bei immer schmaler werdenden Ellipsoiden schon kleine Rundungen starken Einfluß auf die stärkste Ausdehnungsrichtung haben können.

Lemma 8.8 *Für* $0 \leq k \leq N-1$ *gilt* $\frac{\text{volume}(E_{k+1})}{\text{volume}(E_k)} \leq e^{-\frac{1}{5n}}$.

Lemma 8.9 (Khachiyan 1979, Gács and Lovász 1981) *Sind* $A \in \mathbb{Q}^{m \times n}$ *und* $b \in \mathbb{Q}^m$, *so hat das System* $Ax \leq b$ *genau dann eine Lösung, wenn* $Ax \leq b + \epsilon\mathbf{1}$, $-R\mathbf{1} \leq x \leq R\mathbf{1}$ *eine Lösung besitzt für*

$$\frac{1}{\epsilon} = 2n2^{4(\text{size(A)}+\text{size(b)})}$$

und

$$R = 1 + 2^{4(\text{size(A)}+\text{size(b)})}.$$

Sind P *und* P' *die beiden Polyeder, die durch die zwei Ungleichungssysteme beschrieben werden, so folgt aus* $P \neq \emptyset$, *daß* $\text{volume}(P') \geq \left(\frac{2\epsilon}{n2^{\text{size}(A)}}\right)^n$.

Ist $P \neq \emptyset$, so existiert ein $x = (x_1, ..., x_n) \in P$ mit $\text{size}(x_i) \leq 4(\text{size}(A) + \text{size}(b))$ und damit (wegen $\text{size}(n) = 1 + \lceil \log(|n| + 1) \rceil \Rightarrow |n| \leq \frac{1}{2}2^{\text{size}}(n) - 1$)

$$|x_i| \leq \frac{1}{2}2^{4(\text{size}(A)+\text{size}(b))} - 1 = \frac{1}{2}R - \frac{3}{2}.$$

Für $z = (z_1, ..., z_n)$ mit $|z_i| \leq \delta$ sind die Beträge aller Komponenten von Az durch $2^{\text{size}(A)} n\delta$ beschränkt. Da nun

$$\frac{\epsilon}{2^{\text{size}(A)}n} = \frac{1}{2 2^{\text{size}(A)} 2^{4(\text{size}(A)+\text{size}(b))} n^2} < \frac{3}{2}$$

folgt

$$\left\{ z \mid ||z - x||_\infty \leq \frac{\epsilon}{2^{\text{size}(A)}n} \right\} \subseteq P'$$

und somit volume$(P') \geq \left(\frac{2\epsilon}{n 2^{\text{size}(A)}}\right)^n$. □

Satz 8.10 (Khachiyan 1979) *Es existiert ein Algorithmus mit polynomieller Laufzeit, der rationale lineare Programme lößt, d.h. Zulässigkeit und Beschränktheit entscheidet und ggf. eine optimale Lösung findet.*

9 Separierung und Optimierung

Eine wichtige Eigenschaft der Ellipsoidmethode ist es, daß man mit ihr auch LP's lösen kann, die exponentiell viele Restriktionen beinhalten. (Dies liegt daran, daß man immer nur eine verlezte Restriktion finden muss.)

Problem 9.1 Separierungsproblem zum Polytop P
Gegeben: *Ein Vektor $y \in \mathbb{Q}^n$.*
Aufgabe: *Entscheide, ob $y \in P$ liegt oder finde ein $d \in \mathbb{Q}^n$ mit $d^T y > d^T x$ für alle $x \in P$.*

Problem 9.2 Schwaches Separierungsproblem zu einem Polytop P, einem Vektor $c \in \mathbb{Q}^n$ und einem $\epsilon > 0$.
Gegeben: *Ein Vektor $y \in \mathbb{Q}^n$.*
Aufgabe: *Finde entweder ein $y' \in P$ mit $c^T y \leq c^T y' + \epsilon$ oder finde ein $d \in \mathbb{Q}^n$ mit $d^T y > d^T x$ für alle $x \in P$.*

Problem 9.3 Schwaches Optimierungsproblem
Gegeben: *Ein $n \in \mathbb{N}$. Ein Vektor $c \in \mathbb{Q}^n$ und ein $\epsilon > 0$.*
Ein Polytop $P \subseteq \mathbb{R}^n$ mittels eines Orakels für das schwache Separierungsproblem zu P, c und $\frac{\epsilon}{2}$.
Aufgabe: *Finde einen Vektor $y \in P$ mit $c^T y \geq \max\{c^T x \mid x \in P\} - \epsilon$.*

Algorithmus 9.4 Grötschel-Lovász-Schrijver Algorithmus
Gegeben: *Ein $n \in \mathbb{N}$, $n \geq 2$. Ein Vektor $c \in \mathbb{Q}^n$ und ein $0 < \epsilon \leq 1$.*
Ein Polytop $P \subseteq \mathbb{R}^n$ mittels eines Orakels für das schwache Separierungsproblem zu P, c und $\frac{\epsilon}{2}$.
Ein $x_0 \in \mathbb{Q}^n$ und $r, R \in \mathbb{Q}_{>0}$ mit $B(x_0, r) \subseteq R \subseteq B(x_0, R)$.
Aufgabe: *Ein Vektor $y^* \in P$ mit $c^T y^* \geq \max\{c^T x \mid x \in P\} - \epsilon$.*

1) Setze $R := \max\{R, 2\}$, $r := \min\{r, 1\}$, $\gamma := \max\{||c||, 1\}$, $N := 5n^2 \left\lceil \log \frac{4R^2\gamma}{r\epsilon} \right\rceil$ und $y^* = x_0$.

2) Wende Algorithmus 8.3, die **Ellipsoid Methode** *an wobei a_k in Schritt 2 wie folgt bestimmt wird.*
Frage das Orakel für das schwache Separierungsproblem zu P, c und $\frac{\epsilon}{2}$ nach $y = x_k$.

Liefert das Orakel ein $y' \in P$ mit $c^T y \leq c^T y' + \frac{\epsilon}{2}$, dann setze $a_k := c$ und, falls $c^T y' > c^T y^*$, setze $y^* := y'$.

Liefert das Orakel ein d mit $d^T x < d^T y$ für alle $x \in P$, dann setze $a_k := -d$.

Satz 9.5 *Der* **Grötschel-Lovász-Schrijver Algorithmus** *löst das* **schwache Optimierungsproblem** *mit einer Laufzeit, die polynomiell in n, $\alpha := \log\left(\frac{R^2\gamma}{r\epsilon}\right)$, $\mathrm{size}(x_0)$ und*

$$f\left(O(n^3\alpha^2 + n\mathrm{size}(x_0))\right)$$

ist wobei $f(\mathrm{size}(y))$ eine monotone obere Schranke für die Laufzeit des Orakels für das **schwachen Separierungsproblem zu** P**,** c **und** $\epsilon > 0$ *angewendet auf y ist.*

Lemma 9.6 *Es sei $n \in \mathbb{N}$ und $P \subseteq \mathbb{R}^n$ ein rationales Polytop. Sei $x_0 \in \mathbb{Q}^n$ ein innerer Punkt von P. Sei $T \in \mathbb{N}$ so, daß $\mathrm{size}(x_0) \leq \log T$ und $\mathrm{size}(x) \leq \log T$ für alle Ecken x von P gilt.*

Dann folgt

$$B(x_0, r) \subseteq P \subseteq B(x_0, R)$$

für

$$r = \frac{1}{n} T^{-379n^2}$$

und

$$R = 2nT.$$

Ist $c \in \mathbb{Z}^n$ und

$$c' = K^n c + (1, K, K^2, ..., K^{n-1})$$

für $K = 2T^{2n+1}$, so wird $\max\{c'^T x \mid x \in P\}$ in genau einen $x^* \in P$ angenommen. Es gilt $c'^T(x^* - y) > T^{-2n}$ für alle Ecken $y \neq x^*$ von P und $c^T x^* = \max\{c^T x \mid x \in P\}$.

Satz 9.7 *Sei $n \in \mathbb{N}$ und $c \in \mathbb{Z}^n$ (ist keine wesentliche Einschränkung). Sei $P \subseteq \mathbb{R}^n$ ein rationales Polytop und x_0 ein innerer Punkt von P.*

Sei $T \in \mathbb{N}$ so, daß $\mathrm{size}(x_0) \leq \log T$ und $\mathrm{size}(x) \leq \log T$ für alle Ecken x von P gilt.

Sind n, c, x_0, T und ein Orakel polynomieller Laufzeit für das **Separierungsproblem zum Polytop** *P gegeben, so kann man in einer Zeit, die polynomiell in n, $\log T$ und $\mathrm{size}(c)$ beschränkt ist, eine Ecke x^* von P mit $c^T x^* = \max\{c^T x \mid x \in P\}$ finden.*

Bemerkung 9.8 *Wir haben gesehen "Separieren $\Rightarrow$ Optimieren". Es gilt auch die Umkehrung.*

Zu $X \subseteq \mathbb{R}^n$ definiert man

$$X^\circ = \{y \in \mathbb{R}^n \mid x^T y \leq 1 \forall x \in X\}.$$

Ist P ein Polytop mit $0 \in Inn(P)$, so ist auch P° ein Polytop mit $0 \in Inn(P^\circ)$ und $(P^\circ)^\circ = P$. Es gilt

$$y \in P^\circ \Leftrightarrow \max\{y^T x \mid x \in P\} \leq 1.$$

Kann man über P optimieren, so kann man über P° separieren. Daher kann man auch über P° optimieren und somit wieder über $(P^\circ)^\circ = P$ separieren.

10 Das Verfahren von Karmarkar

Karmarkar (1984) betrachtet LP's in der sog. Karmarkar Normalform (KNF)

$$(KNF)\, \min\{x_1 \mid Ax = 0, e^T x = n, x \geq 0\}$$

wobei $A \in \mathbb{R}^{m\times n}$ und $x, e = (1, 1, 1, ..., 1) \in \mathbb{R}^n$ sind und zusätzlich folgende Annahmen gelten

(A1) $Ae = 0$.

(A2) Der optimale Wert des LP ist 0.

Sei (Bild)

$$\Sigma = \left\{ x \in \mathbb{R}^n \mid e^T x = \sum_{i=1}^{n} x_i = n, x \geq 0 \right\}$$

und sei $B^*(e, \rho)$ die $(n-1)$-dimensionale Kugel in der affinen Hülle von Σ mit Mittelpunkt $e \in \Sigma$ und Radius ρ, d.h.

$$B^*(e, \rho) = \{x \in \mathbb{R}^n \mid e^T x = n,\ ||x - e|| \leq \rho\}.$$

Seien

$$R = \min\{\rho \mid \Sigma \subseteq B^*(e, \rho)\}$$

und

$$r = \max\{\rho \mid B^*(e, \rho) \subseteq \Sigma\}.$$

Sei Σ° das relative Innere von Σ, d.h.

$$\Sigma^\circ = \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^{n} x_i = n, x > 0 \right\}.$$

Karmarkar's Algorithmus erzeugt eine Folge $x^{(0)}$, $x^{(1)}$, $x^{(2)}$,... von Punkten in Σ°. Für $k \geq 0$ sei $D_k = \operatorname{diag}(x^{(k)}) = \operatorname{diag}(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)})$.

Algorithmus 10.1 *Karmarkar's Algorithmus.*
Geg.: *Ein LP in (KNF). Ein* $\alpha \in (0, 1)$ *und* $q \in \mathbb{N}$.
Alg.: *Setze* $x^{(0)} := e$, $k := 0$.
Solange $x_1^{(k)} > 2^{-p}$ *führe die folgenden drei Schritte aus*

Schritt 1: Bestimme den eindeutigen Punkt

$$\tilde{x} \in \{x \in B^*(e, \alpha r) \mid AD_k x = 0\},$$

der x_1 *minimiert.*

Schritt 2: Setze

$$D_{k+1} := \frac{n}{\sum_{i=1}^{n} x_i^{(k)} \tilde{x}_i} D_k \operatorname{diag}(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n).$$

Schritt 3: Setze $k := k + 1$.

Bemerkung 10.2 *Geometrische Interpretation von Karmarkar's Algorithmus.*

Sei $x^{(k)}$ bereits bestimmt. Die Abbildung

$$T_k(x) = \frac{n}{\sum_{i=1}^n x_i x_i^{(k)}} \left(\frac{x_1}{x_1^{(k)}}, \frac{x_2}{x_2^{(k)}}, ..., \frac{x_n}{x_n^{(k)}} \right)$$

bildet Σ wieder bijektiv auf Σ ab und ihre Umkehrabbildung ist

$$T_k^{-1}(y) = \frac{n}{\sum_{i=1}^n y_i x_i^{(k)}} \left(y_1 x_1^{(k)}, y_2 x_2^{(k)}, ..., y_n x_n^{(k)} \right).$$

Führen wir den Koordinatenwechsel $x \to y = T_k(x)$ durch, so gilt

$$Ax = 0 \Rightarrow AT_k^{-1}(y) = 0 \Rightarrow AD_k y = 0.$$

Da $T_k(x^{(k)}) = e$, folgt $AD_k e = 0$.

Die zu minimierende Funktion $x \to x_1$ entspricht in y-Koordinaten der Funktion

$$y \to \frac{n x_1^{(k)}}{\sum_{i+1}^n y_i x_i^{(k)}} y_1,$$

die auf Σ nicht-negativ ist und genau dann gleich 0 ist, wenn $y_1 = 0$ gilt.

Das ursprüngliche LP ist daher äquivalent zu dem LP

$$\min\{y_1 \mid AD_k y = 0, e^T y = n, y \geq 0\}$$

wobei zusätzlich folgende Annahmen gelten

(A1) $AD_k e = 0$.

(A2) Der optimale Wert des LP ist 0.

Insb. ist dieses LP wieder in (KNF). Der Punkt $\tilde{x}$ minimiert y_1 nicht über $\{y \mid AD_k y = 0, e^T y = n, y \geq 0\}$ sondern nur über $\{y \mid AD_k y = 0, y \in B^*(e, \alpha r)\}$. Man setzt $x^{(k+1)} = T_k^{-1}(\tilde{x})$ und erhält

$$D_{k+1} = \mathrm{diag}(x^{(k+1)}) = \frac{n}{\sum_{i+1}^n \tilde{x}_i x_i^{(k)}} \mathrm{diag}(x^{(k)}) \mathrm{diag}(\tilde{x}).$$

Satz 10.3 *Für $\alpha = \frac{1}{2}$ terminiert Karmarkar's Algorithmus nach $O(nq)$ vielen Schritten.*

Lemma 10.4 *Der optimale Punkt in Schritt 1 von Karmarkar's Algorithmus ist*

$$\tilde{x} = e - \alpha r \frac{v}{||v||}$$

wobei v die Orthogonalprojektion von $e_1 = (1, 0, 0, ..., 0)$ auf

$$U := \{x \in \mathbb{R}^n \mid AD_k x = 0, e^T x = 0\}$$

ist.

Sei $f : \Sigma^\circ \to \mathbb{R}$ mit

$$f(x) = \frac{x_1^n}{x_1 x_2 x_3 ... x_n}.$$

Es gilt für $x \in \Sigma^\circ$, daß $x_1 x_2 ... x_n \leq \left(\frac{x_1+x_2+...+x_n}{2}\right)^n = 1$ und damit $f(x) \geq x_1^n$.

Lemma 10.5 *Für $x, y \in \Sigma^\circ$ gilt*

$$\frac{f(T_k(x))}{f(T_k(y))} = \frac{f(x)}{f(y)}.$$

Lemma 10.6 *Es gilt (mit obigen Bezeichnungen) $f(\tilde{x}) \leq \frac{e^{-2\alpha}}{1-\alpha}$.*

10.1 Überführung eines LP in Karmarkar Normalform

Sei ein (rationales) LP (P) mit

$$(P) \min\{c^T x \mid Ax \geq b, x \geq 0\}$$

gegeben. Indem man ggf. geeignete Ungleichungen der Form $x_i \leq 2^{\mathrm{poly}(\mathrm{size}(A),\mathrm{size}(b))}$ hinzufügt, kann man o.B.d.A. annehmen, daß

$$\{x \mid Ax \geq b, x \geq 0\}$$

beschränkt ist und daher (P) entweder optimal lösbar ist oder gar keine zulässigen Lösungen besitzt. (Dabei bleibt die Kodierungslänge polynomiell beschränkt.)

Zu (P) konstruiert man mit Hilfe des Dualitätssatzes ein LP (P') mit

$$(P') \min\{x_1 \mid A'x = b', x \geq 0\},$$

für das $e = (1, 1, ..., 1) \in \mathbb{R}^{n-1}$ eine zulässige Lösung und (P) genau dann (optimal) lösbar ist, wenn (P') den optimalen Wert 0 hat. (Vgl. [11]. Man wähle $(1, \bar{x}, \bar{y}, \bar{u}, \bar{v}) = e$. Nach Konstruktion ist (P') immer optimal lösbar.)

Wie oben können wir o.B.d.A. annehmen, daß

$$\{x \mid A'x = b', x \geq 0\}$$

beschränkt ist. Insgesamt bleibt die Kodierungslänge von (P') polynomiell beschränkt in der Kodierungslänge von (P).

Nun führt man mittels

$$(y_1, y_2, ..., y_n) := T(x_1, x_2, ..., x_{n-1}) = \frac{n}{1 + \sum_{i=1}^{n-1} x_i} (x_1, x_2, ..., x_{n-1}, 1)$$

neue Koordinaten ein. Die Abbildung T bildet dabei $[0,\infty)^{n-1}$ bijektiv auf $\Sigma_0 = \{y \mid y \in \Sigma, y_n > 0\}$ ab.

Es ergibt sich ein weiteres LP (P'') mit

$$(P'')\min\{y_1 \mid [A' \mid -b']y = 0, y \in \Sigma\},$$

für das $e = (1,1,...,1) \in \mathbb{R}^n$ eine zulässige Lösung ist. Dabei gilt, daß (P') genau dann den optimalen Wert 0 hat, wenn (P'') den optimalen Wert 0 hat.

Besitzt (P) keine zulässigen Lösungen, so ist der optimale Wert von (P') von 0 verschieden. Ist x nun eine optimale Lösung von (P'), so folgt

$$x_1 \geq 2^{-L}$$

und

$$x_2, ..., x_{n-1} \leq 2^L$$

mit

$$L = \text{poly}(\text{size}(A), \text{size}(b))$$

für ein geeignetes Polynom. Damit folgt

$$y_1 = \frac{nx_1}{1 + \sum_{i=1}^{n-1} x_i} \geq \frac{n2^{-L}}{1 + \sum_{i=1}^{n-1} 2^L} \geq 2^{-2L}.$$

Wendet man nun den Algorithmus von Karmarkar auf (P'') an und wählt $q = 2L$, so kann man die Lösbarkeit von (P) wie folgt entscheiden: Liefert der Algorithmus von Karmarkar einen Vektor y mit $y_1 < 2^{-2L}$, so ist (P) lösbar. Anderenfalls kann (A2) nicht erfüllt gewesen sein und (P) ist nicht lösbar.

Insgesamt ergibt sich, daß man bel. rationale LPs mit Hilfe des Algorithmus von Karmarkar in polynomieller Zeit lösen kann.

11 Beispiele für neuere Innere-Punkt-Verfahren

In diesem Abschnitt betrachten wir ein Innere-Punkt-Verfahren, das sich aus dem Newton-Verfahren zur Bestimmung von Nullstellen motiviert

$$\text{“}f(x) \approx f(x^{(k)}) + Df(x^{(k)})(x - x^{(k)}) \to 0 \Rightarrow x^{(k+1)} := x^{(k)} - Df(x^{(k)})^{-1} f(x^{(k)}).\text{''}$$

Wir betrachten dazu folgendes duale Paar linearer Programme

$$(P) \min\{c^T x \mid Ax = b, x \geq 0\}$$

und

$$(D) \max\{b^T y \mid A^T y + x = c, s \geq 0\}$$

mit $A \in \mathbb{R}^{m \times n}$, $\operatorname{rang}(A) = m$, $b \in \mathbb{R}^m$ und $c \in \mathbb{R}^n$ und erzeugen eine Folge

$$(x^{(k)}, y^{(k)}, s^{(k)}) \in \{(x, y, s) \mid Ax = b, A^T y + s = c, x > 0, s > 0\}$$

“innerer” Punkte, die gegen optimale Lösungen konvergieren.

Da (schwache Dualität)

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

sind nach dem Dualitätssatz die Vektoren (x, y, s) optimaler Lösungen von (P) und (D) genau die Lösungen der nicht-linearen Gleichung

$$\Psi_0(x, y, s) := \begin{pmatrix} Ax - b \\ A^T y + s - c \\ Xs \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

wobei $X = \operatorname{diag}(x)$ und $S = \operatorname{diag}(s)$.

Lemma 11.1 *Ist* $\operatorname{rang}(A) = m$ *und* $(x, s) > 0$, *so ist die Jacobi-Matrix* $D\Psi_0(x, y, s)$ *von* Ψ_0 *in Punkt* (x, y, s) *regulär.*

Beweis: Es gilt

$$D\Psi_0(x, y, s) = \begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{pmatrix}.$$

Wir nehmen an, daß $D\Psi_0(x, y, s)(u, v, w) = 0$ gilt. Aus $Au = 0$ und $A^T v + w$ folgt $u^T w = -u^T A^T v = 0$. Aus $Su + Xw = 0$ folgt $u = -S^{-1} X w$ und daher $0 = u^T w = -w^T X S^{-1} w$. Wegen $(x, s) > 0$ impliziert dies $w = 0$ und $u = 0$. Wegen $\operatorname{rang}(A) = m$ folgt nun aus $A^T v + w = A^T v = 0$ auch $v = 0$. □

Bemerkung 11.2 *Typischerweise konvergiert das Newton-Verfahren schnell (quadratisch) gegen eine Nullstelle, falls man die Iterationen in ausreichender Nähe dieser Nullstelle beginnt und die Jacobi-Matrix in der Nullstelle regulär ist. Dies ist hier nicht gegeben, da alle optimalen Lösungen (x, y, s) die Bedingung $x^T s = 0$ erfüllen und daher $D\Psi_0(x, y, s)$ nicht regulär ist.*

Daher betrachtet man statt

$$\Psi_0(x, y, s) = 0$$

das Problem

$$\Psi_\mu(x, y, s) = 0$$

wobei

$$\Psi_\mu(x, y, s) := \begin{pmatrix} Ax - b \\ A^T y + s - c \\ Xs - \mu e \end{pmatrix}$$

für ein $\mu > 0$ und $e = (1, 1, ..., 1)$ gilt.

Wegen

$$D\Psi_\mu(x, y, s) = D\Psi_0(x, y, s)$$

und

$$Xs > 0 \Rightarrow (x, s) > 0$$

ist in den Lösungen

$$(x(\mu), y(\mu), s(\mu))$$

von $\Psi_\mu(x, y, s) = 0$ die Jacobi-Matrix $D\Psi_\mu$ regulär. Wenn man μ bei jedem Schritt des Newton-Verfahrens passend verkleinert, besteht die Hoffnung, daß die entstehenden Punkte gegen eine optimale Lösung konvergieren.

Wir motivieren nun einen Newton-Schritt des Verfahrens. Dazu seien $x > 0, y, s > 0$ so gegeben, daß

$$Ax = b, A^T y + s = c$$

gilt. Es gilt

$$\begin{aligned}
\Psi_\mu(x + \Delta x, y + \Delta y, s + \Delta s) &= \begin{pmatrix} Ax + A\Delta x - b \\ A^T y + A^T \Delta y + s + \Delta s - c \\ Xs + X\Delta s + S\Delta x + \Delta X \Delta s - \mu e \end{pmatrix} \\
&\approx \begin{pmatrix} Ax + A\Delta x - b \\ A^T y + A^T \Delta y + s + \Delta s - c \\ Xs + X\Delta s + S\Delta x - \mu e \end{pmatrix}.
\end{aligned}$$

Wie beim Newton-Verfahren setzen wir den letzten Term $= 0$. Dies führt zu

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -r \end{pmatrix}$$

für

$$r = Xs - \mu e.$$

Lemma 11.3 *Für $D^2 = XS^{-1}$ (beachte $x, s > 0$) und $q = DX^{-1}r$ gilt*

$$\begin{aligned}
\Delta y &= (AD^2A^T)^{-1}ADq \\
\Delta x &= D^2A^T\Delta y - Dq \\
\Delta s &= -D^{-1}q - D^{-2}\Delta x.
\end{aligned}$$

Beweis: Da XS^{-1} eine Diagonalmatrix mit positiven Einträgen auf der Diagonalen ist, ist D wohldefiniert. Es gilt $D^2 = XS^{-1} = S^{-1}X$ und $r = XD^{-1}q$. Da $\mathrm{rang}(A) = m$ gilt, ist die symmetrische Matrix AD^2A^T positiv definit und insb. invertierbar.

Es gilt $A^T\Delta y + \Delta s = 0$ und $S\Delta x + X\Delta s = -r$. Daraus folgt $\Delta s = -A^T\Delta y$ und $\Delta x = -S^{-1}r - S^{-1}X\Delta s$. Somit folgt mit $A\Delta x = 0$, daß

$$\begin{aligned}
A(-S^{-1}r + S^{-1}XA^T\Delta y) &= -AS^{-1}XD^{-1}q + AS^{-1}XA^T\Delta y \\
&= -ADq + AD^2A^T\Delta y = 0
\end{aligned}$$

und daher $\Delta y = (AD^2A^T)^{-1}ADq$.

Nun folgt

$$\begin{aligned}
\Delta x &= -S^{-1}r - S^{-1}X\Delta s \\
&= -S^{-1}XD^{-1}q + S^{-1}XA^T\Delta y \\
&= -Dq + D^2A^T\Delta y
\end{aligned}$$

und

$$\begin{aligned}
-D^{-1}q - D^{-2}\Delta x &= -D^{-1}q - D^{-2}(-Dq + D^2A^T\Delta y) \\
&= -D^{-1}q + D^{-1}q + A^T\Delta y \\
&= A^T\Delta y \\
&= \Delta s.
\end{aligned}$$

□

Lemma 11.4 *Sind $A \in \mathbb{R}^{m\times n}$ mit $\mathrm{rang}(A) = m$ und $D \in \mathbb{R}^{n\times n}$ symmetrisch und regulär, so ist die Abbildung*

$$x \mapsto \Pi_R x$$

mit

$$\Pi_R := DA^T(AD^2A^T)^{-1}AD$$

die Orthogonalprojektion auf

$$R := \mathrm{Bild}(DA^T) = \{DA^Tw \mid w \in \mathbb{R}^m\}.$$

Mit obigen zwei Lemmata ergibt

$$DA^T \Delta y = \Pi_R q.$$

Sind

$$N := \text{Kern}(AD) = \{y \in \mathbb{R}^n \mid ADy = 0\}$$

und

$$x \mapsto \Pi_N x$$

die Orthogonalprojektion auf N, so gilt (LA 1)

$$\Pi_N = I - \Pi_R$$

und daher

$$\Delta x = D(\Pi_R q - q) = -D\Pi_N q$$

und

$$\Delta s = -D^{-1}(q + D^{-1}\Delta x) = -D^{-1}\Pi_R q.$$

Wir kommen nun zur Analyse des Newton-Schrittes

$$(x, y, s) \to (x + \Delta x, y + \Delta y, s + \Delta s).$$

Es gilt

$$\begin{aligned} \tilde{r} &= (X + \Delta X)(s + \Delta s) - \mu e \\ &= Xs + X\Delta s + S\Delta x + \Delta X \Delta s - \mu e \\ &= Xs - r + \Delta X \Delta s - \mu e \\ &= \Delta X \Delta s \\ &= \Delta \tilde{X} \Delta \tilde{s} \end{aligned}$$

für

$$\Delta \tilde{x} := -D^{-1}\Delta x = \Pi_N q$$

und

$$\Delta \tilde{s} := -D\Delta s = \Pi_R q.$$

Lemma 11.5 *(Seien alle Bezeichnungen wie oben.)*
Sind $x > 0$, y, $s > 0$, $\mu > 0$ und $\beta \in [0, \frac{1}{2}]$ so, daß $||r|| \leq \beta\mu$ gilt, dann folgt $x + \Delta x > 0$, $s + \Delta s > 0$ und $||\tilde{r}|| \leq \beta^2 \mu$.
(Wie typisch beim Newton-Verfahren wird der relative Fehler $\frac{||Xs - \mu e||}{\mu} = \frac{||r||}{\mu}$ bei jedem Schritt quadriert).

Beweis: Es folgt wegen

$$DX^{-1} = (XS^{-1})^{\frac{1}{2}} X^{-1} = (X^{-1}S^{-1})^{\frac{1}{2}} = (R+\mu I)^{-\frac{1}{2}}$$

für $R = \mathrm{diag}(r)$ mit $Xs = r + \mu e$, daß

$$||DX^{-1}||^2 = ||R+\mu I||^{-1} = \max_{1\le i\le n} \frac{1}{|r_i+\mu|} \le \frac{1}{\mu(1-\beta)}.$$

Nun folgt

$$\tilde{\beta} := ||q|| = ||DX^{-1}r|| \le \beta\sqrt{\frac{\mu}{1-\beta}},$$

$$||\Delta\tilde{x}|| = \tilde{\beta}\cos\theta$$

und

$$||\Delta\tilde{s}|| = \tilde{\beta}\sin\theta,$$

wobei θ der Winkel zwischen q und $\Pi_N q$ ist.

Es ergibt sich

$$\begin{aligned}
||\tilde{r}|| &= ||\Delta\tilde{X}\Delta\tilde{s}|| \\
&\le \tilde{\beta}^2\cos\theta\sin\theta \\
&\le \tilde{\beta}^2|\cos\theta\sin\theta| \\
&= \frac{1}{2}\tilde{\beta}^2|\sin(2\theta)| \\
&\le \frac{1}{2}\tilde{\beta}^2 \\
&\le \frac{\mu\beta^2}{2(1-\beta)} \\
&\le \mu\beta^2.
\end{aligned}$$

Weiter folgt

$$\begin{aligned}
||X^{-1}\Delta x|| &= ||X^{-1}D\Delta\tilde{x}|| \\
&\le ||X^{-1}D||||\Delta\tilde{x}|| \\
&\le ||X^{-1}D||\tilde{\beta} \\
&\le \sqrt{\frac{1}{\mu(1-\beta)}}\beta\sqrt{\frac{\mu}{1-\beta}} \\
&= \frac{\beta}{1-\beta} \\
&\le 1
\end{aligned}$$

was

$$x + \Delta x \ge 0$$

impliziert. Analog folgt

$$s + \Delta s \geq 0.$$

Wegen

$$||\tilde{r}|| < \mu$$

folgt

$$(X + \Delta X)(s + \Delta s) = \mu e - \tilde{r} > 0$$

und somit $x + \Delta x > 0$ und $s + \Delta s > 0$. □

Algorithmus 11.6 *Es seien (P) und (D) wie oben.*
Es seien $x^{(0)} > 0$, $y^{(0)}$, $s^{(0)} > 0$ und $\mu^{(0)} > 0$ so gegeben, daß

$$Ax^{(0)} = b, A^T y^{(0)} + s^{(0)} = c, r^{(0)} := X^{(0)} s^{(0)} - \mu^{(0)} e$$

und

$$||r^{(0)}|| \leq \frac{\mu^{(0)}}{2}$$

gilt.
Sei $\epsilon > 0$ gegeben.

(1) Setze $k := 0$.

(2) Setze

$$x^{(k+1)} := x^{(k)} + \Delta x,$$
$$y^{(k+1)} := y^{(k)} + \Delta y$$

und

$$s^{(k+1)} := s^{(k)} + \Delta s$$

mit Δx, Δy und Δs wie in Lemma 11.3.

(3) Setze $\mu^{(k+1)} := \mu^{(k)}(1 - \frac{1}{6\sqrt{n}})$.

(4) Setze $k := k + 1$.

(5) Falls $\mu_k \leq \frac{\epsilon}{n}$ gilt, terminiere mit $(x, y, s) = (x^{(k)}, y^{(k)}, s^{(k)})$, sonst gehe zu Schritt 2.

Satz 11.7 *Der Algorithmus 11.6 terminiert nach höchstens*

$$6\sqrt{n} \ln \left(\frac{n\mu^{(0)}}{\epsilon} \right)$$

vielen Iterationen mit einer strikt zulässigen Näherungslösung x zu (P) und (y, s) zu (D), deren Dualitätslücke

$$c^T x - b^T y \leq 2\epsilon$$

erfüllt.

Beweis: Wir zeigen zunächst per Induktion, daß $||r^{(k)}|| \leq \frac{\mu^{(k)}}{2}$ für alle $k \geq 0$ gilt. Für $k = 0$ ist dies klar. Für $k \geq 1$ folgt per Induktion und mit Lemma 11.5, daß

$$\begin{aligned}
||r^{(k+1)}|| &= ||X^{(k+1)}s^{(k+1)} - \mu^{(k+1)}e|| \\
&= ||X^{(k+1)}s^{(k+1)} - \mu^{(k)}e + (\mu^{(k)} - \mu^{(k+1)})e|| \\
&= ||\tilde{r}^{(k)} + (\mu^{(k)} - \mu^{(k+1)})e|| \\
&= ||\tilde{r}^{(k)} + \frac{\mu^{(k)}}{6\sqrt{n}}e|| \\
&\leq ||\tilde{r}^{(k)}|| + \frac{\mu^{(k)}}{6\sqrt{n}}||e|| \\
&\leq \mu^{(k)}\left(\frac{1}{4} + \frac{\sqrt{n}}{6\sqrt{n}}\right) \\
&= \frac{5}{12}\mu^{(k)} \\
&\leq \frac{1}{2}\mu^{(k+1)}.
\end{aligned}$$

Wegen Lemma 11.5 sind alle Iterierten $(x^{(k)}, y^{(k)}, s^{(k)})$ strikt zulässige Lösungen von (P) und (D).

Da

$$\ln(\mu^{(k)}) = k \ln\left(1 - \frac{1}{6\sqrt{n}}\right) + \ln(\mu_0) \leq -\frac{k}{6\sqrt{n}} + \ln(\mu_0),$$

folgt aus

$$k \geq 6\sqrt{n} \ln\left(\frac{n\mu^{(0)}}{\epsilon}\right),$$

daß $\mu^{(k)} \leq \frac{\epsilon}{n}$ gilt.

Schließlich gilt

$$c^T x - b^T y = x^T s = e^T(Xs) = e^T(\mu e + r) \leq \sqrt{n}\mu + ||r||_1 \leq \sqrt{n}\mu + n||r||$$

woraus die letzte Abschätzung der Aussage folgt. □

Bemerkung 11.8 *Bricht man den Algorithmus 11.6 nicht ab, sondern erzeugt eine unendliche Folge von Iterierten, so entsprechen alle Häufungspunkte dieser Folge optimale Lösungen von (P) und (D).*

Bemerkung 11.9 *(Zentralpfad, analytisches Zentrum) [9] 75f.*

Man nennt die Punktmenge

$$\{(x(\mu), y(\mu), s(\mu)) \mid \mu > 0\}$$

auch zentraler Pfad. Zur Veranschaulichung der Bezeichnung als zentraler Pfad nehmen wir an, daß die Menge

$$D = \{y \mid A^T y \leq c\}$$

beschänkt ist und ein nicht-leeres Inneres

$$\mathrm{Inn}(D) = \{y \mid A^T y < c\}$$

besitzt. Wir betrachten die Punkte $y(\mu) \in \mathrm{Inn}(D)$.

Seien $a_1, ..., a_n$ *die Spalten von* A *und* $c_1, ..., c_n$ *die Komponenten von* c. *Sei*

$$\phi : \mathrm{Inn}(D) \to \mathbb{R}$$

mit

$$\phi(y) = -\sum_{i=1}^{n} \ln\left(c_i - a_i^T y\right).$$

Es gilt

$$\begin{aligned} & \frac{1}{\partial y_l}\frac{1}{\partial y_k}\left(-\ln\left(c_i - \sum_{j=1}^{n} a_{i,j} y_j\right)\right) \\ = \;& \frac{1}{\partial y_l}\left(\frac{a_{i,k}}{c_i - a_i^T y}\right) \\ = \;& \frac{a_{i,k} a_{i,l}}{(c_i - a_i^T y)^2}. \end{aligned}$$

Daher hat die Hesse-Matrix

$$\left(\frac{1}{\partial y_l}\frac{1}{\partial y_k}\phi(y)\right)_{k,l}$$

die Gestalt

$$\sum_{i=1}^{n} \frac{a_i a_i^T}{(c_i - a_i^T y)^2}.$$

Da D *beschränkt ist, existiert zu jedem* $\Delta y \neq 0$ *ein Index* i *mit* $a_i^T \Delta y \neq 0$. *Daraus folgt, daß die Hesse-Matrix positiv definit ist und somit* ϕ *in* $\mathrm{Inn}(D)$ *streng konvex ist, d.h. für alle* $y_1, y_2 \in \mathrm{Inn}(D)$ *mit* $y_1 \neq y_2$ *und* $\lambda \in (0,1)$ *gilt*

$$\phi(\lambda y_1 + (1-\lambda) y_2) < \lambda\phi(y_1) + (1-\lambda)\phi(y_2).$$

Weiter divergiert ϕ *am Rand von* D *gegen* ∞, *da einer der Terme* $c_i - a_i^T y$ *dort gegen* 0 *geht. Man nennt die Funktion* ϕ *auch Barrierefunktion für* D.

Es folgt, daß ϕ *in* $\mathrm{Inn}(D)$ *ein eindeutiges Minimum* $\bar{y}$ *besitzt, daß man analytisches Zentrum von* D *nennt.*

Der Punkt $\bar{y}$ *maximiert das Produkt der Euklidischen Abstände* $\frac{c_i - a_i^T y}{||a_i||}$ *von* y *zu den*

Hyperebenen, die durch die Nebenbedingungen gegeben sind, denn

$$\begin{aligned}
& -\sum_{i=1}^{n} \ln\left(c_i - a_i^T y\right) \\
= \quad & -\ln\left(\prod_{i=1}^{n} c_i - a_i^T y\right) \\
= \quad & -\ln\left(\prod_{i=1}^{n} ||a_i||\right) - \ln\left(\prod_{i=1}^{n} \frac{c_i - a_i^T y}{||a_i||}\right).
\end{aligned}$$

Die Funktion

$$\phi(y) - \frac{b^T y}{\mu}$$

ist ebenfalls streng konvex und ihr eindeutiges Minimum in D stimmt mit $y(\mu)$ überein. Dies sieht man leicht ein, indem man die Ableitung

$$\sum_{i=1}^{n} \frac{a_i}{c_i - a_i^T y}$$

gleich 0 *setzt,* $x_i := \frac{\mu}{c_i - a_i^T y}$ *und* $s_i := c_i - a_i^T y$ *setzt und* $\Psi_\mu(x, y, s) = 0$ *nachrechnet.*

Geht nun μ *gegen* 0, *so wird der zweite Term in* $\phi(y) - \frac{b^T y}{\mu}$ *immer dominanter und* $y(\mu)$ *sollte gegen eine optimale Lösung von* (D) *konvergieren.*

12 Anwendung: Netzwerksimplex

In diesem Abschnitt beschreiben wir eine Variante des Simplexalgorithmus für das sogenannte "Min Cost Flow" Problem.

Definition 12.1 *Ein gerichteter Graph D ist ein Paar (V, E) bestehend aus einer endlichen Menge V ("Ecken") und einer Menge $E \subseteq V \times V$ ("gerichtete Kanten").*

Sei $D = (V, E)$ ein gerichteter Graph und sei $P: v_0 v_1 v_2 ... v_l$ eine endliche Folge von Ecken $v_i \in V$ mit $(v_i, v_{i+1}) \in E$ für $0 \leq i \leq l-1$. P ist ein gerichteter Weg der Länge l, falls $|\{v_0, v_1, ..., v_l\}| = l+1$ gilt. P ist ein gerichteter Kreis der Länge l, falls $|\{v_0, v_1, ..., v_{l-1}\}| = l$ und $v_0 = v_l$ gilt. Ungerichtete Wege und Kreise definiert man analog wobei die Bedingung "$(v_i, v_{i+1}) \in E$" durch "$(v_i, v_{i+1}) \in E$ oder $(v_{i+1}, v_i) \in E$" ersetzt wird.

D ist stark zusammenhängend, falls es in D zwischen je zwei Ecken einen gerichteten Weg gibt.

D ist zusammenhängend, falls es in D zwischen je zwei Ecken einen ungerichteten Weg gibt.

Eine Menge $T \subseteq E$ ist ein spannender Baum, falls (V, T) zusammenhängend ist und keinen ungerichteten Kreis enthält. Eine Menge $T \subseteq E$ ist ein gerichteter spannender Baum mit Wurzel $r \in V$, falls (V, T) keinen ungerichteten Kreis enthält und in (V, T) zu jeder Ecke $u \in V \setminus \{r\}$ ein gerichteter Weg von r nach u existiert.

Die Inzidenzmatrix $A = (a_{v,e})_{v \in V, e \in E} \in \mathbb{R}^{V \times E}$ ist definiert durch

$$a_{v,e} = \begin{cases} -1, & e = (v, u) \text{ für ein } u \in V \\ 1, & e = (u, v) \text{ für ein } u \in V \\ 0, & \text{sonst} \end{cases}$$

Lemma 12.2 *Ist $D = (V, E)$ ein gerichteter Graph mit Inzidenzmatrix A und $F \subseteq E$, so enthält (V, F) genau dann keinen ungerichteten Kreis, wenn die Spalten von A mit Index in F $\{a_{\cdot,e} \mid e \in F\}$ linear unabhängig sind.*

Folgerung 12.3 *Ist $D = (V, E)$ ein zusammenhängender, gerichteter Graph mit Inzidenzmatrix A so entsprechen die Basen des Spaltenraumes von A genau den spannenden Bäumen von D.*

Weiter gilt $\text{rang}(A) = |V| - 1$.

Sei $D = (V, E)$ ein gerichteter Graph mit Inzidenzmatrix A. Sei $c, u \in \mathbb{R}^E$ mit $u > 0$. Seien $r, s \in V$, $f \in \mathbb{R}_{>0}$ und $b \in \mathbb{R}^E$ mit $b_r = -f$, $b_s = f$ und $b_u = 0$ für $u \in V \setminus \{r, s\}$.

Das Min Cost Flow Problem ist folgendes LP

$$(P) \min\{c^T x \mid x \in \mathbb{R}^E, Ax = b, 0 \leq x \leq u\}.$$

(Interpretieren) Wir machen folgende Annahmen.

- D besitzt einen gerichteten spannenden Baum $T \subseteq E$ mit Wurzel r.

- $e^* = (r,s) \in E$ mit $u_{e^*} = f$ und $c_{e^*} > \left|\sum_{e \in E \setminus \{e^*\}} c_e\right|$.

Eine zulässige Lösung x von (P) ($Ax = b$ und $0 \leq x \leq u$) nennt man Baumlösung, falls für

$$F = \{e \in E \mid 0 < x_e < u_e\}$$

der gerichtete Graph (V, F) keine ungerichteten Kreise enthält, d.h. F kann zu einem spannenden Baum T_F von D erweitert werden. Setzt man $x_{e^*} = f$ und $x_e = 0$ für $e \in E \setminus \{e^*\}$, so ist x eine Baumlösung.

Lemma 12.4 *Eine zulässige Lösung von (P) ist genau dann eine Ecke von $\{x \mid Ax = b, 0 \leq x \leq u\}$, wenn sie eine Baumlösung ist.*

Wir stellen das zu (P) duale LP auf.

$$c^T x \quad \geq \quad y^T(Ax) - z^T x = (A^T y - z)^T x \geq y^T b - z^T u \tag{2}$$

$$(D) \max\{y^T b - z^T u \mid y^T A - z^T \leq c^T, z \geq 0\}.$$

Da $u > 0$, gilt für alle optimalen Lösungen (y, z) von (D), daß $z_e = \max\{0, y^T A_{\cdot,e} - c_e\}$.

Sei nun $x \in \mathbb{R}^E$ eine zulässige Baumlösung. Sei $T \subseteq E$ ein spannender Baum mit

$$F = \{e \in E \mid 0 < x_e < u_e\} \subseteq T.$$

Es existiert ein eindeutiges $y \in \mathbb{R}^V$ mit $y_r = 0$ und

$$y^T A_{\cdot,e} = c_e$$

für alle $e \in T$, i.e. $y_w = y_v + c_{(v,w)}$ für $(v, w) \in T$. Setzt man

$$z_e = \max\{0, y^T A_{\cdot,e} - c_e\}$$

für alle $e \in E$, so folgt, daß (y, z) eine zulässige Lösung für (D) ist.

Lemma 12.5 *Sind x, T und (y, z) wie oben, so sind x und (y, z) genau dann optimale Lösungen von (P) und (D), wenn für alle $e = (v, w) \in E \setminus T$ folgende zwei Bedingungen gelten*

(i) $y_w > y_v + c_{(v,w)} \Rightarrow x_e = u_e$.

(i) $y_w < y_v + c_{(v,w)} \Rightarrow x_e = 0$.

Sei nun $e = (v, w) \in E \setminus T$ mit $y_w > y_v + c_{(v,w)}$ und $x_e < u_e$ (Fall 1) oder $y_w < y_v + c_{(v,w)}$ und $x_e = 0$ (Fall 2). In Fall 1 gilt $x_e = 0$ und in Fall 2 gilt $x_e = u_e$.

Die Kante e liegt in genau einem Kreis C, dessen Kanten aus $T \cup \{e\}$ stammen. Orientiere C so, daß er die Kante e im Fall 1 (Fall 2) entsprechend (entgegen) ihrer Orientierung durchläuft. Nun kann man x entlang C um $\epsilon > 0$ erhöhen ($+\epsilon$ auf Vorwärtskanten und $-\epsilon$ auf Rückwärtskanten) bis eine Kante h auf C $x_h = 0$ oder $x_h = u_h$ erfüllt.

Die neue zulässige Lösung x' ist wieder eine Baumlösung zum Baum $T' = (T \setminus \{h\}) \cup \{e\}$.

Algorithmus 12.6 *(Netzwerk Simplex Algorithmus) Beginne mit der Baumlösung x mit $x_{e^*} = f$ und $x_e = 0$ für $e \in E \setminus \{e^*\}$.*

Solange Kanten existieren, die die Bedingungen aus Lemma 12.5 verletzen, führe obige Modifikationen aus.

Bemerkung 12.7 *Wie üblich bei der Simplexmethode kann man durch geschickte Wahl der Kante e verhindern, daß der Algorithmus in einen Zyklus läuft.*

Dies impliziert, daß der Algorithmus nach endlicher Zeit terminiert. Die erzeugte Lösung ist dann natürlich optinmal.

Bemerkung 12.8 *Spezialfälle und ähnliche Probleme: Assignment Problem, Min Cost Path.*

Literatur

[1] M. Aigner, Diskrete Mathematik, Vieweg.

[2] M.S. Bazaraa, H.D. Sherali und C.M. Shetty, Nonlinear Programming - Theory and Algorithms, John Wiley & Sons, 1993.

[3] K.H. Borgwardt, Optimierung Operations Research Spieltheorie, Birkhäuser 2001.

[4] V. Chvátal, Linear Programming, Freeman, New York, 1983.

[5] G.B. Dantzig, Linear Programming and Extentions, Princeton University Press, Princeton, 1963.

[6] U. Faigle, W. Kern und G. Still, Algorithmic Principles of Mathematical Programming, Kluwer Academic Publishers, 2002.

[7] M. Grötschel, L. Lovász und L. Schrijver, Geometrical Algorithms and Combinatorial Optimization, Springer 1988.

[8] A. Iouditski, Optimisation, Vorlesungsskript, http://www-lmc.imag.fr/lmc-sms/Anatoli.Iouditski/teaching/cours.htm.

[9] F. Jarre und J. Stoer, Optimierung, Springer 2004.

[10] H. Jongen, Optimization, Skript.

[11] H. Jongen, K. Meer und E. Triesch, Optimization Theory, Kluwer Academic Publishers, 2004.

[12] M. Kochol, A note on approximation of a ball by polytopes, *Discrete Optimization* **1** (2004), 229-231.

[13] B. Korte und J. Vygen, Combinatorial Optimization: Theory and Algorithms, Springer, Zweite Auflage 2002.

[14] M. Padberg, Linear Optimization and Extensions, Springer-Verlag, Berlin, 1995.

[15] C.H. Papadimitriou und K. Steiglitz, Combinatorial Optimization: Algorithms and Complexity, Prentice Hall, 1982.

[16] R.T. Rockafellar, Convex Analysis, Princeton University Press 1970.

[17] A. Schrijver, Theory of Linear and Integer Programming, Wiley, Chichester, 1986.

[18] A. Schrijver, Combinatorial Optimization: Polyhedra and Efficiency. Springer 2002.

[19] J. Stoer und C. Witzgall, Convexity and Optimization in Finite Dimensions I, Springer, 1970.

[20] E. Triesch, Vorlesungsskript.

[21] V.V. Vazirani, Approximation Algorithms, Springer, 1999.

[22] L.A. Wolsey, Integer Programming, Wiley-Interscience, 1998.

[23] G.M. Ziegler, Lectures on Polytopes, Graduate Texts in Mathematics 152, Springer-Verlag, New York 1995,