Optimierung I

Einführung in die Optimierung

Skript zur Vorlesung

von

Prof. Dr. Mirjam Dür

Prof. Dr. Alexander Martin

Prof. Dr. Stefan Ulbrich

Wintersemester 2008/2009

TU Darmstadt

Überarbeitete Version vom 10. Februar 2009

Vorwort

Das vorliegende Skript fasst den Stoff der Vorlesung *Optimierung I* (d.h. der *Einführung in die Optimierung*) vom Wintersemester 2006/2007 zusammen. Teile des Skripts (Kapitel 3 bis 5) gehen auf Vorlesungen von Prof. Dr. Alexander Martin zur *Diskrete Optimierung I und II*, die er an der TU Darmstadt in den Jahren 2000 bis 2003 hielt, und von Prof. Martin Grötschel über *Lineare Optimierung*, die er im Wintersemester 1984/1985 an der Universität Augsburg hielt, zurück.
Kapitel 6 ist dem Buch von Grötschel, Lovasz, und Schrijver [GLS88] entnommen bzw. beruht teilweise auf einem Skript von Prof. M. Grötschel zur Vorlesung *Lineare Optimierung* vom Wintersemester 2003/2004.
Kapitel 7 und 8 orientieren sich an den Büchern [Ho79] von R. Horst, [GK99, GK00] von C. Geiger und C. Kanzow, [NW99] von J. Nocedal und S.J. Wright und an Vorlesungen zur Optimierung, die S. Ulbrich in den Jahren 2000 bis 2004 an der TU München gehalten hat.
Besonderer Dank bei der Verfassung des Skripts gilt Thorsten Materne, der große Teile davon geschrieben hat, sowie Markus Möller, der das Skript Korrektur gelesen hat. Dennoch ist das Skript noch nicht vollständig überarbeitet und es ist sicherlich nicht frei von Fehlern. Für Hinweise auf solche sind wir immer dankbar.

Darmstadt, Oktober 2008
Prof. Dr. Mirjam Dür Prof. Dr. Alexander Martin Prof. Dr. Stefan Ulbrich

Inhaltsverzeichnis

Kapitel 1

Einleitung

Optimierung beschäftigt sich damit, Minima oder Maxima einer Funktion f über einer Menge $\mathcal{X}$ zu finden. Aus der Analysis ist der Satz bekannt, dass eine stetige Funktion über einer kompakten Menge ihr Minimum und ihr Maximum in Punkten x_{min} und x_{max} annimmt. Dieser Satz ist aber ein reiner Existenzsatz. Er sagt nichts darüber aus, wie man diese Punkte x_{min} und x_{max} praktisch berechnen kann. Optimierung im weitesten Sinn beschäftigt sich mit diesem Problem.

Die Funktion, deren Minimum oder Maximum gefunden werden soll, wird Zielfunktion genannt, die Menge $\mathcal{X}$ heißt zulässige Menge. Die Elemente $x \in \mathcal{X}$ heißen zulässige Punkte oder zulässige Lösungen. Die zulässige Menge kann der ganze Raum $\mathbb{R}^n$ sein (dann spricht man von unrestringierter Optimierung bzw. Optimierung ohne Nebenbedingungen); sie kann aber auch eine Teilmenge des Raumes sein, die durch sogenannte Nebenbedingungen beschrieben wird.

In diesem Skript werden zwei verschiedene Schreibweisen für ein Minimierungsproblem verwendet:

$$\min\{f(x) : x \in \mathcal{X}\},$$

bzw.

$$\begin{array}{c} \min f(x) \\ \text{s.t. } x \in \mathcal{X}. \end{array}$$

Beides bedeutet dasselbe: Wir suchen das Minimum der Funktion f über der Menge $\mathcal{X}$. Die Abkürzung "s.t." steht für das englische "subject to", was soviel bedeutet wie "unter den Nebenbedingungen".

Oft ist die zulässige Menge durch Gleichungen und Ungleichungen beschrieben, also

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

mit Funktionen $h : \mathbb{R}^n \to \mathbb{R}^k$, $g : \mathbb{R}^n \to \mathbb{R}^m$.
Die Menge der Punkte aus $\mathcal{X}$, in denen das Minimum angenommen wird, bezeichnen wir mit $\mathrm{Argmin}(f, \mathcal{X})$. Formal:

$$\alpha = \min\{f(x) : x \in \mathcal{X}\} \quad \Longleftrightarrow \quad \mathrm{Argmin}(f, \mathcal{X}) = \{x \in \mathcal{X} : f(x) = \alpha\}.$$

Beispiel 1.1 *Wir suchen das Minimum der Funktion* $f(x) = x^3 - 7.5x^2 + 18x - 10.5$ *über der Menge aller Punkte, die folgende zwei Nebenbedingungen erfüllen:* $x - 1 \geq 0$ *und* $x^2 - 5x \leq 0$.
Die Zielfunktion in diesem Beispiel ist also die Funktion $f(x) = x^3 - 7.5x^2 + 18x - 10.5$*, die zulässige Menge ist die Menge* $\mathcal{X} = \{x \in \mathbb{R} : x - 1 \geq 0, x^2 - 5x \leq 0\} = [1, 5]$*. Formal wird das Optimierungsproblem geschrieben als*

$$\min\{x^3 - 7.5x^2 + 18x - 10.5 : x - 1 \geq 0, x^2 - 5x \leq 0\},$$

oder auch

$$\begin{array}{rrcl} \min & x^3 - 7.5x^2 + 18x - 10.5 & & \\ \text{s.t.} & x - 1 & \geq & 0 \\ & x^2 - 5x & \leq & 0. \end{array}$$

Siehe Abbildung 1.1 für eine Skizze.

Das Minimum wird im Punkt $x = 1$ *angenommen, d.h.* $\mathrm{Argmin}(f, \mathcal{X}) = \{1\}$*, der zugehörige Funktionswert ist* $f(1) = 1$*.*

In der Graphik gut zu sehen ist auch ein Phänomen, das sehr oft beobachtet werden kann: Es gibt ein sogenanntes lokales Minimum im Punkt $x = 3$. Diese Beobachtung führt uns zur folgenden Definition:

Definition 1.2 *Ein Punkt* $\bar{x} \in \mathcal{X}$ *heißt* lokaler Minimalpunkt *der Funktion* f *über der Menge* $\mathcal{X}$*, wenn eine offene Umgebung* $\mathcal{U}(\bar{x})$ *von* $\bar{x}$ *existiert, so dass*

$$f(\bar{x}) \leq f(x) \quad \forall\, x \in \mathcal{U}(\bar{x}) \cap \mathcal{X}.$$

Der Punkt $\bar{x} \in \mathcal{X}$ *heißt* globaler Minimalpunkt *der Funktion* f *über der Menge* $\mathcal{X}$*, wenn*

$$f(\bar{x}) \leq f(x) \quad \forall\, x \in \mathcal{X}.$$

Lokale *und* globale Maximalpunkte *sind analog definiert.*

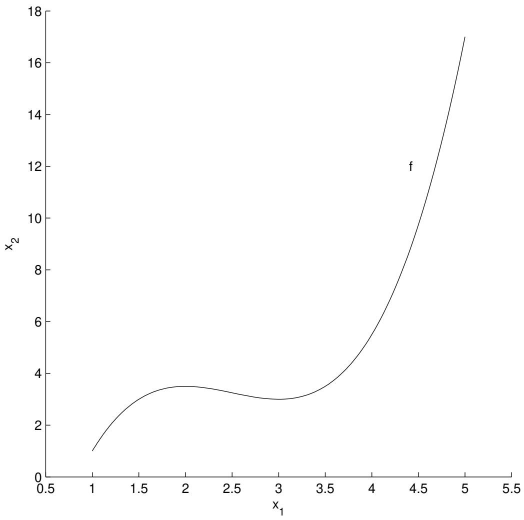

Abbildung 1.1: Die Funktion $f(x) = x^3 - 7.5x^2 + 18x - 10.5$ im Intervall $[1, 5]$.

Eine weitere wichtige Beobachtung ist die, dass jedes Maximierungsproblem auch als Minimierungsproblem geschrieben werden kann. Dabei verwendet man die Relation

$$\max\{f(x) : x \in \mathcal{X}\} = -\min\{-f(x) : x \in \mathcal{X}\}.$$

Statt also ein Maximierungsproblem $\max\{f(x) : x \in \mathcal{X}\}$ zu lösen, kann man das Minimierungsproblem $\min\{-f(x) : x \in \mathcal{X}\}$ lösen. Um dann den korrekten Wert des Maximums zu erhalten, muss die Lösung des Minimierungsproblems noch mit (-1) multipliziert werden.

Beispiel 1.3 *(Fortsetzung von Beispiel 1.1)*
Suchen wir nun das Maximum der Funktion $f(x) = x^3 - 7.5x^2 + 18x - 10.5$ *über der Menge* $\mathcal{X} = \{x \in \mathbb{R} : x - 1 \geq 0, x^2 - 5x \leq 0\} = [1, 5]$. *Wie man aus Abbildung 1.1 sieht, wird das Maximum im Punkt* $x = 5$ *angenommen, der zugehörige Zielfunktionswert ist* $f(5) = 17$.
Bestimmen wir dieses Maximum nun über den Umweg des Minimierungsproblems:

$$\begin{aligned} \max \quad & \{x^3 - 7.5x^2 + 18x - 10.5 : x - 1 \geq 0, x^2 - 5x \leq 0\} = \\ -\min \quad & \{-(x^3 - 7.5x^2 + 18x - 10.5) : x - 1 \geq 0, x^2 - 5x \leq 0\}, \end{aligned}$$

Graphisch ist dies in Abb. 1.3 dargestellt.

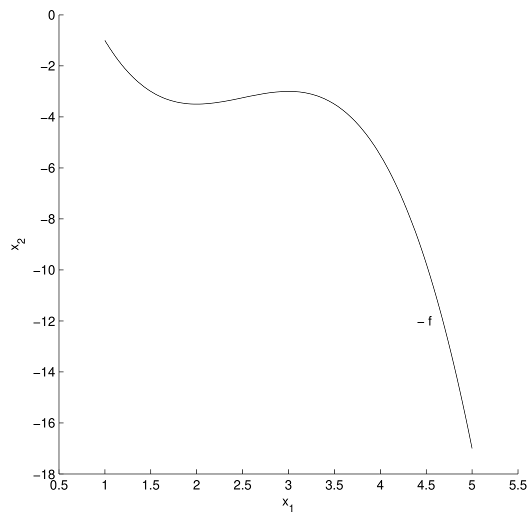

Abbildung 1.2: Die Funktion $-f(x) = -(x^3 - 7.5x^2 + 18x - 10.5)$ im Intervall $[1, 5]$.

Das Minimum von $-f$ *über* $[1, 5]$ *wird im Punkt* $x = 5$ *angenommen, der zugehörige Funktionswert ist* $f(5) = -17$*. Um wieder auf das Maximum der ursprünglichen Funktion zu kommen, multiplizieren wir jetzt* (-17) *mit* (-1)*.*

Je nachdem, welche Form die Zielfunktion und die zulässige Menge haben, ist ein Optimierungsproblem verschieden schwer zu lösen. Der folgende Überblick soll eine grobe Einteilung von Optimierungsproblemen bieten:

Lineare Optimierungsprobleme:

Von einem linearen Optimierungsproblem, auch lineares Problem (LP) genannt, spricht man, wenn sowohl die Zielfunktion als auch die Nebenbedingungen lineare Funktionen vom $\mathbb{R}^n$ nach $\mathbb{R}$ sind. Wie man sich denken kann, ist dies die einfachste Klasse von Optimierungsproblemen. Es gibt eine Reihe von Algorithmen zur Lösung von LPs, am bekanntesten ist wohl der Simplex-Algorithmus.

Ein lineares Optimierungsproblem in Standardform ist von der Gestalt:

$$\min c^T x \quad \text{s.t.} \qquad Ax = b, \quad x \geq 0$$

mit $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m,n}$, $x \in \mathbb{R}^n$.

Beispiel 1.4 *(Ein Transportproblem)*
Ein Chemieunternehmen hat m Fabriken $F_1, \ldots, F_m$ und r Verkaufsstellen $V_1, \ldots, V_r$. Jede Fabrik F_i kann pro Woche a_i Tonnen eines gewissen chemischen Produkts herstellen. a_i heißt Kapazität *der Fabrik F_i. Jede Verkaufsstelle V_j hat einen bekannten wöchentlichen* Bedarf *von b_j Tonnen des Produkts. Die Kosten, um eine Tonne des Produkts von Fabrik F_i an Verkaufsstelle V_j zu transportieren, sind c_{ij}.*

Problemstellung: *Welche Menge des Produkts muss man von jeder Fabrik zu jeder Verkaufsstelle transportieren, so dass die Kapazitäten der Fabriken eingehalten, der Bedarf aller Verkaufsstellen gedeckt und die Kosten hierbei minimal sind?*

Modellierung als Optimierungsproblem: *Sei $x_{ij} \geq 0$, $1 \leq i \leq m$, $1 \leq j \leq r$, die Zahl der Tonnen, die von F_i nach V_j transportiert werden. Dann kann man das Problem wie folgt formulieren:*

$$\begin{aligned}
\min \quad & \sum_{i=1}^{m} \sum_{j=1}^{r} c_{ij} x_{ij} && \textit{(Transportkosten)} \\
\text{s.t.} \quad & \sum_{j=1}^{r} x_{ij} \leq a_i, \quad i = 1, \ldots, m, && \textit{(Kapazitäten einhalten)} \\
& \sum_{i=1}^{m} x_{ij} \geq b_j, \quad j = 1, \ldots, r, && \textit{(Bedarf erfüllen)} \\
& x_{ij} \geq 0, \quad i = 1, \ldots, m, \quad j = 1, \ldots, r.
\end{aligned}$$

Offensichtlich handelt es sich um ein Lineares Optimierungsproblem.
In praktischen Anwendungen kommen oft noch Produktions- und Lagerhaltungskosten dazu.

Diskrete Optimierungsprobleme:

Ein diskretes Optimierungsproblem hat sehr oft eine lineare Zielfunktion. Die Schwierigkeit liegt darin, dass die zulässige Menge nun so genannte Integerbedingungen enthält, d.h. Bedingungen der Art $x_i \in \mathbb{Z}$, oder $x_i \in \{0, 1\}$. Das klingt im ersten Moment leichter, schließlich hat ein solches Problem ja nur endlich viele zulässige Punkte, man könnte also einfach alle durchprobieren und würde so das Optimum ganz einfach finden. Der Grund, weshalb das nicht möglich ist, liegt am meistens exponentiellen Wachstum der Anzahl der zulässigen Punkte, sobald die Dimension des Problems groß wird. Man braucht daher spezielle Lösungstechniken für diskrete Probleme.

Kontinuierliche Optimierungsprobleme:

sind Optimierungsprobleme, bei denen keine Integerbedingungen auftreten.

Konvexe Optimierungsprobleme:

Hier taucht zum ersten Mal der Begriff der Konvexität auf, der in Kapitel 2 ausführlich behandelt wird. Im Wesentlichen besagt er, dass bei konvexen Optimierungsproblemen jedes lokale Optimum bereits das globale Optimum ist. Es ist also möglich, Algorithmen zu entwickeln, die auf lokaler Information basieren, wie zum Beispiel den Gradienten von Zielfunktion und Nebenbedingungen.

Globale Optimierungsprobleme:

Das sind Probleme, bei denen lokale Minima auftreten, die nicht gleich dem globalen Minimum sind. Hier genügt es nicht mehr, lokale Informationen, wie Gradienten, zu betrachten. Es müssen spezielle globale Optimierungstechniken entwickelt werden.

Diese Vorlesung soll Grundlagen vermitteln, die sowohl für die diskrete als auch für die kontinuierliche Optimierung gebraucht werden. Auf dieser Basis ist später eine Spezialisierung in die eine oder andere Richtung möglich.

Beispiel 1.5 Portfoliooptimierung

Ein Investor möchte einen Betrag $B > 0$ so in ein Portfolio aus n Aktien investieren, dass die erwartete Rendite mindestens ρ und das Risiko minimal ist. Bezeichne r_i die Rendite der i-ten Aktie nach einem Jahr (dies ist eine Zufallsvariable) und $x \in \mathbb{R}^n$ mit

$$\sum_{i=1}^{n} x_i = 1, \quad x \geq 0,$$

die Anteile der Aktien am Portfolio (der Anleger investiert x_iB in Aktie i), dann ist die Rendite des Portfolios

$$R(x) = r^T x, \quad r = (r_1, \ldots, r_n)^T.$$

Wir nehmen an, dass der Zufallsvektor $r = (r_1, \ldots, r_n)^T$ den Erwartungswert $\mu \in \mathbb{R}^n$ und die Kovarianzmatrix $\Sigma \in \mathbb{R}^{n,n}$ habe. Dann ist die erwartete Rendite des Portfolios

$$E(R(x)) = \mu^T x$$

und seine Varianz

$$V(R(x)) = x^T \Sigma x.$$

Suchen wir nun das Portfolio mit erwarteter Rendite $\geq \rho$, das minimale Varianz hat, so führt dies auf das Optimierungsproblem

$$\min \quad x^T \Sigma x \quad \text{u. d. Nebenbedingung} \quad \sum_{i=1}^{n} x_i = 1, \quad x \geq 0, \quad \mu^T x \geq \rho.$$

Dies ist ein konvexes quadratisches Optimierungsproblem.
Suchen wir alternativ das Portfolio mit Varianz $\leq \nu$, das die maximale erwartete Rendite hat, so erhalten wir das Optimierungsproblem

$$\max \quad \mu^T x \quad \text{u. d. Nebenbedingung} \quad \sum_{i=1}^{n} x_i = 1, \quad x \geq 0, \quad x^T \Sigma x \leq \nu.$$

Dies ist ein konvexes Optimierungsproblem mit linearen und quadratischen Nebenbedingungen.

Kapitel 2

Konvexe Mengen und Funktionen

Ein grundlegender Begriff für die gesamte Optimierung ist der Begriff der Konvexität. Wie bereits angedeutet, ist das Vorhandensein oder Nichtvorhandensein von Konvexität mitentscheidend dafür, wie schwierig ein Optimierungsproblem ist. Der Begriff "konvex" wird sowohl auf Mengen als auch auf Funktionen angewendet. Wir beginnen mit konvexen Mengen.

2.1 Konvexe Mengen

Definition 2.1 *Eine Menge $\mathcal{C} \subset \mathbb{R}^n$ heißt* konvex*, wenn mit je zwei Punkten aus $\mathcal{C}$ auch die gesamte Verbindungsstrecke zwischen den Punkten in $\mathcal{C}$ liegt, d.h. wenn aus $x_1, x_2 \in \mathcal{C}$ und $\lambda \in [0,1]$ folgt*

$$\lambda x_1 + (1-\lambda)x_2 \in \mathcal{C}.$$

Beispiel 2.2 *Folgende Mengen sind konvex (Übung):*
a) die leere Menge; die Menge, die nur aus einem einzigen Element $x \in \mathbb{R}^n$ besteht; der ganze Raum $\mathbb{R}^n$;
b) jede Hyperebene *$\mathcal{H}$, das ist eine Menge der Form*

$$\mathcal{H} = \{x \in \mathbb{R}^n : a^T x = \alpha\},$$

wobei $a \in \mathbb{R}^n, a \neq 0$ und $\alpha \in \mathbb{R}$. a heißt dabei Normalvektor *zu $\mathcal{H}$.*
c) jeder von einer Hyperebene $\mathcal{H}$ erzeugte abgeschlossene Halbraum

$$\mathcal{H}^a = \{x \in \mathbb{R}^n : a^T x \geq \alpha\},$$

und jeder dazu gehörende offene Halbraum

$$\mathcal{H}^o = \{x \in \mathbb{R}^n : a^T x > \alpha\};$$

d) die Lösungsmenge eines linearen Gleichungssystems $Ax = b$, *mit* $A \in \mathbb{R}^{m \times n}$ *und* $b \in \mathbb{R}^m$*;*
e) jede abgeschlossene (und auch jede offene) Kugel um einen gegebenen Punkt $x_0 \in \mathbb{R}^n$ *vom Radius* $\alpha > 0$

$$\mathcal{B}_\alpha(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\| \leq \alpha\}.$$

Den Punkt $z = \lambda x_1 + (1 - \lambda)x_2$ mit $\lambda \in [0, 1]$ nennt man Konvexkombination von x_1 und x_2. Bei dieser Definition muss man sich jedoch nicht auf Punktepaare beschränken, man kann allgemein Konvexkombinationen von p Punkten betrachten:

Definition 2.3 *Gegeben seien Punkte* $x_1, \ldots, x_p \in \mathbb{R}^n$ *und Zahlen* $\lambda_1, \ldots, \lambda_p \in \mathbb{R}_+$ *mit der Eigenschaft* $\sum_{i=1}^{p} \lambda_i = 1$. *Dann heißt der Punkt*

$$z = \lambda_1 x_1 + \ldots + \lambda_p x_p$$

eine Konvexkombination *der Punkte* $x_1, \ldots, x_p$. *Gilt zusätzlich* $0 < \lambda_i < 1$ *für alle* λ_i*, so heißt* z echte Konvexkombination *von* $x_1, \ldots, x_p$.

Satz 2.4 *Eine Menge* $\mathcal{C} \subset \mathbb{R}^n$ *ist konvex genau dann, wenn für jede beliebige Zahl* $p \in \mathbb{N}$ *alle Konvexkombination von* p *Punkten* $x_1, \ldots, x_p$ *aus* $\mathcal{C}$ *wieder in* $\mathcal{C}$ *enthalten ist.*

Beweis. $\boxed{(\Longrightarrow):}$ Angenommen, $\mathcal{C}$ ist konvex. Beweis mittels Induktion nach p:
Für $p = 1$ ist die Aussage offensichtlich wahr. Nehmen wir nun an, dass sie für $p > 1$ wahr ist und betrachten $x_1, \ldots, x_{p+1} \in \mathcal{C}$, $\lambda_1, \ldots, \lambda_{p+1} \in \mathbb{R}_+$ mit $\sum_{i=1}^{p+1} \lambda_i = 1$ und

$$x = \lambda_1 x_1 + \ldots + \lambda_p x_p + \lambda_{p+1} x_{p+1}.$$

Ohne Beschränkung der Allgemeinheit können wir $\lambda_{p+1} \neq 1$ annehmen (sonst wäre bereits $x = x_{p+1} \in \mathcal{C}$). Mit

$$z = \frac{\lambda_1}{1 - \lambda_{p+1}} x_1 + \ldots + \frac{\lambda_p}{1 - \lambda_{p+1}} x_p$$

können wir x ausdrücken als

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

Es gilt

$$\frac{\lambda_1}{1-\lambda_{p+1}} \geq 0, \ldots, \frac{\lambda_p}{1-\lambda_{p+1}} \geq 0 \qquad \text{und} \qquad \sum_{i=1}^{p} \frac{\lambda_i}{1-\lambda_{p+1}} = 1,$$

daher ist, laut Induktionsannahme, $z \in \mathcal{C}$. Da $\mathcal{C}$ konvex ist, ist auch $x \in \mathcal{C}$.

$\boxed{(\Longleftarrow):}$ Ist umgekehrt jede Konvexkombination von p Punkten aus $\mathcal{C}$ wieder in $\mathcal{C}$ enthalten, dann gilt das insbesondere für $p = 2$. Daher ist $\mathcal{C}$ eine konvexe Menge nach Definition 2.1. $\square$

Eine einfache und leicht zu beweisende geometrische Eigenschaft konvexer Mengen beschreibt der folgende Satz.

Satz 2.5 *Der Durchschnitt einer beliebigen Familie konvexer Mengen ist wieder eine konvexe Menge.*

Beweis. Übung. $\square$

Es ist leicht zu sehen, dass die Vereinigung konvexer Mengen im Allgemeinen nicht konvex ist.
Die Summe konvexer Mengen und deren skalare Vielfache sind konvexe Mengen:

Satz 2.6 *Seien $\mathcal{C}$ und $\mathcal{D}$ konvexe Mengen im $\mathbb{R}^n$ und $\alpha \in \mathbb{R}$. Dann sind auch die Mengen*

$$\mathcal{C} + \mathcal{D} := \{x + y : x \in \mathcal{C}, y \in \mathcal{D}\}$$

sowie

$$\alpha\mathcal{C} := \{\alpha x : x \in \mathcal{C}\}$$

konvex.

Beweis. Übung. $\square$

Betrachtet man eine nichtkonvexe Menge $\mathcal{M} \subset \mathbb{R}^n$, so lässt sich diese "konvexifizieren", indem man konvexe Obermengen von $\mathcal{M}$ betrachtet. Der Durchschnitt all dieser Mengen ist die kleinste konvexe Menge, die $\mathcal{M}$ enthält:

Definition 2.7 *Der Durchschnitt aller konvexen Mengen, die $\mathcal{M}$ enthalten, heißt die* konvexe Hülle *von $\mathcal{M}$ und wird mit* $\operatorname{conv}\mathcal{M}$ *bezeichnet.*

Der folgende Satz zeigt, dass die konvexe Hülle einer Menge $\mathcal{M}$ dasselbe ist wie die Menge aller Konvexkombinationen von Punkten aus $\mathcal{M}$.

Satz 2.8 *Die konvexe Hülle einer Menge $\mathcal{M}$ ist die Menge aller Konvexkombinationen von Punkten aus $\mathcal{M}$.*

Beweis. Übung. □

Jeder Punkt in $\operatorname{conv}\mathcal{M}$ ist also die Konvexkombination von Punkten in $\mathcal{M}$. Der folgende Satz sagt, dass man für diese Darstellung höchstens $(n+1)$ Punkte benötigt, wobei n die Dimension des Raumes ist.

Satz 2.9 (Satz von Carathéodory) *Die konvexe Hülle einer Menge $\mathcal{M} \subset \mathbb{R}^n$ ist die Menge aller Konvexkombinationen von $(n+1)$-elementigen Teilmengen von $\mathcal{M}$.*

Beweis. Sei $\bar{x} \in \operatorname{conv}\mathcal{M}$. Wegen Satz 2.8 bedeutet das: Es existieren $x_1, \ldots, x_p \in \mathcal{M}$ so, dass

$$\bar{x} = \sum_{i=1}^{p} \lambda_i x_i, \qquad \text{mit } \sum_{i=1}^{p} \lambda_i = 1, \quad \lambda_i \geq 0 \;\forall\, i.$$

Wenn bereits $p \leq n+1$ gilt, dann ist nichts mehr zu zeigen. Gilt $p > n+1$, dann zeigen wir, dass man für die Darstellung von x auf einen der p Punkte verzichten kann:
Betrachten wir dazu die $(p-1)$ Vektoren $y_i = x_i - x_p$ $(i = 1, \ldots, p-1)$. Für $p > n+1$ sind die y_i linear abhängig, d.h. es existieren Zahlen $\alpha_1, \ldots, \alpha_{p-1}$, die nicht alle verschwinden, so dass

$$\sum_{i=1}^{p-1} \alpha_i y_i = 0.$$

Anders gesagt,

$$\sum_{i=1}^{p-1} \alpha_i (x_i - x_p) = 0$$

oder

$$\sum_{i=1}^{p-1} \alpha_i x_i + \left(-\sum_{i=1}^{p-1} \alpha_i\right) x_p = 0.$$

Mit der Definition $\alpha_p = \left(-\sum_{i=1}^{p-1} \alpha_i\right)$ haben wir also

$$\sum_{i=1}^{p} \alpha_i x_i = 0, \qquad \sum_{i=1}^{p} \alpha_i = 0.$$

Wegen $(\alpha_1, \ldots, \alpha_p) \neq (0, \ldots, 0)$ gibt es daher mindestens ein $\alpha_i > 0$.
Sei nun i_0 definiert durch

$$\frac{\lambda_{i_0}}{\alpha_{i_0}} = \min\left\{\frac{\lambda_i}{\alpha_i} : \alpha_i > 0\right\}.$$

Dann gilt

$$\lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}}\alpha_i \geq 0 \quad \forall\, i, \qquad \text{und} \qquad \sum_{i=1}^{p}\left(\lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}}\alpha_i\right) = 1.$$

Somit ist

$$\bar{x} = \sum_{i=1}^{p} \lambda_i x_i = \sum_{i=1}^{p}\left(\lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}}\alpha_i\right) x_i = \sum_{\substack{i=1 \\ i \neq i_0}}^{p} \left(\lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}}\alpha_i\right) x_i$$

eine Darstellung von $\bar{x}$ als Konvexkombination von echt weniger als p Punkten aus $\mathcal{M}$. □

2.2 Extrempunkte und Extremrichtungen

Definition 2.10 *Sei $\mathcal{C} \neq \emptyset$ eine konvexe Menge im $\mathbb{R}^n$. Ein Punkt x heißt* Extrempunkt von $\mathcal{C}$, *wenn er nicht als echte Konvexkombination von verschiedenen Punkten aus $\mathcal{C}$ dargestellt werden kann, d.h. wenn aus $x = \lambda x_1 + (1-\lambda)x_2$ mit $x_1, x_2 \in \mathcal{C}$ und $\lambda \in (0,1)$ folgt: $x = x_1 = x_2$. Hat $\mathcal{C}$ nur endlich viele Extrempunkte, so nennt man diese auch* Ecken.

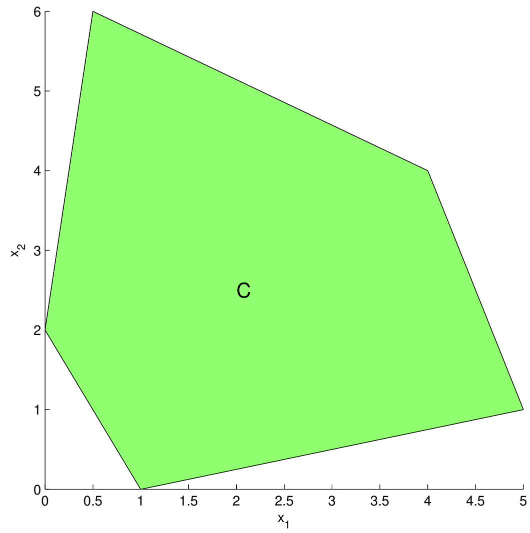

Abbildung 2.1: Konvexe Menge zu Beispiel 2.12

Beispiel 2.11 *Betrachte die konvexe Menge* $\mathcal{C} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$. *Die Menge ihrer Extrempunkte ist* $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$.

Beispiel 2.12 *Betrachte die konvexe Menge aus Abbildung 2.1. Ihre fünf Extrempunkte sind die Ecken* $(0, 2)$, $(1, 0)$, $(5, 1)$, $(4, 4)$ *und* $(0.5, 6)$.

Konvexe Mengen, die nur endlich viele Extrempunkte besitzen, nennt man Polyeder bzw. Polytope. Sie spielen eine wichtige Rolle in der linearen Optimierung und werden daher in einem eigenen Kapitel (Kapitel 3) ausführlicher behandelt.
In den obigen beiden Beispielen kann man zeigen, dass jeder Punkt der Menge als Konvexkombination der Extrempunkte darstellbar ist. Tatsächlich gilt folgender Satz, den wir ohne Beweis anführen:

Satz 2.13 *Sei* $\mathcal{C} \neq \emptyset$ *eine kompakte konvexe Menge im* $\mathbb{R}^n$. *Dann ist* $\mathcal{C}$ *die konvexe Hülle ihrer Extrempunkte.*

Beweis. Siehe, z.B. Rockafellar [Ro70]. $\square$

Dieser Satz gilt jedoch nur für kompakte konvexe Mengen. Bei unbeschränkten konvexen Mengen genügen die Extrempunkte nicht mehr zur Darstellung der gesamten Menge, wie man am folgenden Beispiel sieht:

Beispiel 2.14 *Betrachte die abgeschlossene konvexe Menge* $\mathcal{C} = \{(x,y) \in \mathbb{R}^2 : y \geq |x|\}$. *Der Ursprung ist der einzige Extrempunkt dieser Menge, die Menge besteht aber nicht nur aus Konvexkombinationen dieses Extrempunktes.*

Um auch unbeschränkte konvexe Mengen beschreiben zu können, benötigt man den Begriff der Extremrichtung.

Definition 2.15 *Sei* $\mathcal{C} \neq \emptyset$ *eine abgeschlossene konvexe Menge im* $\mathbb{R}^n$. *Ein Vektor* $d \in \mathbb{R}^n$, $d \neq 0$ *heißt* Richtung von $\mathcal{C}$, *wenn für jedes* $x \in \mathcal{C}$ *gilt:* $x + \alpha d \in \mathcal{C}$ *für jedes* $\alpha > 0$.
Zwei Richtungen d_1, d_2 *von* $\mathcal{C}$ *heißen* verschieden, *wenn* $d_1 \neq \beta d_2$ *für alle* $\beta > 0$.
Eine Richtung d *von* $\mathcal{C}$ *heißt* Extremrichtung, *wenn* d *nicht darstellbar ist als positive Linearkombination von zwei verschiedenen Richtungen; d.h. falls* $d = \beta_1 d_1 + \beta_2 d_2$ *mit* $\beta_1, \beta_2 > 0$, *dann ist* $d_1 = \gamma d_2$ *für ein* $\gamma > 0$.

Beispiel 2.16 *(Fortsetzung von Beispiel 2.14.)*
Betrachten wir nochmals die Menge $\mathcal{C} = \{(x,y) \in \mathbb{R}^2 : y \geq |x|\}$. *Richtungen*

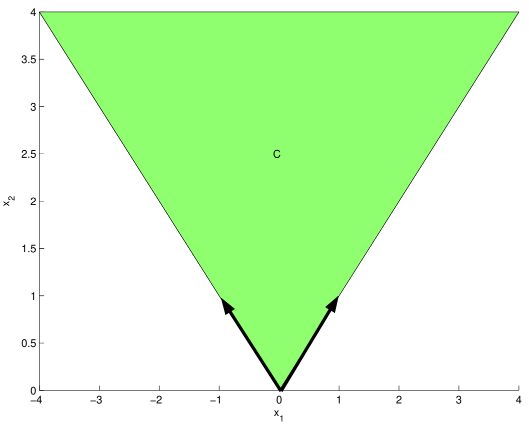

Abbildung 2.2: Konvexe Menge zu Beispiel 2.16

von $\mathcal{C}$ sind jene Vektoren, die mit dem Vektor $(0,1)^T$ einen Winkel $\leq 45°$ einschlie\'ßen. Die Extremrichtungen sind genau die Vektoren $d_1 = (-1,1)^T$ und $d_2 = (1,1)^T$. Jede andere Richtung von $\mathcal{C}$ ist als Linearkombination von d_1 und d_2 darstellbar.

Mehr über Extrempunkte und Extremrichtungen gibt es im Kapitel 3.

2.3 Trennungssätze

Trennungssätze beschreiben den anschaulich einleuchtenden Sachverhalt, dass es möglich ist, zwischen zwei disjunkte konvexe Mengen $\mathcal{C}_1$ und $\mathcal{C}_2$ eine Hyperebene zu legen, die die beiden Mengen "trennt", d.h. die den Raum so in zwei Halbräume teilt, dass $\mathcal{C}_1$ in einem Halbraum liegt und $\mathcal{C}_2$ im anderen. Man unterscheidet dabei zwischen "Trennung" und "strikter Trennung":

Definition 2.17 *Seien $\mathcal{M}_1$ und $\mathcal{M}_2$ beliebige Mengen im $\mathbb{R}^n$. Eine Hyperebene $\mathcal{H}$* trennt *$\mathcal{M}_1$ und $\mathcal{M}_2$, wenn $\mathcal{M}_1$ und $\mathcal{M}_2$ in gegenüberliegenden abgeschlossenen Halbräumen liegen, die von $\mathcal{H}$ erzeugt werden.*
Die Trennung heißt strikt*, wenn das Entsprechende für die von $\mathcal{H}$ erzeugten offenen Halbräume gilt.*

Wir beweisen zunächst eine Proposition, die zum Beweis des strikten Trennungssatzes (Satz 2.19) benötigt wird.

Proposition 2.18 *Sei $\mathcal{C}$ eine nichtleere, abgeschlossene, konvexe Menge im $\mathbb{R}^n$, welche den Ursprung nicht enthält. Dann existiert eine Hyperebene, die $\mathcal{C}$ und den Ursprung strikt trennt.*

Beweis. Sei $\mathcal{B}_\alpha$ eine abgeschlossene Kugel um den Ursprung,

$$\mathcal{B}_\alpha = \{x \in \mathbb{R}^n : \|x\| \leq \alpha\},$$

so, dass $\mathcal{C} \cap \mathcal{B}_\alpha \neq \emptyset$. Dieser Durchschnitt ist eine kompakte Menge, daher nimmt die stetige Funktion $\|x\|$ ihr Minimum über $\mathcal{C} \cap \mathcal{B}_\alpha$ in einem Punkt $\bar{x} \in \mathcal{C} \cap \mathcal{B}_\alpha$ an. Wegen $0 \notin \mathcal{C}$ gilt $\|x\| > 0$ für jedes $x \in \mathcal{C}$ und somit auch $\|\bar{x}\| > 0$.
Nun sei x ein beliebiger Punkt aus $\mathcal{C}$. Da $\mathcal{C}$ konvex ist, gilt für jedes $\lambda \in [0,1]$ $(\lambda x + (1-\lambda)\bar{x}) \in \mathcal{C}$ und

$$\|\lambda x + (1-\lambda)\bar{x}\|^2 \geq \|\bar{x}\|^2,$$

da $\bar{x}$ minimalen Abstand von 0 hat. Anders gesagt,

$$(\lambda x + (1-\lambda)\bar{x})^T(\lambda x + (1-\lambda)\bar{x}) \geq \bar{x}^T\bar{x} \qquad \forall\, \lambda \in [0,1],$$

oder

$$\lambda^2(x-\bar{x})^T(x-\bar{x}) + 2\lambda\bar{x}^T(x-\bar{x}) \geq 0 \qquad \forall\, \lambda \in [0,1]. \tag{2.1}$$

Wir zeigen jetzt, dass $\bar{x}^T(x-\bar{x}) \geq 0$: Angenommen, es wäre $\bar{x}^T(x-\bar{x}) = -\varepsilon < 0$. Dann können wir $\lambda \in (0,1)$ so klein wählen, dass

$$2\varepsilon > \lambda(x-\bar{x})^T(x-\bar{x}) > 0.$$

Dann würde folgen, dass

$$2\lambda\bar{x}^T(x-\bar{x}) < -\lambda^2(x-\bar{x})^T(x-\bar{x}),$$

ein Widerspruch zu (2.1). Daher muss für jedes $x \in \mathcal{C}$

$$\bar{x}^T(x-\bar{x}) \geq 0$$

sein, das heißt

$$\bar{x}^T x \geq \bar{x}^T\bar{x} > 0 \qquad \forall x \in \mathcal{C}.$$

Sei $\beta = \frac{1}{2}\bar{x}^T\bar{x}$. Damit trennt die Hyperebene

$$\mathcal{H} = \{x \in \mathbb{R}^n : \bar{x}^T x = \beta\}$$

strikt die Menge $\mathcal{C}$ und den Ursprung. □

Mit Hilfe dieser Proposition können wir den eigentlichen Trennungssatz für abgeschlossene Mengen beweisen.

Satz 2.19 (Strikter Trennungssatz) *Seien $\mathcal{C}_1$ und $\mathcal{C}_2$ zwei disjunkte, nichtleere, abgeschlossene, konvexe Mengen im $\mathbb{R}^n$, und sei $\mathcal{C}_2$ kompakt. Dann existiert eine Hyperebene, die $\mathcal{C}_1$ und $\mathcal{C}_2$ strikt trennt.*

Beweis. Da $\mathcal{C}_2$ kompakt ist, ist die Menge $\mathcal{C}_1 - \mathcal{C}_2$ abgeschlossen und (nach Satz 2.6) konvex. Da $\mathcal{C}_1$ und $\mathcal{C}_2$ disjunkt sind, ist der Ursprung nicht in $\mathcal{C}_1 - \mathcal{C}_2$ enthalten. Nach Proposition 2.18 existiert also eine Hyperebene

$$\mathcal{H}_{\mathcal{C}_1-\mathcal{C}_2} = \{x \in \mathbb{R}^n : \bar{x}^T x = \alpha\}$$

die den Ursprung strikt von $\mathcal{C}_1 - \mathcal{C}_2$ trennt. Hierbei minimiert $\bar{x} \in \mathcal{C}_1 - \mathcal{C}_2$ die Distanz von $\mathcal{C}_1 - \mathcal{C}_2$ zum Ursprung, und $\alpha = \frac{1}{2}\bar{x}^T\bar{x}$.
Für jedes $x \in \mathcal{C}_1 - \mathcal{C}_2$ gilt

$$\bar{x}^T x > \alpha > 0.$$

Für alle $u \in \mathcal{C}_1, v \in \mathcal{C}_2$ ist $x = u - v \in \mathcal{C}_1 - \mathcal{C}_2$ und damit

$$\bar{x}^T(u - v) > \alpha > 0.$$

Daher

$$\bar{x}^T u > \bar{x}^T v + \alpha > \bar{x}^T v \quad \forall\, u \in \mathcal{C}_1, v \in \mathcal{C}_2.$$

Es folgt, dass

$$\inf_{u \in \mathcal{C}_1} \bar{x}^T u \geq \sup_{v \in \mathcal{C}_2} \bar{x}^T v + \alpha > \sup_{v \in \mathcal{C}_2} \bar{x}^T v.$$

Es existiert daher eine Zahl β so, dass

$$\inf_{u \in \mathcal{C}_1} \bar{x}^T u > \beta > \sup_{v \in \mathcal{C}_2} \bar{x}^T v.$$

Damit trennt die Hyperebene $\{x \in \mathbb{R}^n : \bar{x}^T x = \beta\}$ strikt die Mengen $\mathcal{C}_1$ und $\mathcal{C}_2$. □

Für strikte Trennbarkeit ist die Voraussetzung, dass eine der beiden Mengen kompakt ist, unverzichtbar, wie man an folgendem Beispiel sieht:

Beispiel 2.20 *Sei $\mathcal{C}_1 := \{(x, y) \in \mathbb{R}^2 : y \leq 0\}$ und $\mathcal{C}_2 := \{(x, y) \in \mathbb{R}^2 : y \geq e^x\}$. Die Mengen sind disjunkt, beide sind konvex und abgeschlossen, trotzdem ist keine strikte Trennung möglich.*
$\mathcal{C}_1$ und $\mathcal{C}_2$ sind aber trennbar durch die Hyperebene $\mathcal{H} = \{(x, y) \in \mathbb{R}^2 : y = 0\}$.

In diesen beiden Sätzen war es wichtig, dass die vorkommenden Mengen abgeschlossen waren, deshalb war die strikte Trennung möglich. Verzichtet man auf die Abgeschlossenheit der Mengen, so muss man auch auf die Striktheit der Trennung verzichten. Dies beschreiben die folgende Proposition bzw. der nächste Satz.

Proposition 2.21 *Sei $\mathcal{C}$ eine nichtleere, konvexe Menge im $\mathbb{R}^n$, welche den Ursprung nicht enthält. Dann existiert eine Hyperebene, die $\mathcal{C}$ und den Ursprung trennt.*

Beweis. Für jedes $x \in \mathcal{C}$ sei

$$\mathcal{Y}(x) = \{y \in \mathbb{R}^n : y^T y = 1, y^T x \geq 0\}.$$

$\mathcal{Y}(x)$ ist nichtleer und abgeschlossen. Seien $x_1, \ldots, x_k$ endlich viele Punkte aus $\mathcal{C}$. Da $\mathcal{C}$ konvex ist, ist nach Satz 2.4 die Menge aller x , die darstellbar sind als

$$x = \sum_{i=1}^{k} \alpha_i x_i \qquad \text{mit} \quad \sum_{i=1}^{k} \alpha_i = 1, \quad \alpha_i \geq 0$$

eine konvexe Teilmenge von $\mathcal{C}$. Sie ist außerdem abgeschlossen. Nach Proposition 2.18 existiert daher ein $\bar{y} \neq 0$ so, dass

$$\bar{y}^T x_i > 0 \qquad \forall\, i = 1, \ldots, k.$$

O.B.d.A. nehmen wir $\bar{y}^T \bar{y} = 1$ an. Damit ist $\bar{y}$ in jeder der Mengen $\mathcal{Y}(x_i)$ enthalten, und daher

$$\bigcap_{i=1}^{k} \mathcal{Y}(x_i) \neq \emptyset.$$

Die Mengen $\mathcal{Y}(x)$ sind kompakt, da sie abgeschlossene Teilmengen der kompakten Menge $\mathcal{Y} = \{y \in \mathbb{R}^n : y^T y = 1\}$ sind. Aus der endlichen Durchschnittseigenschaft[1] folgt daher:

$$\bigcap_{x \in \mathcal{C}} \mathcal{Y}(x) \neq \emptyset.$$

Wähle nun ein beliebiges $\hat{y} \in \bigcap_{x \in \mathcal{C}} \mathcal{Y}(x)$. Dann gilt $\hat{y}^T x \geq 0$ für alle $x \in \mathcal{C}$. Daher trennt die Hyperebene

$$\{x \in \mathbb{R}^n : \hat{y}^T x = 0\}$$

die Menge $\mathcal{C}$ und den Ursprung.
Ein anderer, elementarer Beweis: Mit der Menge $\mathcal{C}$ ist auch ihr Abschluss $\overline{\mathcal{C}} = \mathcal{C} \cup \partial\mathcal{C}$ konvex.

[1]Die endliche Durchschnittseigenschaft ist ein topologischer Begriff und besagt Folgendes: Betrachte eine kompakte Menge $\mathcal{Y}$ und ein System $\mathcal{S}$ von abgeschlossenen Teilmengen von $\mathcal{Y}$ mit der Eigenschaft, dass der Durchschnitt von jeweils endlich vielen Mengen aus $\mathcal{S}$ nichtleer ist. Dann ist auch der Durchschnitt aller Mengen aus $\mathcal{S}$ nichtleer.
Nachzulesen ist dies z.B. in H. Heuser: Lehrbuch der Analysis, Teil 2 ([He03]).

1. Fall: $0 \notin \overline{\mathcal{C}}$
Dann liefert Proposition 2.18 eine Hyperebene, die $\overline{\mathcal{C}}$ und 0 strikt trennt, also erst recht $\mathcal{C}$ und 0 strikt trennt.
2. Fall: $0 \in \overline{\mathcal{C}}$, also $0 \in \partial\mathcal{C}$ wegen $0 \notin \mathcal{C}$.
Wegen $0 \in \partial\mathcal{C}$ existiert eine Folge $(x_k)_{k\in\mathbb{N}}$ mit $x_k \notin \overline{\mathcal{C}}$ für alle k und $\lim_{k\to\infty} x_k = 0$.
Nach Satz 2.19 existieren Hyperebenen $\mathcal{H}_k = \{x : a_k^T x = \alpha_k\}$, die jeweils $\overline{\mathcal{C}}$ und x_k strikt trennen. Hierbei können wir ohne Einschränkung $\|a_k\| = 1$ wählen. Es gilt also

$$a_k^T x > \alpha_k > a_k^T x_k \quad \forall\, x \in \overline{\mathcal{C}}.$$

Wegen $0 \in \overline{\mathcal{C}}$ gilt insbesondere

$$0 = a_k^T 0 > \alpha_k > a_k^T x_k \to 0 \quad \text{für } k \to \infty,$$

wobei wir $|a_k^T x_k| \le \|a_k\|\|x_k\| = \|x_k\| \to 0$ benutzt haben. Es gilt also

$$\lim_{k\to\infty} \alpha_k = 0.$$

Nun liegt (a_k) im Kompaktum $\{y \in \mathbb{R}^n : \|y\| = 1\}$ und enthält daher eine konvergente Teilfolge $(a_k)_{k\in K} \subset (a_k)$, also

$$\lim_{k\in K\to\infty} a_k = a, \quad \|a\| = 1.$$

Grenzübergang $k \in K \to \infty$ in

$$a_k^T x > \alpha_k \quad \forall\, x \in \overline{\mathcal{C}}$$

liefert nun

$$a^T x = \lim_{k\in K\to\infty} a_k^T x \ge \lim_{k\to\infty} \alpha_k = 0 \quad \forall\, x \in \overline{\mathcal{C}}.$$

Damit trennt $H = \{x : a^T x = 0\}$ die Menge $\overline{\mathcal{C}}$ von 0 und somit auch $\mathcal{C}$ von 0. $\square$

Wieder können wir mit dieser Proposition einen Trennungssatz für zwei Mengen zeigen.

Satz 2.22 *Seien $\mathcal{C}_1$ und $\mathcal{C}_2$ zwei disjunkte, nichtleere, konvexe Mengen im $\mathbb{R}^n$. Dann existiert eine Hyperebene, die $\mathcal{C}_1$ und $\mathcal{C}_2$ trennt.*

Beweis. Die Menge $\mathcal{C}_1 - \mathcal{C}_2$ erfüllt die Voraussetzungen von Proposition 2.21. Daher existiert ein Vektor $\hat{y}$ so, dass für alle $x \in \mathcal{C}_1 - \mathcal{C}_2$ gilt: $\hat{y}^T x \geq 0$. Dies ist äquivalent dazu, dass aus $u \in \mathcal{C}_1, v \in \mathcal{C}_2$ folgt: $\hat{y}^T(u - v) \geq 0$. Daher existiert eine Zahl β so, dass

$$\inf_{u \in \mathcal{C}_1} \hat{y}^T u \geq \beta \geq \sup_{v \in \mathcal{C}_2} \hat{y}^T v.$$

Die Hyperebene $\{x \in \mathbb{R}^n : \hat{y}^T x = \beta\}$ trennt daher die Mengen $\mathcal{C}_1$ und $\mathcal{C}_2$. □

2.4 Stützeigenschaften

Bisher haben wir eine konvexe Menge durch eine "innere" Eigenschaft beschrieben, nämlich durch Konvexkombinationen von ihren Elementen. Es gibt eine zweite, äquivalente Beschreibung durch "äußere" Eigenschaften:

Definition 2.23 *Sei $\mathcal{C} \neq \emptyset$ eine abgeschlossene, konvexe Menge im $\mathbb{R}^n$. Eine Hyperebene $\mathcal{H} = \{x \in \mathbb{R}^n : a^T x = \alpha\}$ heißt* Stützhyperebene von $\mathcal{C}$*, wenn*

$$\mathcal{C} \cap \mathcal{H} \neq \emptyset \qquad \textit{und} \qquad \mathcal{C} \subseteq \mathcal{H}^+ \quad \textit{oder} \quad \mathcal{C} \subseteq \mathcal{H}^-,$$

wobei

$$\mathcal{H}^+ = \{x \in \mathbb{R}^n : a^T x \geq \alpha\} \qquad \textit{und} \qquad \mathcal{H}^- = \{x \in \mathbb{R}^n : a^T x \leq \alpha\}$$

die beiden von $\mathcal{H}$ erzeugten abgeschlossenen Halbräume sind. $\mathcal{H}^+$ bzw. $\mathcal{H}^-$ heißen dann Stützhalbraum von $\mathcal{C}$*. Für $\mathcal{C} \subseteq \mathcal{H}^+$ heißt $-a$* äußere Normale von $\mathcal{H}$*, für $\mathcal{C} \subseteq \mathcal{H}^-$ heißt a* äußere Normale.

Satz 2.24 *Sei $\mathcal{C} \neq \emptyset$ eine kompakte, konvexe Menge und $a \in \mathbb{R}^n \setminus \{0\}$. Dann existiert eine Stützhyperebene von $\mathcal{C}$ mit äußerer Normale a.*

Beweis. Da $\mathcal{C} \neq \emptyset$ kompakt und die Funktion $f : \mathbb{R}^n \to \mathbb{R}$, $f(y) = a^T y$ stetig ist, existiert $\alpha = \max_{y \in \mathcal{C}} a^T y$. Damit ist

$$\mathcal{H} = \{x \in \mathbb{R}^n : a^T x = \alpha\}$$

die gesuchte Stützhyperebene. □

Satz 2.25 *Jede nichtleere, abgeschlossene, konvexe Menge $\mathcal{C}$ im $\mathbb{R}^n$ ist Durchschnitt ihrer Stützhalbräume.*

Beweis. Bezeichnen wir mit $\mathcal{H}^+$ einen Stützhalbraum, und mit $\bigcap \mathcal{H}^+$ den Durchschnitt aller Stützhalbräume. Da $\mathcal{C} \subseteq \mathcal{H}^+$ für jeden Stützhalbraum, gilt

$$\mathcal{C} \subseteq \bigcap \mathcal{H}^+.$$

Angenommen, $\mathcal{C} \subsetneq \bigcap \mathcal{H}^+$, d.h. es existiert ein $\bar{x} \in \bigcap \mathcal{H}^+ \setminus \mathcal{C}$.
Wegen Satz 2.19 existiert eine Hyperebene $\mathcal{G} = \{x : \ a^T x = \alpha\}$, die $\mathcal{C}$ und $\{\bar{x}\}$ trennt, d.h. $\mathcal{C} \subseteq \mathcal{G}^+ = \{x : \ a^T x \geq \alpha\}$, aber $\bar{x} \notin \mathcal{G}^+$. Mit $\tilde{\alpha} = \min_{x \in \mathcal{C}} a^T x$ gilt $\tilde{\alpha} \geq \alpha$ und $\mathcal{H} = \{x : \ a^T x = \tilde{\alpha}\}$ ist Stützhyperbene von $\mathcal{C}$ mit $\mathcal{C} \subseteq \mathcal{H}^+$ (Parallelverschiebung der Hyperbene $\mathcal{G}$ bis sie $\mathcal{H}$ berührt). Wegen $\bar{x} \notin \mathcal{G}^+$ gilt jedoch $a^T \bar{x} < \alpha \leq \tilde{\alpha}$, also auch $\bar{x} \notin \mathcal{H}^+$. Somit gilt $\bar{x} \notin \bigcap \mathcal{H}^+$, was im Widerspruch zu $\bar{x} \in \bigcap \mathcal{H}^+$ steht.
Daher muss $\bigcap \mathcal{H}^+ \setminus \mathcal{C} = \emptyset$ sein, und somit $\bigcap \mathcal{H}^+ = \mathcal{C}$. □

2.5 Konvexe Funktionen

Definition 2.26 *Sei $\mathcal{C} \subset \mathbb{R}^n$ konvex. Eine Funktion $f : \mathcal{C} \to \mathbb{R}$ heißt* konvex, *wenn für alle $x_1, x_2 \in \mathcal{C}$ und $\lambda \in [0, 1]$ gilt:*

$$f\Big(\lambda x_1 + (1-\lambda)x_2\Big) \leq \lambda f(x_1) + (1-\lambda) f(x_2).$$

Sie heißt strikt konvex, *wenn für $x_1 \neq x_2$ und $\lambda \in (0, 1)$ die Ungleichung strikt ist. Die Funktion f heißt* (strikt) konkav, *wenn $-f$ (strikt) konvex ist.*

Bemerkung 2.27 *Äquivalent dazu ist folgende Definition: $f : \mathcal{C} \to \mathbb{R}$ heißt konvex, wenn für Punkte $x_1, \ldots, x_p \in \mathcal{C}$ und $\lambda_1, \ldots, \lambda_p \geq 0$ mit $\sum_{i=1}^p \lambda_i = 1$ gilt:*

$$f\left(\sum_{i=1}^p \lambda_i x_i\right) \leq \sum_{i=1}^p \lambda_i f(x_i).$$

Übung: *Zeigen Sie, dass diese beiden Definitionen äquivalent sind.*

Anschaulich bedeutet Konvexität einer Funktion, dass für je zwei Punkte auf dem Funktionsgraphen die Verbindungsstrecke nirgends unterhalb der Funktion liegt (nirgends oberhalb bei einer konkaven Funktion).
Lineare Funktionen $c^T x + \gamma$ sind offensichtlich konvex und konkav.

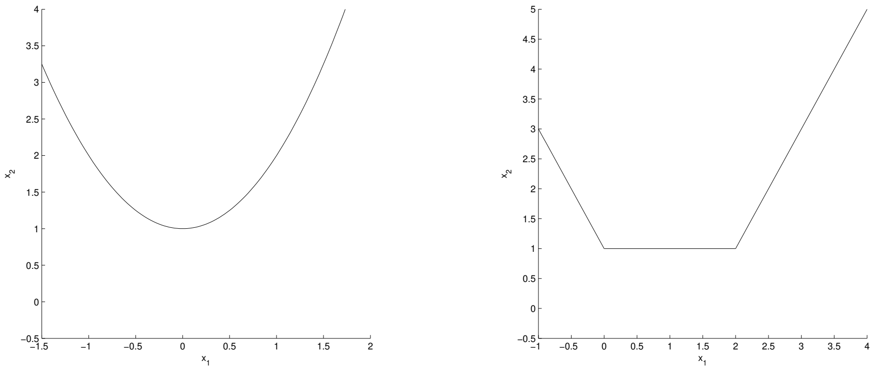

Abbildung 2.3: Links: strikt konvexe Funktion; Rechts: konvexe Funktion.

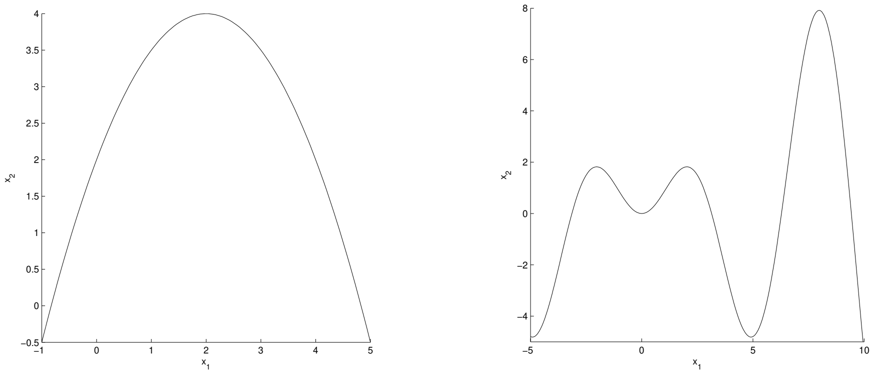

Abbildung 2.4: Links: konkave Funktion; Rechts: weder konvex noch konkav.

Es ist leicht, sich folgende Eigenschaften zu überlegen:

Satz 2.28 *Sei $\mathcal{C} \subset \mathbb{R}^n$ konvex, seien $f_1, f_2 : \mathcal{C} \to \mathbb{R}$ konvexe Funktionen und sei $\alpha > 0$. Dann sind auch αf_1, $f_1 + f_2$ und $\max[f_1, f_2]$ konvex auf $\mathcal{C}$.*

Beweis. Übung. □

Differenz, Produkt und Minimum konvexer Funktionen sind im Allgemeinen nicht konvex.
Zu jeder Funktion lassen sich zwei sie charakterisierende Mengen definieren. Ist die Funktion konvex, sind diese Mengen ihrerseits konvex:

Definition 2.29 *Sei $\mathcal{C} \subset \mathbb{R}^n$ und $f : \mathcal{C} \to \mathbb{R}$ eine Funktion. Dann heißt die Menge*

$$\mathcal{E}(f) = \{(x, \alpha) \in \mathcal{C} \times \mathbb{R} : f(x) \leq \alpha\}$$

der Epigraph von f. *Für $\beta \in \mathbb{R}$ heißt die Menge*

$$\mathcal{L}(f, \beta) = \{x \in \mathcal{C} : f(x) \leq \beta\}$$

(untere) Niveaumenge von f zum Niveau β.

Satz 2.30 *Sei $f : \mathbb{R}^n \to \mathbb{R}$. Dann gilt:*

(a) f ist konvex $\iff$ $\mathcal{E}(f)$ ist konvex.

(b) f ist konvex $\implies$ $\mathcal{L}(f, \beta)$ ist konvex für jedes $\beta \in \mathbb{R}$.
Die Umkehrung gilt nicht.

Beweis. Übung. □

Funktionen, deren Niveaumengen $\mathcal{L}(f, \beta)$ für jedes β konvex sind, heißen quasikonvex. Jede konvexe Funktion ist also quasikonvex, aber nicht umgekehrt. Konvexe Funktionen sind (bis auf Randpunkte) stetig. Sie sind jedoch nicht notwendigerweise differenzierbar, wie man leicht am Beispiel $f(x) = |x|$ sieht.

Satz 2.31 *Sei $\mathcal{C} \neq \emptyset$ eine konvexe Menge im $\mathbb{R}^n$, und sei $f : \mathcal{C} \to \mathbb{R}$ konvex. Dann ist f im Inneren von $\mathcal{C}$ stetig.*

Beweis. siehe Rockafellar [Ro70]. □

2.6 Differenzierbare konvexe Funktionen

Hier beschäftigen wir uns mit zwei charakteristischen Eigenschaften für differenzierbare konvexe Funktionen.

Satz 2.32 *Sei $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^1$.*
(a) f ist genau dann konvex über der konvexen Menge $\mathcal{C} \subseteq \mathbb{R}^n$, wenn für alle $x_1, x_2 \in \mathcal{C}$ gilt:

$$f(x_2) \geq f(x_1) + (x_2 - x_1)^T \nabla f(x_1). \tag{2.2}$$

(b) f ist strikt konvex $\Longleftrightarrow$ (2.2) gilt strikt für alle $x_1 \neq x_2 \in \mathcal{C}$.

Beweis. $\boxed{\text{(a, } \Longleftarrow\text{):}}$ Es gelte (2.2) für alle $x_1, x_2 \in \mathcal{C}$. Wählen wir zwei beliebige Punkte $x, y \in \mathcal{C}$ und $\lambda \in (0, 1)$. Wegen der Konvexität von $\mathcal{C}$ ist dann auch

$$z = \lambda x + (1 - \lambda) y \tag{2.3}$$

in $\mathcal{C}$. Wegen (2.2) gilt für $x, z \in \mathcal{C}$

$$f(x) \geq f(z) + (x - z)^T \nabla f(z), \tag{2.4}$$

und aus demselben Grund gilt

$$f(y) \geq f(z) + (y - z)^T \nabla f(z). \tag{2.5}$$

Multiplizieren wir nun (2.4) mit λ und (2.5) mit $(1 - \lambda)$, und addieren die beiden Ungleichungen, so erhalten wir

$$\lambda f(x) + (1 - \lambda) f(y) \geq f(z) + [\lambda(x - z) + (1 - \lambda)(y - z)]^T \nabla f(z).$$

Wegen (2.3) verschwindet die eckige Klammer, und die Konvexität von f ist gezeigt.
$\boxed{\text{(a, } \Longrightarrow\text{):}}$ Sei f konvex. Wir wählen $x, y \in \mathcal{C}$ und definieren die Hilfsfunktion $h : \mathbb{R} \to \mathbb{R}$ als

$$h(\lambda) = (1 - \lambda) f(x) + \lambda f(y) - f\Big((1 - \lambda) x + \lambda y\Big).$$

Wegen der Konvexität von f gilt für $x \neq y$ und $0 < \lambda < 1$, dass $h(\lambda) \geq 0$. Ausserdem ist $h(0) = 0$. Daher gilt für die Ableitung von h and der Stelle $\lambda = 0$:

$$\left.\frac{dh}{d\lambda}\right|_{\lambda=0} = -f(x) + f(y) - (y - x)^T \nabla f(x) \geq 0,$$

und damit gilt auch (2.2).
$\boxed{\text{(b):}}$ analog, nur jeweils $\geq$ durch $>$ ersetzen. $\square$

Ist eine Funktion zwei mal stetig differenzierbar, so lässt sich mit Hilfe der Hesse–Matrix feststellen, ob sie konvex ist oder nicht:

Satz 2.33 *Sei $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^2$. f ist genau dann konvex, wenn die Hesse–Matrix $\nabla^2 f(x)$ für alle $x \in \mathbb{R}^n$ positiv semidefinit ist.*

Beweis. $\boxed{(\Longleftarrow):}$ Die Hesse–Matrix $\nabla^2 f(x)$ sei überall positiv semidefinit. Nach dem Satz von Taylor gilt für alle $x, y \in \mathbb{R}^n$

$$f(y) = f(x) + (y-x)^T \nabla f(x) + \tfrac{1}{2}(y-x)^T \nabla^2 f\Big(x + t(y-x)\Big)(y-x) \quad (2.6)$$

wobei t eine reelle Zahl ist, $0 \leq t \leq 1$. Da $\nabla^2 f(x)$ überall positiv semidefinit ist, ist der letzte Summand in (2.6) nichtnegativ, daher

$$f(y) \geq f(x) + (y-x)^T \nabla f(x). \quad (2.7)$$

Aus Satz 2.32 folgt nun, dass f konvex ist.
$\boxed{(\Longrightarrow):}$ f sei konvex über dem $\mathbb{R}^n$. Angenommen, es existiert ein $x \in \mathbb{R}^n$, in dem die Hesse–Matrix nicht positiv semidefinit ist. Dann muss es ein $y \in \mathbb{R}^n$ geben, so dass

$$(y-x)^T \nabla^2 f(x)(y-x) < 0.$$

Wegen der Stetigkeit von $\nabla^2 f$ kann y so gewählt werden, dass für alle reellen t mit $0 \leq t \leq 1$

$$(y-x)^T \nabla^2 f\Big(x + t(y-x)\Big)(y-x) < 0$$

ist. Mit (2.6) folgt, dass für diese x, y (2.7) nicht gilt. Nach Satz 2.32 kann f also nicht konvex über $\mathbb{R}^n$ sein. $\square$

Zusatz: Man kann durch eine leichte Modifikation des ersten Teils des Beweises zeigen, dass zudem gilt

$$\nabla^2 f(x) \quad \text{pos. definit } \forall x \in \mathbb{R}^n \quad \Longrightarrow \quad f : \mathbb{R}^n \to \mathbb{R} \text{ strikt konvex.}$$

Achtung: Die Umkehrung gilt im Allgemeinen nicht, wie das Beispiel $f(x) = x^4$ zeigt.

Beispiel 2.34 *Betrachten wir eine quadratische Funktion* $f(x) = \alpha + c^T x + \frac{1}{2}x^T Qx$ *mit* $Q \in \mathbb{R}^{n\times n}$*symmetrisch,* $c \in \mathbb{R}^n$ *und* $\alpha \in \mathbb{R}$*. Ihre Hesse-Matrix ist*

$$\nabla^2 f(x) = Q.$$

Daher:

$$\begin{array}{rcl} f \text{ ist konvex} & \iff & Q \text{ ist positiv semidefinit.} \\ f \text{ ist konkav} & \iff & Q \text{ ist negativ semidefinit.} \end{array}$$

Zudem gelten bei quadratischen Funktionen auch die Äquivalenzen

$$\begin{array}{rcl} f \text{ ist strikt konvex} & \iff & Q \text{ ist positiv definit.} \\ f \text{ ist strikt konkav} & \iff & Q \text{ ist negativ definit.} \end{array}$$

Es gibt aber auch quadratische Funktionen von $\mathbb{R}^n$ *nach* $\mathbb{R}$ *(wenn* $n \geq 2$*), die weder konvex noch konkav sind, zum Beispiel die Funktion* $f(x_1, x_2) = x_1^2 + x_2^2 - 4x_1x_2$*. Deren Hesse-Matrix ist*

$$\nabla^2 f(x) = \begin{pmatrix} 2 & -4 \\ -4 & 2 \end{pmatrix},$$

die die Eigenwerte -2 *und* 6 *hat und daher indefinit ist.*

2.7 Optimalitätsresultate für konvexe Optimierungsprobleme

Satz 2.35 *Sei* $\mathcal{C} \subseteq \mathbb{R}^n$ *eine konvexe Menge und* $f : \mathcal{C} \to \mathbb{R}$ *eine konvexe Funktion. Dann ist jedes lokale Minimum von* f *über* $\mathcal{C}$ *bereits ein globales Minimum.*

Beweis. Sei $\bar{x} \in \mathcal{C}$ ein lokaler Minimalpunkt. Angenommen, es existiert ein Punkt $x^* \in \mathcal{C}$ mit $f(x^*) < f(\bar{x})$. Aus der Konvexität von f folgt, dass

$$f(\bar{x} + \lambda(x^* - \bar{x})) \leq \lambda f(x^*) + (1 - \lambda)f(\bar{x}) < f(\bar{x})$$

für alle $\lambda \in (0, 1)$. Dies widerspricht jedoch der lokalen Optimalität von $\bar{x}$, da ein $\bar{\lambda} > 0$ existieren muss, so dass $f(\bar{x}+\lambda(x^*-\bar{x})) \geq f(\bar{x})$ für $0 < \lambda < \bar{\lambda}$. □

Satz 2.36 *Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine konvexe Funktion, $\mathcal{C} \subseteq \mathbb{R}^n$ eine konvexe Menge. Dann ist* $\mathrm{Argmin}(f, \mathcal{C})$*, d.h. die Menge der Punkte, wo f ihr Minimum über $\mathcal{C}$ annimmt, konvex.*

Beweis. Übung. □

Korollar 2.37 *Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine strikt konvexe Funktion, $\mathcal{C} \subseteq \mathbb{R}^n$ eine konvexe Menge. Wenn das Minimum von f über $\mathcal{C}$ angenommen wird, dann in einem eindeutigen Punkt.*

Beweis. Angenommen, das Minimum wird an zwei verschiedenen Punkten $x_1, x_2 \in \mathcal{C}$ angenommen, und sei $\bar{\alpha} = f(x_1) = f(x_2)$. Aus Satz 2.36 folgt, dass $f(\lambda x_1 + (1-\lambda)x_2) = \bar{\alpha}$ für alle $\lambda \in (0,1)$. Das ist jedoch ein Widerspruch zur strikten Konvexität von f. □

Der nächste Satz gibt Auskunft darüber, wie das Minimum einer differenzierbaren konvexen Funktion über dem $\mathbb{R}^n$ (wenn also keine einschränkenden Nebenbedingungen erfüllt werden müssen) gefunden werden kann:

Satz 2.38 *Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine differenzierbare konvexe Funktion. Dann ist $\bar{x} \in \mathbb{R}^n$ ein globaler Minimalpunkt von f über $\mathbb{R}^n$ genau dann, wenn $\nabla f(\bar{x}) = 0$.*

Beweis. Ist $\bar{x} \in \mathbb{R}^n$ ein globaler Minimalpunkt von f dann ist bekanntlich $\nabla f(\bar{x}) = 0$ (auch wenn f nicht konvex ist!).

(Zur Erinnerung: In einem lokalen Minimum $\bar{x}$ gilt für alle $v \in \mathbb{R}^n$

$$0 \leq \lim_{t \searrow 0} \frac{f(\bar{x} + tv) - f(\bar{x})}{t} = \nabla f(\bar{x})^T v,$$

also $\nabla f(\bar{x})^T v \geq 0$ für alle v und somit $\nabla f(\bar{x}) = 0$.)
Sei nun umgekehrt $\nabla f(\bar{x}) = 0$ und zudem f konvex. Aus Satz 2.32 folgt, dass

$$f(x) - f(\bar{x}) \geq (x - \bar{x})^T \nabla f(\bar{x}) = 0 \quad \forall\, x \in \mathbb{R}^n,$$

und somit

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

Daher ist $\bar{x}$ ein globaler Minimalpunkt von f über $\mathbb{R}^n$. □

Entsprechende Resultate gelten nicht für das Maximum einer konvexen Funktion, wie man an Abbildung 2.5 sieht.

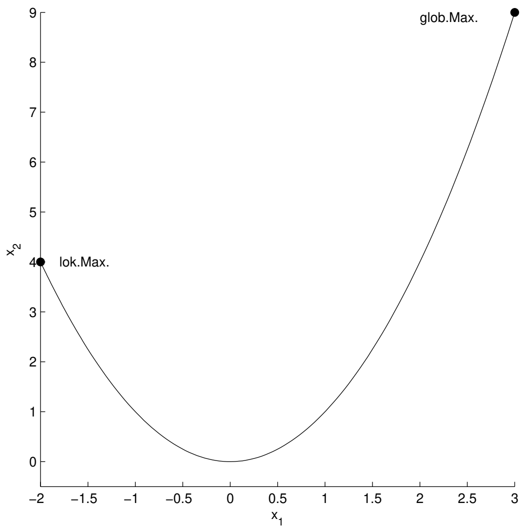

Abbildung 2.5: Lokales und globales Maximum einer konvexen Funktion.

In diesem Beispiel wird das Maximum in einem Randpunkt, genauer: in einem Extrempunkt angenommen. Dies gilt immer, wie der nächste Satz zeigt.

Satz 2.39 *Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine konvexe Funktion, $\mathcal{C} \subseteq \mathbb{R}^n$ eine kompakte konvexe Menge. Dann nimmt f ihr Maximum über $\mathcal{C}$ in einem Extrempunkt an.*

Beweis. Sei x ein beliebiger Punkt aus $\mathcal{C}$ und $\mathcal{E}$ die Menge der Extrempunkte von $\mathcal{C}$. Nach Satz 2.13 kann x als Konvexkombination von Punkten aus $\mathcal{E}$ dargestellt werden, nach Satz 2.9 sind dazu höchstens $(n+1)$ solche Punkte notwendig. Es existieren also $v_0, v_1, \ldots, v_n \in \mathcal{E}$ und $\lambda_0, \lambda_1, \ldots, \lambda_n \geq 0$ mit $\sum_{i=0}^n \lambda_i = 1$, so dass

$$x = \sum_{i=0}^{n} \lambda_i v_i.$$

Aus der Konvexität von f folgt die Abschätzung

$$\begin{aligned} f(x) = f\left(\sum_{i=0}^{n} \lambda_i v_i\right) &\leq \sum_{i=0}^{n} \lambda_i f(v_i) \\ &\leq \sum_{i=0}^{n} \lambda_i \max\{f(v_i) : i = 0, \ldots, n\} \\ &= \max\{f(v_i) : i = 0, \ldots, n\} \sum_{i=0}^{n} \lambda_i \\ &= \max\{f(v_i) : i = 0, \ldots, n\}. \end{aligned}$$

Zu einem beliebigen Punkt gibt es also immer einen Extrempunkt mit größerem Funktionswert. Das Maximum von f über $\mathcal{C}$ wird also in einem Extrempunkt angenommen. □

Korollar 2.40 *Eine lineare Funktion nimmt sowohl ihr Maximum als auch ihr Minimum über einer kompakten konvexen Menge in einem Extrempunkt an.*

Beweis. Da eine lineare Funktion konvex ist, folgt die Aussage über das Maximum direkt aus dem vorigen Satz 2.39.
Für das Minimierungsproblem verwenden wir die Beziehung

$$\min\{f(x) : x \in \mathcal{C}\ \} = -\max\{-f(x) : x \in \mathcal{C}\}.$$

Da mit f auch $-f$ linear und damit konvex ist, folgt auch diese Aussage aus Satz 2.39. □

Satz 2.41 *Sei $\mathcal{M} \subset \mathbb{R}^n$ kompakt, und sei $f : \mathbb{R}^n \to \mathbb{R}$ konvex. Dann gilt:*

$$\max\{f(x) : x \in \mathcal{M}\} = \max\{f(x) : x \in \operatorname{conv} \mathcal{M}\}.$$

Beweis. Übung. □

Kapitel 3

Polytope und Polyeder

Jede konvexe Menge ist also Durchschnitt von Halbräumen. Es ist klar, dass man für diese Darstellung im Allgemeinen unendlich viele Halbräume benötigt. Solche Mengen, die durch endlich viele Halbräume dargestellt werden können, verdienen deshalb spezielle Betrachtung:

Definition 3.1 *Eine Menge $\mathcal{P} \subseteq \mathbb{R}^n$ heißt* Polyeder*, wenn sie der Durchschnitt endlich vieler abgeschlossener Halbräume ist. Ein Polyeder heißt* Polytop*, wenn es beschränkt ist.*

Bemerkung 3.2 *Jedes Polyeder ist konvex (weil Durchschnitt endlich vieler Halbräume, die ja konvexe Mengen sind).*

Bemerkung 3.3 *Jedes Polyeder lässt sich in der Form*

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

schreiben mit $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.

Beispiel 3.4 *Betrachten wir das Polyeder, das von folgenden fünf Halbräumen erzeugt wird:*

$$\begin{array}{lrcrcr}
(1) & 2x_1 & & & \leq & 5 \\
(2) & & - & 2x_2 & \leq & 1 \\
(3) & -x_1 & - & x_2 & \leq & -1 \\
(4) & 2x_1 & + & 9x_2 & \leq & 23 \\
(5) & 6x_1 & - & 2x_2 & \leq & 13
\end{array}$$

Wie man sieht (Abb. 3.1), ist diese Menge beschränkt; das Polyeder ist also ein Polytop. Es hat die Ecken: $(-2, 3)$; $(1.5, -0.5)$; $(2, -0.5)$; $(2.5, 1)$ und $(2.5, 2)$.

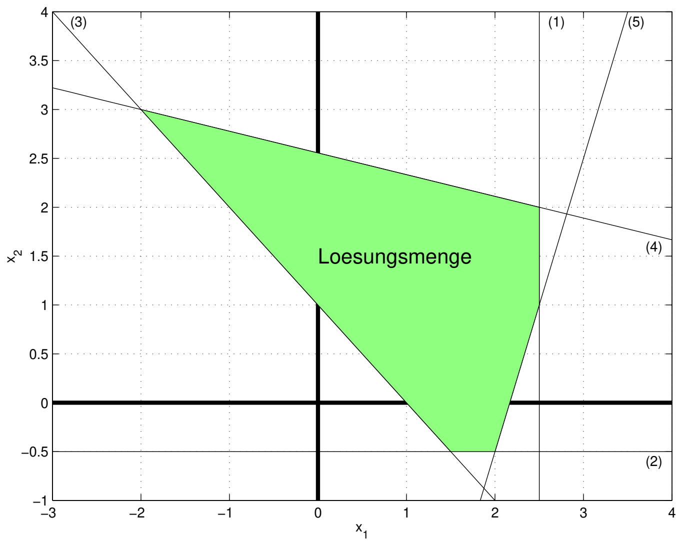

Abbildung 3.1: Das Polyeder aus Beispiel 3.4.

Das obige System (1)–(5) lässt sich auch in Matrixschreibweise als

$$Ax \leq b$$

schreiben, wobei

$$A = \begin{pmatrix} 2 & 0 \\ 0 & -2 \\ -1 & -1 \\ 2 & 9 \\ 6 & -2 \end{pmatrix}, \qquad b = \begin{pmatrix} 5 \\ 1 \\ -1 \\ 23 \\ 13 \end{pmatrix}.$$

$Ax \leq b$ wird auch definierendes System des Polyeders genannt. Dabei sind A und b nicht eindeutig bestimmt.

Die zulässigen Mengen linearer Probleme treten nicht immer in der Form $Ax \leq b$ auf. Häufig gibt es Gleichungen oder Vorzeichenbeschränkungen auf Variablen.

Allgemein gilt:

Bemerkung 3.5 *Die Lösungsmenge des Systems*

$$\begin{array}{rcrcl} Bx & + & Cy & = & c \\ Dx & + & Ey & \leq & d \\ x & & & \geq & 0 \\ x & & & \in & \mathbb{R}^p \\ & & y & \in & \mathbb{R}^q \end{array} \tag{3.1}$$

ist ein Polyeder.

Beweis. Setze $n = p + q$ und

$$A = \begin{pmatrix} B & C \\ -B & -C \\ D & E \\ -I & 0 \end{pmatrix} \qquad \text{und} \qquad b = \begin{pmatrix} c \\ -c \\ d \\ 0 \end{pmatrix}.$$

Dann ist $\mathcal{P}(A, b)$ die Lösungsmenge des Systems (3.1). □

Ein spezieller Polyedertyp wird uns häufig begegnen, insbesondere bei der Entwicklung von Lösungsverfahren für lineare Probleme. Hierzu führen wir die folgende Bezeichnung ein:

$$\mathcal{P}^{=}(A, b) = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}.$$

Beachte nochmals, dass Polyeder unterschiedliche Darstellungsformen haben können. Wir werden im Rahmen dieser Vorlesung Sätze über Polyeder beweisen, die darstellungsabhängig sind bzw. sich in der einen oder anderen Darstellung leichter beweisen lassen bzw. einsichtiger sind. Wir werden die Sätze dann nur für eine Darstellungsform beweisen und geben nachfolgend Transformationen an, wie man von einer Darstellung in eine andere kommt.

Bemerkung 3.6 Transformationen.

Regel I: *Einführung von Schlupfvariablen.*

Gegeben seien $a \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$. *Wir schreiben*

$$a^T x \leq \alpha \qquad \text{als} \qquad a^T x + y = \alpha, \;\; y \geq 0.$$

$y \in \mathbb{R}$ *heißt* Schlupfvariable.

Regel II: *Einführung von vorzeichenbeschränkten Variablen.*

Eine nicht vorzeichenbeschränkte Variable $x \in \mathbb{R}$ können wir auch schreiben als $x = x^+ - x^-$ mit $x^+, x^- \geq 0$.

Damit kann ein Polyeder $\mathcal{P}(A, b)$ jederzeit in ein Polyeder $\mathcal{P}^=(D, d)$ überführt werden. Wir bezeichnen zwei Polyeder, die mit obigen Regeln ineinander transformiert werden können, als äquivalent.

Bevor wir uns der Dimension und den Seitenflächen von Polyedern widmen, betrachten wir noch einen speziellen Polyedertyp, der uns häufig begegnen wird.

Definition 3.7 *Eine Menge $\mathcal{K} \subseteq \mathbb{R}^n$ heißt* Kegel*, wenn für jedes $x \in \mathcal{K}$ und $\alpha \geq 0$ gilt: $\alpha x \in \mathcal{K}$. Ein Kegel $\mathcal{K} \subseteq \mathbb{R}^n$ heißt* polyedrisch*, wenn $\mathcal{K}$ ein Polyeder ist.*

Bemerkung 3.8 *Ein nichtleerer Kegel $\mathcal{K} \subseteq \mathbb{R}^n$ ist genau dann polyedrisch, wenn es eine Matrix A gibt mit*

$$\mathcal{K} = \mathcal{P}(A, 0).$$

Beweis.
$\boxed{(\Longleftarrow):}$ Ist $\mathcal{K} = \mathcal{P}(A, 0)$, so ist $\mathcal{K}$ ein nichtleeres Polyeder und offensichtlich auch ein Kegel.
$\boxed{(\Longrightarrow):}$ Hier verwenden wir die Bezeichnung $A_{i.}$ für die i-te Zeile von A.
Sei $\mathcal{K}$ ein nichtleerer polyedrischer Kegel. Dann existiert eine Matrix A und ein Vektor b mit $\mathcal{K} = \mathcal{P}(A, b)$. Da $0 \in \mathcal{K}$ gilt, folgt $b \geq 0$.
Angenommen, es existiert ein $\bar{x} \in \mathcal{K}$ mit $A\bar{x} \not\leq 0$, d.h. es existiert eine Zeile $A_{i.}$ von A mit $t = A_{i.}\bar{x} > 0$. Da $\mathcal{K}$ ein Kegel ist, gilt $\lambda\bar{x} \in \mathcal{K}$ $\forall \lambda \geq 0$. Also gilt $A_{i.}\lambda\bar{x} = \lambda t \leq b_i$ für alle $\lambda \geq 0$, was wegen $t > 0$ auf einen Widerspruch führt. Also erfüllen alle $x \in \mathcal{K}$ auch $Ax \leq 0$. Wegen $b \geq 0$ folgt daraus nun $\mathcal{K} = \mathcal{P}(A, 0)$. □

3.1 Seitenflächen von Polyedern

Definition 3.9 *Es seien $\mathcal{M} \subseteq \mathbb{R}^n, a \in \mathbb{R}^n, \alpha \in \mathbb{R}$. Die Ungleichung $a^T x \leq \alpha$ heißt* gültig *bezüglich $\mathcal{M}$, falls*

$$\mathcal{M} \subseteq \{x \in \mathbb{R}^n : a^T x \leq \alpha\}.$$

Man beachte, dass $a = 0$ in Definition 3.9 zugelassen ist.

Definition 3.10 *Sei $\mathcal{P} \subseteq \mathbb{R}^n$ ein Polyeder. Eine Menge $\mathcal{F} \subseteq \mathcal{P}$ heißt* Seitenfläche *von $\mathcal{P}$, wenn es eine bezüglich $\mathcal{P}$ gültige Ungleichung $d^T x \leq \delta$ gibt mit*

$$\mathcal{F} = \mathcal{P} \cap \{x \in \mathbb{R}^n : d^T x = \delta\}.$$

Eine Seitenfläche heißt echt*, wenn $\mathcal{F} \neq \mathcal{P}$ gilt. $\mathcal{F}$ heißt* nichttrivial*, wenn $\emptyset \neq \mathcal{F} \neq \mathcal{P}$ gilt. Ist $d^T x \leq \delta$ gültig für $\mathcal{P}$, dann heißt $\mathcal{P} \cap \{x \in \mathbb{R}^n : d^T x = \delta\}$ die von $d^T x \leq \delta$* induzierte Seitenfläche.

Beachte wiederum, dass $d = 0$ in Definition 3.10 zugelassen ist.

Beispiel 3.11 *Sei $\mathcal{P}(A, b)$ gegeben durch*

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \textit{und} \qquad b = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

vgl. Abbildung 3.2.

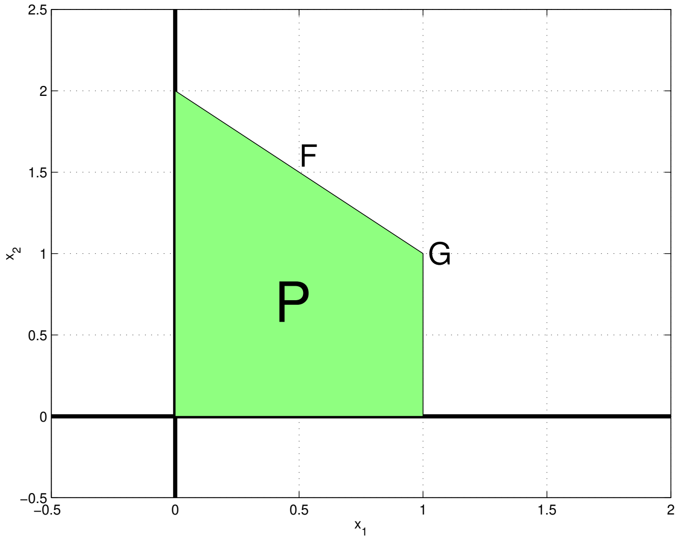

Abbildung 3.2: Grafische Darstellung zu Beispiel 3.11

Das Geradenstück zwischen den Punkten $\binom{1}{1}$ *und* $\binom{0}{2}$ *definiert eine Seitenfläche* $\mathcal{F}$ *von* $\mathcal{P}$*, denn* $\mathcal{F} = \mathcal{P} \cap \{x \in \mathbb{R}^2 : x_1 + x_2 = 2\}$ *ist durch die Ungleichung* $x_1 + x_2 \leq 2$ *induziert. Ebenso ist die Menge bestehend aus dem Punkt* $\mathcal{G} = (1,1)^T$ *eine Seitenfläche von* $\mathcal{P}$*, denn*

$$\mathcal{G} = \mathcal{P} \cap \{x : 2x_1 + x_2 = 3\} \qquad \textit{und} \qquad 2x_1 + x_2 \leq 3 \textit{ ist gültige Ungleichung,}$$

oder auch:

$$\mathcal{G} = \mathcal{P} \cap \{x : 3x_1 + x_2 = 4\} \qquad \textit{und} \qquad 3x_1 + x_2 \leq 4 \textit{ ist gültige Ungleichung.}$$

Daraus folgt, dass eine Seitenfläche durch völlig unterschiedliche Ungleichungen induziert werden kann.

Bemerkung 3.12 *Sei* $\mathcal{P} \subseteq \mathbb{R}^n$ *ein Polyeder. Dann gilt:*

(a) $\mathcal{P}$ *ist eine Seitenfläche von sich selbst.*

(b) $\emptyset$ *ist eine Seitenfläche von* $\mathcal{P}$*.*

(c) Ist $\mathcal{F} = \{x : d^T x = \delta\} \cap \mathcal{P}$ *eine nichttriviale Seitenfläche von* $\mathcal{P}$*, dann gilt* $d \neq 0$*.*

Beweis.

(a) $\mathcal{P} = \mathcal{P} \cap \{x \in \mathbb{R}^n : 0^T x = 0\}$,

(b) Sei $\delta > 0$ beliebig. $0^T x \leq \delta$ ist eine gültige Ungleichung für $\mathcal{P}$ und es gilt $\emptyset = \mathcal{P} \cap \{x \in \mathbb{R}^n : 0^T x = \delta\}$.

(c) Ist klar, da für $d = 0$ nur der Fall $\delta \geq 0$ eine gültige Ungleichung $d^T x \leq \delta$ liefert und somit einer der beiden ersten Fälle eintritt.

□

Satz 3.13 *Sei* $\mathcal{P} = \mathcal{P}(A,b) \neq \emptyset$ *ein Polyeder und* $c \in \mathbb{R}^n$*. Betrachte das lineare Optimierungsproblem*

$$\max\{c^T x : x \in \mathcal{P}\}.$$

Sei $\mathcal{O}$ *die Lösungsmenge und im Fall* $\mathcal{O} \neq \emptyset$ *bezeichne* $\bar{z} = \max\{c^T x : x \in \mathcal{P}\}$ *den Optimalwert. Dann gilt:*

(a) Ist $\mathcal{O} \neq \emptyset$, dann ist $\mathcal{O} = \{x \in \mathcal{P} : c^T x = \bar{z}\}$ eine nichtleere Seitenfläche von $\mathcal{P}$ und $\{x \in \mathbb{R}^n : c^T x = \bar{z}\}$ ist eine Stützhyperebene von $\mathcal{P}$, falls $c \neq 0$.

(b) Die Menge der Optimallösungen des linearen Problems $\max\{c^T x : x \in \mathcal{P}\}$ ist eine Seitenfläche des durch die Nebenbedingungen definierten Polyeders.

Beweis.

$\boxed{\text{(a):}}$ Sei $\mathcal{O} \neq \emptyset$. Nach Definition von $\bar{z}$ ist $c^T x \leq \bar{z}$ für alle $x \in \mathcal{P}$ $c^T x \leq \bar{z}$ ist also gültig für $\mathcal{P}$. Ferner gilt $\mathcal{O} = \{x \in \mathcal{P} : c^T x = \bar{z}\}$, somit ist $\mathcal{O} \neq \emptyset$ nach Definition eine nichtleere Seitenfläche von $\mathcal{P}$. Zudem folgt unmittelbar, dass die Menge $\{x : c^T x = \bar{z}\}$ eine Stützhyperebene ist, falls $c \neq 0$.

$\boxed{\text{(b):}}$ Im Fall $\mathcal{O} = \emptyset$ ist dies trivial und andernfalls ist (b) nur eine Umformulierung von (a).

□

Notation: Zu einer Matrix $A \in \mathbb{R}^{m,n}$ mit Zeilenindexmenge M bezeichne $A_{i\cdot}$, $i \in M$, Zeile i von A. Für $I \subset M$ bezeichne $A_{I\cdot}$ die aus den Zeilen $i \in I$ gebildete Teilmatrix.

Definition 3.14 *Sei $\mathcal{P} = \mathcal{P}(A, b) \subseteq \mathbb{R}^n$ ein Polyeder und M die Zeilenindexmenge von A. Für $\mathcal{F} \subseteq \mathcal{P}$ sei*

$$eq(\mathcal{F}) = \{i \in M : A_{i\cdot}x = b_i \ \forall x \in \mathcal{F}\},$$

d.h. $eq(\mathcal{F})$ ist die Menge der für alle $x \in \mathcal{F}$ bindenden Restriktionen *(engl.: equality set). Für $I \subseteq M$ bezeichne*

$$\mathrm{fa}(I) = \{x \in \mathcal{P} : A_{I\cdot}x = b_I\}.$$

die von I induzierte Seitenfläche *(engl. induced face).*

$\mathrm{fa}(I)$ ist tatsächlich eine Seitenfläche von $\mathcal{P} = \mathcal{P}(A, b)$, denn mit

$$c^T = \sum_{i \in I} A_{i\cdot} \qquad \text{und} \qquad \gamma = \sum_{i \in I} b_i$$

ist $c^T x \leq \gamma$ gültig und $\text{fa}(I) = \{x \in \mathcal{P} : c^T x = \gamma\}$.

Betrachte nochmals Beispiel 3.11, so gilt mit $M = \{1, 2, 3, 4\}$

$$\text{fa}(\{1,2\}) = \mathcal{G} \qquad \text{und} \qquad \text{eq}\left(\left\{\begin{pmatrix}0\\2\end{pmatrix}\right\}\right) = \{1,3\}.$$

Definition 3.15 *Sei $\mathcal{P} \neq \emptyset$ ein Polyeder im $\mathbb{R}^n$. Die* Dimension $\dim(\mathcal{P})$ von *$\mathcal{P}$ ist definiert durch*

$$\dim(\mathcal{P}) = \max\{d \in \mathbb{N}_0 \;:\; \exists x_0, \ldots, x_d \in \mathcal{P} \text{ mit } x_1 - x_0, x_2 - x_0, \ldots, x_d - x_0 \text{ linear unabhängig}\},$$

mit anderen Worten

$$\dim(\mathcal{P}) = \dim(\text{aff}\,(\mathcal{P})).$$

Hierbei ist aff $(\mathcal{P})$ *die affine Hülle der Menge $\mathcal{P}$, also der kleinste affine Raum, der $\mathcal{P}$ enthält.*
$\mathcal{P}$ heißt volldimensional, *wenn $\dim(\mathcal{P}) = n$ gilt.*

3.2 Ecken, Facetten, Redundanz

Definition 3.16 *Im Folgenden sei $Ax \leq b$ ein Ungleichungssystem und M sei die Zeilenindexmenge von A.*

(a) *Sei $I \subseteq M$, dann heißt das System $A_{I.}\, x \leq b_I$* unwesentlich *oder* redundant *bezüglich $Ax \leq b$, falls gilt*

$$\mathcal{P}(A, b) = \mathcal{P}(A_{M\setminus I.}, b_{M\setminus I}).$$

(b) *Enthält $Ax \leq b$ ein unwesentliches Teilsystem $A_{I.}\, x \leq b_I$, dann heißt $Ax \leq b$* redundant, *andernfalls* irredundant.

(c) *Eine Ungleichung $A_{i.}\, x \leq b_i$ heißt* wesentlich *oder* nicht redundant *bezüglich $Ax \leq b$, falls gilt*

$$\mathcal{P}(A, b) \neq \mathcal{P}(A_{M\setminus\{i\}.}, b_{M\setminus\{i\}}).$$

(d) *Eine Ungleichung $A_{i.}\, x \leq b_i$ heißt* implizite Gleichung *bezüglich $Ax \leq b$, falls $i \in eq(\mathcal{P}(A, b))$.*

(e) *Ein System* $Ax \leq a$, $Bx = b$ *heißt* irredundant, *wenn* $Ax \leq a$ *keine unwesentliche Ungleichung bezüglich des Systems*

$$Ax \leq a, \quad Bx \leq b, \quad -Bx \leq -b$$

enthält und B *vollen Zeilenrang hat.*

(f) *Eine nichttriviale Seitenfläche* F *von* $\mathcal{P}(A, b)$ *heißt* Facette *von* $\mathcal{P}(A, b)$, *falls* F *in keiner anderen echten Seitenfläche von* $\mathcal{P}(A, b)$ *enthalten ist.*

Beispiel 3.17 *Betrachte* $\mathcal{P}(A, b)$ *mit*

$$A = \begin{pmatrix} 0 & 1 \\ 0 & -1 \\ -1 & 0 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{und} \qquad b = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 2 \\ 1 \end{pmatrix}.$$

Es gilt:

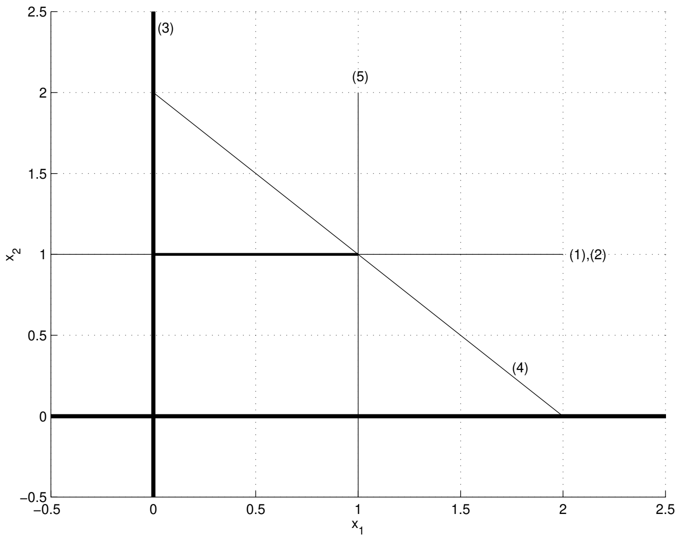

Abbildung 3.3: Grafische Darstellung zu Beispiel 3.17.

- *Die zu* $\{4\}$ *gehörende Ungleichung* $A_{\{4\}.}\,x \leq b_{\{4\}}$ *ist unwesentlich. Wir schreiben in diesem Beispiel für diesen Fall kurz:* $\{(4)\}$ *ist unwesentlich.*

- $\{(5)\}$ *ist unwesentlich, aber nicht* $\{(4),(5)\}$.
- $\{(1)\}$, $\{(2)\}$, *und* $\{(3)\}$ *sind wesentlich.*
- $\mathcal{F} = \left\{\binom{0}{1}\right\}$ *ist eine Facette.*
- *Das System* $x_2 = 1$, $\begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} x \leq \binom{0}{1}$ *ist irredundant.*

Definition 3.18 *Sei* $\mathcal{P} \subseteq \mathbb{R}^n$ *ein Polyeder. Eine nulldimensionale (also einpunktige) Seitenfläche von* $\mathcal{P}$ *heißt* Ecke.

Wir zeigen nun unter anderem, dass Ecken genau die Extrempunkte von Polyedern sind.

Satz 3.19 *Sei* $\mathcal{P} = \mathcal{P}(A,b) = \{x \in \mathbb{R}^n : Ax \leq b\} \subseteq \mathbb{R}^n$ *ein Polyeder und* $x \in \mathcal{P}$. *Dann sind die folgenden Aussagen äquivalent:*

(1) x *ist eine Ecke von* $\mathcal{P}$.

(2) $\{x\}$ *ist eine nulldimensionale Seitenfläche von* $\mathcal{P}$.

(3) x *ist keine echte Konvexkombination von Elementen von* $\mathcal{P}$, *d.h. für* $y, z \in \mathcal{P}$, $y \neq z, 0 < \lambda < 1$ *gilt* $x \neq \lambda y + (1-\lambda) z$.

(4) $\operatorname{rang}(A_{eq(\{x\})\cdot}) = n$.

(5) Es existiert ein $c \in \mathbb{R}^n \setminus \{0\}$, *so dass* x *die eindeutig bestimmte Optimallösung des linearen Problems* $\max\{c^T y : y \in \mathcal{P}\}$ *ist.*

Beweis. Die Äquivalenz der Aussagen (1) und (2) ist gerade die Definition einer Ecke. Der Ringbeweis besteht aus den Teilen (1)⇒(5), (5)⇒(3), (3)⇒(4) und (4)⇒(2).

$\boxed{(1)\Longrightarrow(5):}$ Nach Definition ist x eine Seitenfläche, also existiert eine bezüglich $\mathcal{P}$ gültige Ungleichung $c^T x \leq \gamma$, so dass $\{y \in \mathcal{P} : c^T y = \gamma\} = \{x\}$. Folglich ist x die eindeutig bestimmte Optimallösung von $\max\{c^T y : y \in \mathcal{P}\}$. Ist $\mathcal{P} \neq \{x\}$, so ist $c \neq 0$, andernfalls kann $c \neq 0$ gewählt werden.

$\boxed{(5)\Longrightarrow(3):}$ Sei x die eindeutige Optimallösung von $\max\{c^T u : u \in \mathcal{P}\}$ mit Wert γ. Gilt $x = \lambda y + (1-\lambda)z$ für $y, z \in \mathcal{P}$, $y \neq z$, $0 < \lambda < 1$, dann folgt

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

Daraus folgt nun $c^T y = \gamma = c^T z$ und dies ist ein Widerspruch zur Eindeutigkeit von x.

$\boxed{(3)\Longrightarrow(4):}$ Gilt (4) nicht, dann existiert $v \neq 0$ mit $A_{\mathrm{eq}(\{x\})\cdot} v = 0$. Für kleines $t > 0$ gilt dann

$$A(x \pm tv) \leq b,$$

denn die bindenden Nebenbedingungen bleiben aktiv, die inaktiven Nebenbedingungen bleiben für $t > 0$ klein inaktiv. Mit $y = x - tv$, $z = x + tv$ gilt dann $y, z \in \mathcal{P}$ und $x = \frac{1}{2}y + \frac{1}{2}z$. Also gilt auch (3) nicht.

$\boxed{(4)\Longrightarrow(2):}$ Wir wissen, dass $\mathrm{fa}(\mathrm{eq}(\{x\}))$ eine Seitenfläche ist und wegen $\mathrm{rang}\,(A_{\mathrm{eq}(\{x\})\cdot}) = n$ gilt

$$\mathrm{fa}(\mathrm{eq}(\{x\})) = \{y \in \mathcal{P} \ : \ A_{\mathrm{eq}(\{x\})\cdot} y = b_{\mathrm{eq}(\{x\})}\} = \{x\}.$$

Also ist $\{x\}$ eine Seitenfläche und ihre Dimension ist offensichtlich 0.

$\square$

Für Polyeder der Form $\mathcal{P}^=(A, b)$ gibt es ebenfalls eine Charakterisierung der Ecken. Dazu sei für $x \in \mathbb{R}^n$

$$\mathrm{supp}\,(x) = \{i \in \{1, 2, \ldots, n\} : x_i \neq 0\}$$

der Träger (engl.: support) von x.

Satz 3.20 *Für $x \in \mathcal{P}^=(A, b) = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\} \subseteq \mathbb{R}^n$ sind folgende Aussagen äquivalent:*

(1) x ist eine Ecke von $\mathcal{P}^=(A, b)$.

(2) $\operatorname{rang}(A_{\cdot \operatorname{supp}(x)}) = \left|\operatorname{supp}(x)\right|$.

(3) Die Spaltenvektoren $A_{\cdot j}$, $j \in \operatorname{supp}(x)$, *sind linear unabhängig.*

Beweis. Übung. $\square$

Definition 3.21 *Sei* $\mathcal{P} \subseteq \mathbb{R}^n$ *ein Polyeder. Eine eindimensionale Seitenfläche von* $\mathcal{P}$ *heißt* Kante. *Ist* $\mathcal{F}$ *eine Kante und gibt es ein* $x \in \mathbb{R}^n$ *und* $0 \neq z \in \mathbb{R}^n$ *mit*

$$\left\{ \begin{array}{l} \mathcal{F} = \{x\} + \operatorname{lin}(\{z\}) \\ \mathcal{F} = \{x\} + \operatorname{cone}(\{z\}) \end{array} \right\}, \quad \textit{so heißt } \mathcal{F} \qquad \left\{ \begin{array}{l} \textit{Extremallinie} \\ \textit{Extremalstrahl} \end{array} \right\}.$$

Dabei ist $\operatorname{lin}(\{z\})$ *die lineare Hülle von* $\{z\}$ *und* $\operatorname{cone}(\{z\}) := \{\alpha z : \alpha \geq 0\}$ *die konische Hülle von* $\{z\}$.

Kanten sind also entweder Extremallinien, Extremalstrahlen oder Verbindungsstrecken zweier Ecken. Sind zwei Ecken $x, y \in \mathcal{P}$ durch eine Kante verbunden, so nennt man x und y adjazent auf $\mathcal{P}$.

3.3 Dimensionsformel und Darstellung von Seitenflächen

Um Rechenformeln für die Dimension von Polyedern beweisen zu können, ist der Begriff von inneren Punkten nützlich.

Definition 3.22 *Ein Element* x *eines Polyeders* $\mathcal{P}$ *heißt* innerer Punkt *von* $\mathcal{P}$, *falls* x *in keiner echten Seitenfläche von* $\mathcal{P}$ *enthalten ist.*

Beachte: Innere Punkte von Polyedern sind nicht notwendigerweise innere Punkte im Sinne der natürlichen Topologie. Sie sind lediglich innere Punkte im Sinne der Relativtopologie auf $\mathcal{P}$.

Satz 3.23 *Zu jedem nichtleeren Polyeder* $\mathcal{P} = \mathcal{P}(A, b)$ *existiert ein Punkt* $\bar{x} \in \mathcal{P}$ *mit*

$$A_{eq(\mathcal{P})\cdot}\bar{x} = b_{eq(\mathcal{P})}, \quad A_{M\setminus eq(\mathcal{P})\cdot}\bar{x} < b_{M\setminus eq(\mathcal{P})},$$

wobei M die Zeilenindexmenge von A bezeichne (im Fall $M \setminus eq(\mathcal{P}) = \emptyset$ entfällt die Ungleichung). Jeder solche Punkt ist innerer Punkt von $\mathcal{P}$. Insbesondere besitzt jedes nichtleere Polyeder innere Punkte und es gilt

$$\dim(\mathcal{P}) = n - \operatorname{rang}(A_{eq(\mathcal{P})\cdot}) = \dim(\operatorname{Kern}(A_{eq(\mathcal{P})\cdot})).$$

Beweis. Sei $I = \operatorname{eq}(\mathcal{P})$ und $J = M \setminus I$. Wir zeigen zunächst die Existenz von $\bar{x} \in \mathcal{P}$ mit

$$A_{I\cdot}\bar{x} = b_I, \quad A_{J\cdot}\bar{x} < b_J.$$

Gilt $\boxed{I = M,}$ dann ist

$$\mathcal{P} = \{x \ : \ Ax = b\} = \{x \ : \ A_{I\cdot}x = b_I\}, \quad J = \emptyset.$$

Wegen $\mathcal{P} \neq \emptyset$ existiert also ein $\bar{x} \in \mathcal{P}$ mit $A_I\bar{x} = b_I$, es ist sogar jedes $\bar{x} \in \mathcal{P}$ geeignet.
Gilt $\boxed{I \neq M,}$ dann ist das System $Ax \leq b$ äquivalent zu

$$\begin{array}{rcl} A_{I\cdot}\,x & = & b_I \\ A_{J\cdot}\,x & \leq & b_J. \end{array}$$

Zu jedem $j \in J$ existiert wegen $j \notin \operatorname{eq}(\mathcal{P})$ ein $x_j \in \mathcal{P}$ mit $A_{j\cdot}\,x_j < b_j$. Da $\mathcal{P}$ konvex ist, liegt die Konvexkombination

$$\bar{x} = \frac{1}{|J|}\sum_{i \in J} x_i$$

wieder in $\mathcal{P}$, erfüllt also insbesondere $A_{I\cdot}\bar{x} = b_I$. Zudem gilt für alle $j \in J$

$$A_{j\cdot}\bar{x} = A_{j\cdot}\left(\frac{1}{|J|}\sum_{i \in J} x_i\right) = \frac{1}{|J|}\sum_{i \in J} A_{j\cdot}x_i < \frac{1}{|J|}\sum_{j \in J} b_j = b_j.$$

Also ist $\bar{x} \in \mathcal{P}$ mit $A_{I\cdot}\,\bar{x} = b_I$ und $A_{J\cdot}\,\bar{x} < b_J$.

Jedes solche $\bar{x}$ ist ein innerer Punkt: Sei $\mathcal{F} = \{x \in \mathcal{P} \ : \ d^T x = \delta\}$ eine beliebige Seitenfläche mit $\bar{x} \in \mathcal{F}$. Wir müssen zeigen, dass $\mathcal{F} = \mathcal{P}$ gilt. Wegen

$$A_{I\cdot}\,\bar{x} = b_I, \quad A_{J\cdot}\,\bar{x} < b_J$$

(die Ungleichung entfällt für $J = \emptyset$) finden wir $\epsilon > 0$ mit

$$A_{I\cdot}\,(\bar{x} + y) = b_I, \quad A_{J\cdot}\,(\bar{x} + y) < b_J \quad \forall\, y \in \operatorname{Kern}(A_{I\cdot}), \quad \|y\| \leq \epsilon.$$

Damit gilt

$$\{\bar{x}+y \ : \ y \in \operatorname{Kern}(A_{I\cdot}), \quad \|y\| \leq \epsilon\} \subset \mathcal{P} \subset \{x \ : \ d^T x \leq \delta\}$$

und folglich

$$d^T(\bar{x}+y) = \delta + d^T y \leq \delta \quad \forall\, y \in \operatorname{Kern}(A_{I\cdot}), \quad \|y\| \leq \epsilon.$$

Dies erzwingt $d^T y = 0$ für alle $y \in \operatorname{Kern}(A_{I\cdot})$ und somit gilt

$$d^T(\bar{x}+y) = \delta \quad \forall\, y \in \operatorname{Kern}(A_{I\cdot}).$$

Somit folgt

$$\mathcal{P} \subset \{x : A_{I\cdot}x = b_I\} = \{\bar{x}+y \ : \ y \in \operatorname{Kern}(A_{I\cdot})\} \subset \{x \ : \ d^T x = \delta\}$$

und dies zeigt $\mathcal{F} = \mathcal{P}$.
Zur Dimensionsformel: Wir haben gezeigt, dass gilt

$$\{\bar{x}+y \ : \ y \in \operatorname{Kern}(A_{I\cdot}), \quad \|y\| \leq \epsilon\} \subset \mathcal{P} \subset \{\bar{x}+y \ : \ y \in \operatorname{Kern}(A_{I\cdot})\}.$$

Somit gilt $\dim(\operatorname{Kern}(A_{I\cdot})) \leq \dim(\mathcal{P}) \leq \dim(\operatorname{Kern}(A_{I\cdot}))$. Schließlich wissen wir nach dem Dimensionssatz über lineare Abbildungen, dass $n = \operatorname{rang}(A_{I\cdot}) + \dim(\operatorname{Kern}(A_{I\cdot}))$. $\square$

Wir verfeinern nun die Dimensionsformel für Seitenflächen.

Satz 3.24 *Ist $\mathcal{F} \neq \emptyset$ eine Seitenfläche des Polyeders $\mathcal{P}(A,b) \subseteq \mathbb{R}^n$, dann gilt*

$$\dim(\mathcal{F}) = n - \operatorname{rang}(A_{eq(\mathcal{F})\cdot}).$$

Beweis. Sei M die Zeilenindexmenge von A und setze $I = \operatorname{eq}(\mathcal{F})$, $J = M \setminus I$. Sei $\mathcal{F}$ von $d^T x \leq \delta$ induziert. Dann gilt

$$\mathcal{F} = \{x \ : \ d^T x = \delta, \quad Ax \leq b\}$$

und die bindenden Nebenbedingungen sind

$$d^T x = \delta, \quad A_{I\cdot}x = b_I,$$

also

$$\mathcal{F} = \{x \ : \ d^T x = \delta, \quad A_{I\cdot}x = b_I, \quad A_{J\cdot}x \leq b_J\}.$$

Nach Satz 3.23 existiert ein innerer Punkt $\bar{x}$ von $\mathcal{F}$ mit

$$d^T\bar{x} = \delta, \quad A_{I\cdot}\bar{x} = b_I, \quad A_{J\cdot}\bar{x} < b_J$$

und es gilt

$$\dim(\mathcal{F}) = \dim(\text{Kern}\begin{pmatrix} A_{I\cdot} \\ d^T \end{pmatrix}).$$

Es bleibt zu zeigen, dass Kern $\binom{A_{I\cdot}}{d^T}$ = Kern $(A_{I\cdot})$. Annahme, es existiert $y \in$ Kern $(A_{I\cdot})$ mit $d^Ty \neq 0$, o.E. $d^Ty > 0$. Wir finden dann $\epsilon > 0$ mit

$$A_{I\cdot}(\bar{x} + \epsilon y) = b_I, \quad A_{J\cdot}(\bar{x} + \epsilon y) < b_J,$$

also $\bar{x} + \epsilon y \in \mathcal{P}$, aber $d^T(\bar{x} + \epsilon y) = \delta + \epsilon d^Ty > \delta$. Dies ist ein Widerspruch zu $\mathcal{P} \subset \{x \ : \ d^Tx \leq \delta\}$. □

Folgerung 3.25 *Sei $\mathcal{P} = \mathcal{P}(A, b) \subseteq \mathbb{R}^n$ ein nichtleeres Polyeder. Dann gilt:*

(a) Ist eq($\mathcal{P}$) $= \emptyset$, dann ist $\mathcal{P}$ volldimensional.

(b) Ist $\mathcal{F}$ eine Seitenfläche von $\mathcal{P}$, dann gilt nach dem Beweis von Satz 3.24 Kern $\binom{A_{I\cdot}}{d^T}$ = Kern $(A_{I\cdot})$ *und somit*

$$\mathcal{F} = \{x \in \mathcal{P} : A_{eq(\mathcal{F})\cdot}x = b_{eq(\mathcal{F})}\}. = \text{fa}(eq(\mathcal{F})).$$

(c) Ist $\mathcal{F}$ eine echte Seitenfläche von $\mathcal{P}$, dann gilt $\dim(\mathcal{F}) \leq \dim(\mathcal{P}) - 1$.

Satz 3.26 *Sei $\mathcal{F}$ eine Seitenfläche eines Polyeders $\mathcal{P}(A, b)$ und $\bar{x} \in \mathcal{F}$. Der Vektor $\bar{x}$ ist genau dann ein innerer Punkt von $\mathcal{F}$, wenn eq($\{\bar{x}\}$) = eq($\mathcal{F}$) gilt.*

Beweis. $\bar{x}$ ist genau dann ein innerer Punkt von $\mathcal{F}$, wenn die kleinste (im Sinne der Mengeninklusion) Seitenfläche $\mathcal{G}$ von $\mathcal{F}$, die $\bar{x}$ enthält, $\mathcal{F}$ selbst ist. Nach Folgerung 3.25, b) gilt

$$\mathcal{F} = \{x : A_{\text{eq}(\mathcal{F})\cdot}x = b_{\text{eq}(\mathcal{F})}, \ A_{M\setminus\text{eq}(\mathcal{F})\cdot}x \leq b_{M\setminus\text{eq}(\mathcal{F})}\}.$$

Nach Folgerung 3.25, b) ist jede $\bar{x}$ enthaltende Seitenfläche $\mathcal{G}$ von der Form

$$\mathcal{G} = \{x : A_{I\cdot}x = b_I, \ A_{M\setminus I\cdot}x \leq b_{M\setminus I}\}, \quad \text{eq}(\mathcal{F}) \subset I \subset \text{eq}(\bar{x})$$

und die minimale solche Seitenfläche ist

$$\mathcal{G} = \{x : A_{\mathrm{eq}(\bar{x})\cdot}x = b_{\mathrm{eq}(\bar{x})},\ A_{M\setminus\mathrm{eq}(\bar{x})\cdot}x \leq b_{M\setminus\mathrm{eq}(\bar{x})}\} = \mathrm{fa}(\mathrm{eq}(\{\bar{x}\})).$$

Daher gilt $\mathcal{F} = \mathcal{G}$ genau dann, wenn $\mathrm{eq}(\mathcal{F}) = \mathrm{eq}(\bar{x})$. □

Im nächsten Satz bezeichnet $\mathrm{aff}\,(\mathcal{F})$ die affine Hülle der Menge $\mathcal{F}$, also den kleinsten affinen Raum, der $\mathcal{F}$ enthält.

Satz 3.27 *Sei $\mathcal{F} \neq \emptyset$ eine Seitenfläche des Polyeders $\mathcal{P}(A,b)$. Dann gilt*

$$\begin{aligned} \mathcal{F} &= \left\{x \in \mathcal{P} : A_{eq(\mathcal{F})\cdot}\, x = b_{eq(\mathcal{F})}\right\} = \mathrm{fa}(eq(\mathcal{F})), \\ \mathrm{aff}\,(\mathcal{F}) &= \left\{x \in \mathbb{R}^n : A_{eq(\mathcal{F})\cdot}\, x = b_{eq(\mathcal{F})}\right\}. \end{aligned}$$

Beweis. $\mathcal{F} = \mathrm{fa}(\mathrm{eq}(\mathcal{F}))$ wurde bereits in Folgerung 3.25, b) gezeigt.
Sei $I = \mathrm{eq}(\mathcal{F})$ und $T = \{x : A_{I\cdot}\, x = b_I\}$. Offenbar ist T ein affiner Raum und wegen $\mathcal{F} \subseteq T$ gilt $\mathrm{aff}\,(\mathcal{F}) \subseteq \mathrm{aff}\,(T) = T$. Sei $s = \dim(\mathcal{F})$. Aus Satz 3.24 folgt $\dim(\mathrm{Kern}\,(A_{I\cdot})) = s$ und somit $\dim(T) = s$. Aus $\dim(\mathrm{aff}\,(\mathcal{F})) = \dim(T)$ und $\mathrm{aff}\,(\mathcal{F}) \subseteq T$ folgt nun $\mathrm{aff}\,(\mathcal{F}) = T$. □

Kapitel 4

Grundlagen der Linearen Optimierung

Wie bereits erwähnt, ist ein lineares Optimierungsproblem eines, bei dem sowohl die Zielfunktion als auch alle Nebenbedingungen lineare Funktionen sind. Man nennt lineare Optimierungsprobleme oft einfach lineare Probleme oder auch lineare Programme und kürzt dies mit LP ab.
Ein allgemeines lineares Optimierungsproblem ist also von der Form

$$\min c^T x \quad \text{s.t.} \quad x \in \mathcal{P}$$

mit einem Polyeder $\mathcal{P} \subset \mathbb{R}^n$ und $c \in \mathbb{R}^n$.
Üblicherweise überführt man ein LP unter Umständen durch Verwendung der Transformationsregeln I und II in eine der folgenden beiden Normalformen:

Natürliche Form:

$$\min c^T x \quad \text{s.t.} \quad Ax \leq b \tag{LP}$$

mit $A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$.

Standardform:

$$\min c^T x \quad \text{s.t.} \quad Ax = b, \quad x \geq 0 \tag{LPS}$$

mit $A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$. Diese Form wird von dem in Kürze besprochenen Simplex-Verfahren verwendet.

In diesem Kapitel werden wir zeigen, dass es zu einem gegebenen LP immer ein dazugehörendes zweites LP gibt, das duale LP. Diese beiden LPs hängen eng zusammen, und man kann aus dem Optimum des einen auf das Optimum des anderen schließen. Zudem werden wir Optimalitätsbedingungen herleiten, die die Basis leistungsfähiger Optimierungsverfahren darstellen.
Insbesondere in der Dualitätstheorie ist es manchmal bequem, anstelle der Minimierungsprobleme (LP) und (LPS) die entsprechenden Maximierungsprobleme zu verwenden. Diese können dann mit $-c$ anstelle c natürlich auch als Minimierungsprobleme geschrieben werden.

4.1 Duales Problem und schwacher Dualitätssatz

Betrachte das lineare Optimierungsproblem

$$\min \quad c^T x \quad \text{s.t.} \quad Ax \le b \tag{LP}$$

mit $A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$. Zur Erinnerung: Ein Vektor $x \in \mathbb{R}^n$ heißt zulässig für (LP), falls $x \in \mathcal{P}(A,b)$. Ein Vektor $\bar{x} \in \mathbb{R}^n$ heißt optimal, falls $c^T x \ge c^T \bar{x}$ für alle $x \in \mathcal{P}(A,b)$.
Wir nehmen einmal an, dass (LP) eine optimale Lösung $\bar{x}$ besitzt. Die Grundidee der Dualitätstheorie besteht darin, eine zweites LP herzuleiten, das ein Maximierungsproblem ist und dessen Optimalwert eine scharfe Unterschranke für den Optimalwert $c^T\bar{x}$ von (LP) liefert.
Wir wollen uns zunächst überlegen, wie wir unter Verwendung der Nebenbedingungen von (LP) zu guten Unterschranken kommen. Wir stellen zunächst fest, dass gilt

$$\begin{aligned} x \in \mathcal{P}(A,b) \quad &\Longleftrightarrow \quad Ax \le b \quad \Longleftrightarrow \quad -Ax \ge -b \\ &\Longleftrightarrow \quad \forall\, y \in \mathbb{R}^m, \quad y \ge 0: \quad -y^T Ax \ge -y^T b. \end{aligned}$$

Somit gilt für alle $y \ge 0$ mit $y^T A = -c^T$ (wir werden sehen, dass es solche y gibt, falls (LP) eine Lösung hat)

$$c^T x = -y^T Ax \ge -y^T b \quad \forall\, x \in \mathcal{P}(A,b).$$

Die beste Unterschranke erhalten wir, wenn wir $-y^T b$ über alle diese y maximieren. Dies führt uns auf folgendes

Duales Problem von (LP)

$$\max -b^T y \quad \text{s.t.} \quad A^T y = -c, \quad y \geq 0. \tag{DP}$$

Nach unseren Vorüberlegungen ist es nicht überraschend, dass der folgende schwache Dualitätssatz gilt.

Satz 4.1 (Schwacher Dualitätssatz) *Sei $A \in \mathbb{R}^{m\times n}$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$. Betrachte das Paar von primalem Problem (P) und dualem Problem (D)*

$$(P) \quad \begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax \leq b \end{array} \qquad\qquad (D) \quad \begin{array}{ll} \max & -b^T y \\ \text{s.t.} & A^T y = -c \\ & y \geq 0 \end{array}$$

Ist $x \in \mathbb{R}^n$ zulässig für (P) und $y \in \mathbb{R}^m$ zulässig für (D), so gilt

$$c^T x \geq -b^T y.$$

Genauer haben wir

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

Beweis. Da y zulässig für (D) ist, gilt $y^T A = -c^T$, $y \geq 0$. und die Zulässigkeit von x für (P) liefert $b - Ax \geq 0$. Damit folgt sofort

$$c^T x - (-b^T y) \overset{-y^T A = c^T}{=} -y^T Ax + b^T y = (b - Ax)^T y \overset{b-Ax\geq 0, y\geq 0}{\geq} 0.$$

□

Wir erhalten ein erstes einfaches Kriterium, um folgende wichtige Frage zu klären:

Kann man auf einfache Weise erkennen, ob überhaupt noch eine Verbesserung möglich ist oder ob bereits ein Optimum erreicht ist?

Korollar 4.2 *Ist in Satz 4.1 $\bar{x} \in \mathbb{R}^n$ zulässig für (P), $\bar{y} \in \mathbb{R}^m$ zulässig für (D) und gilt*

$$c^T \bar{x} = -b^T \bar{y},$$

dann ist $\bar{x}$ optimal für (P) und $\bar{y}$ optimal für (D).

Beweis. Nach dem schwachen Dualitätssatz gilt

$$x \text{ zulässig für (P)} \implies c^T x \geq -b^T \bar{y} = c^T \bar{x}.$$

Also ist $\bar{x}$ optimale Lösung von (P). Analog gilt nach dem schwachen Dualitätssatz

$$y \text{ zulässig für (D)} \implies -b^T y \leq c^T \bar{x} = -b^T \bar{y}$$

und damit ist $\bar{y}$ optimale Lösung von (D). □

Ist das Primalproblem nicht in der natürlichen Form (LP) gegeben, dann kann man es in die Form (LP) bringen. Das zugehörige duale Problem läßt sich dann oft vereinfachen. Eine schöne Symmetrie zwischen Primal- und Dualproblem erhält man in folgendem Fall:

Satz 4.3 *Das zu*

$$\min \quad c^T x \quad s.t. \quad Ax \geq b, \quad x \geq 0 \tag{4.1}$$

duale Problem kann geschrieben werden in der Form

$$\max \quad b^T y \quad s.t. \quad A^T y \leq c, \quad y \geq 0. \tag{4.2}$$

Beweis. (4.1) kann in der Form (LP) geschrieben werden

$$\min c^T x \ \text{ s.t. } \ \begin{pmatrix} -A \\ -I \end{pmatrix} x \leq \begin{pmatrix} -b \\ 0 \end{pmatrix}.$$

Das duale Problem hierzu lautet (bachte $I^T = I$)

$$\max -(-b^T, 0^T) \begin{pmatrix} y \\ z \end{pmatrix} \ \text{ s.t. } \ (-A^T, -I) \begin{pmatrix} y \\ z \end{pmatrix} = -c, \quad y \geq 0, \quad z \geq 0.$$

Nun ist $A^T y + z = c$, $z \geq 0$ äquivalent zu $A^T y \leq c$ und somit ist das duale Problem äquivalent zu

$$\max b^T y \ \text{ s.t. } \ A^T y \leq c, \quad y \geq 0.$$

□

Ein Beispiel: Ölraffinerie

In Ölraffinerien wird Rohöl durch chemische/physikalische Verfahren in gewisse Komponenten zerlegt. Die Ausbeute an Komponenten hängt vom Verfahren (Crackprozess) ab.
Wir nehmen an, dass die Raffinerie A aus Rohöl drei Komponenten (schweres Öl; mittelschweres Öl; leichtes Öl) herstellen will. Sie hat dazu zwei Crackverfahren C_1 und C_2 zur Verfügung, die die folgenden Ausbeuten liefern (ME: Mengeneinheiten; GE: Geldeinheiten):

C_1 :	2 ME schweres Öl	
	2 ME mittelschweres Öl	Kosten: 3 GE
	1 ME leichtes Öl	
C_2 :	1 ME schweres Öl	
	2 ME mittelschweres Öl	Kosten: 5 GE
	4 ME leichtes Öl	

Darüber hinaus hat die Firma folgende Lieferverpflichtungen:
3 ME schweres Öl,
5 ME mittelschweres Öl,
4 ME leichtes Öl.

Diese Mengen sollen so kostengünstig wie möglich produziert werden.

Mathematische Modellierung des Problems

Variablen:

$$\begin{aligned} x_1 &: \text{ Produktionsniveau von } C_1, \\ x_2 &: \text{ Produktionsniveau von } C_2. \end{aligned}$$

Nebenbedingungen:

$$\begin{aligned} 2x_1 + x_2 &\geq 3 && \text{schweres Öl} \\ 2x_1 + 2x_2 &\geq 5 && \text{mittelschweres Öl} \\ x_1 + 4x_2 &\geq 4 && \text{leichtes Öl} \\ x_1 &\geq 0 \\ x_2 &\geq 0. \end{aligned}$$

Minimierung der Kosten:

$$\min 3x_1 + 5x_2$$

Wir können unser LP also in folgender Form schreiben:

$$\begin{aligned} \min\ & 3x_1 + 5x_2 && (4.3)\\ \text{s.t.}\ & 2x_1 + \ \ x_2 \geq 3 && (4.3a)\\ & 2x_1 + 2x_2 \geq 5 && (4.3b)\\ & x_1 + 4x_2 \geq 4 && (4.3c)\\ & x_1 \geq 0 && (4.3d)\\ & x_2 \geq 0 && (4.3e) \end{aligned}$$

Wie bereits in Kapitel 1 erwähnt, heißt die Funktion $3x_1 + 5x_2$ in (4.3) Zielfunktion. Jeder Punkt $(x_1, x_2)^T$, der $(4.3a)$-$(4.3e)$ erfüllt, heißt zulässiger Punkt des Problems (manchmal auch zulässige Lösung genannt).

Grafische Darstellung des Problems

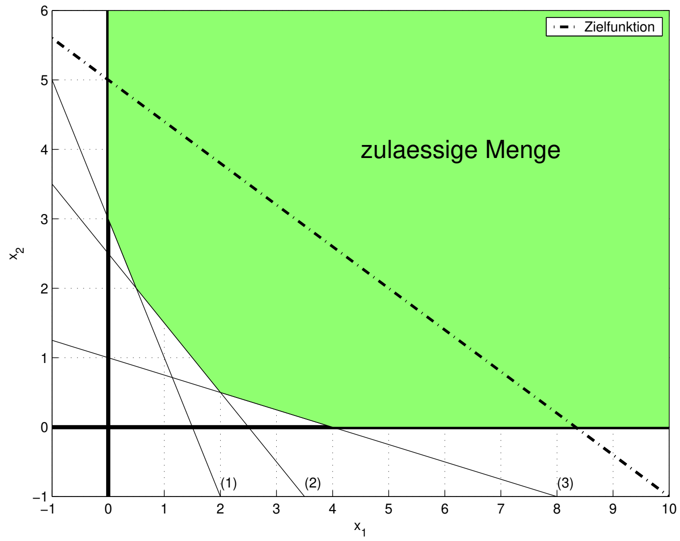

Abbildung 4.1: Grafische Darstellung zu (4.3)

Auf der Geraden

$$G_\gamma = \{x \in \mathbb{R}^2 : 3x_1 + 5x_2 = \gamma\}$$

hat die Zielfunktion den Wert $c^T x = \gamma$. Verschiebt man die Gerade G_{25} in Abb. 4.1 in Richtung des Ursprungs $(0,0)^T$, bis die verschobene Gerade die zulässige Menge nur noch tangiert, erhält man die zu G_{25} parallele Gerade

$$G_{8.5} = \{x \in \mathbb{R}^2 : 3x_1 + 5x_2 = 8.5\},$$

welche die zulässige Menge genau im Punkt $x^* = (2, \frac{1}{2})^T$ schneidet. Jede weitere Verschiebung erzeugt einen leeren Schnitt der Geraden mit der zulässigen Menge. Daraus folgt, dass x^* die Optimallösung des Problems ist, d.h. die minimalen Kosten betragen 8.5 GE.
Eine grafische Lösung des Problems ist i.A. in höheren Dimensionen (mehr als zwei Variablen) nicht durchführbar. Auf der Suche nach anderen Lösungsmethoden betrachten wir zunächst eine etwas allgemeinere Zielfunktion

$$ax_1 + bx_2.$$

Hier können wir sehen, dass unser Problem gar keine Optimallösung besitzen muss. Zum Beispiel gilt mit $a < 0$ und der Folge $x^n = (n, 0)^T$

$$ax_1^n + bx_2^n = an \to -\infty \quad \text{für } n \to \infty.$$

Für unser reales Problem ist $a < 0$ natürlich unsinnig, aber mathematisch müssen solche Fälle behandelt werden. Ist $a \geq 0$ und $b \geq 0$, so gibt es immer eine Optimallösung.
Wir haben bereits gesehen, dass die Optimallösung eines LPs (wenn sie existiert) immer in einer Ecke angenommen wird (vgl. Korollar 2.40). Um das LP zu lösen, könnte man also alle Ecken berechnen und ausprobieren, wo der beste Zielfunktionswert auftritt. Ein bester unter diesen Punkten (er muss nicht eindeutig sein) ist eine Optimallösung. Dieses Verfahren ist zwar endlich, aber nicht sehr effizient und in der Praxis nicht einsetzbar. Dennoch steckt darin der Kern eines der wichtigsten Verfahren zur Lösung linearer Probleme, des Simplexverfahrens. Man startet dabei mit einer zulässigen Ecke, die der Durchschnitt von n Hyperebenen ist, und versucht dann systematisch, die Lösung zu verbessern, indem man entlang einer Kante zu einer benachbarten Ecke wandert.

Duales Problem für das Beispiel

Um zu zeigen, dass $(2, \frac{1}{2})^T$ optimal ist, verwenden wir das duale Problem und Korollar 4.2.

Unser Problem (4.3) ist bereits in der Form (4.1)

$$\min \quad c^T x \quad \text{s.t.} \quad Ax \geq b, \quad x \geq 0, \tag{LP}$$

wobei

$$A = \begin{pmatrix} 2 & 1 \\ 2 & 2 \\ 1 & 4 \end{pmatrix}, \quad b = \begin{pmatrix} 3 \\ 5 \\ 4 \end{pmatrix}, \quad c = \begin{pmatrix} 3 \\ 5 \end{pmatrix}.$$

Das duale Problem lautet nach Satz 4.3

$$\max b^T y \quad \text{s.t.} \quad A^T y \leq c, \quad y \geq 0,$$

also

$$\begin{array}{lrcrcrcl} \max & 3y_1 & + & 5y_2 & + & 4y_3 & & \\ \text{s.t.} & 2y_1 & + & 2y_2 & + & y_3 & \leq & 3 \\ & y_1 & + & 2y_2 & + & 4y_3 & \leq & 5 \\ & y_1 & & & & & \geq & 0 \\ & & & y_2 & & & \geq & 0 \\ & & & & & y_3 & \geq & 0 \end{array} \tag{4.4}$$

Betrachten wir nun den Punkt $y^* = \left(0, \frac{7}{6}, \frac{2}{3}\right)^T$, so ist dies ein zulässiger Punkt des dualen Problems und wir erhalten

$$3y_1 + 5y_2 + 4y_3 = 8.5 = 3x_1 + 5x_2,$$

d.h. x^* ist tatsächlich optimal für (4.3) nach Korollar 4.2 und y^* ist auch optimal für (4.4).

Ökonomische Interpretation

Ein konkurrierender Anbieter B überlegt sich, wie teuer er die drei Ölsorten anbieten darf, damit er

- nicht teurer ist als die Crackprozesse $C1$ und $C2$,
- hierbei maximalen Gewinn macht.

Seien nun y_1, y_2, y_3 die Preise von Lieferant B für die Ölsorten S (schweres Öl), M (mittelschweres Öl) und L (leichtes Öl). Der Verkauf der benötigten Mengen bringt die Einnahme

$$3y_1 + 5y_2 + 4y_3 \quad \text{Geldeinheiten}$$

Um nicht teurer als der Crackprozess $C1$ zu sein, müßte

$$2y_1 + 2y_2 + y_3 \leq 3$$

gelten. Für den Crackprozess C_2 ergibt sich analog die Bedingung

$$y_1 + 2y_2 + 4y_3 \leq 5.$$

Wenn wir die Einnahmen $3y_1 + 5y_2 + 4y_3$ von Anbieter B unter den obigen Nebenbedingungen maximieren, ergibt sich genau das duale Problem (4.4). Wie wir bald sehen werden (vgl. Dualitätssatz unten), stimmt der maximale Erlös $b^T\bar{y}$ für Anbieter B genau überein mit dem besten Preis $c^T\bar{x}$, den Raffinerie A anbieten kann. Die Optimallösungen $y^* = \left(0, \frac{7}{6}, \frac{2}{3}\right)^T$ werden auch Schattenpreise genannt.

Interpretation der Lösung:

$y_1^* = 0$: will Raffinerie A die von S produzierte Menge infinitesimal erhöhen, dann kostet sie das nichts, da davon immer genügend produziert wird.
$y_2^* = \frac{7}{6}$: heißt, um die von M produzierte Menge infinitesimal zu erhöhen, entstehen Zusatzkosten in Höhe von $\frac{7}{6}$ Geldeinheiten pro Mengeneinheit.
$y_3^* = \frac{2}{3}$: analog.

Der schwache Dualitätssatz ist noch unbefriedigend und wird im Folgenden verschärft. Wir wollen insbesondere

- zeigen, dass primales und duales Problem immer Optimallösungen mit gleichem Optimalwert haben, falls sie beide zulässige Lösungen besitzen,
- griffige Optimalitätsedingungen angeben.

Wir starten zunächst mit einer ersten Optimalitätsbedingung für (LP).

Satz 4.4 *$\bar{x}$ ist genau dann Optimallösung von*

$$\min c^T x \quad \text{s.t.} \quad Ax \leq b, \tag{LP}$$

$A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$, wenn gilt

$$\begin{aligned} &A\bar{x} \leq b, \\ &c^T s \geq 0 \quad \forall\, s \in \mathcal{Z}(\bar{x}) := \{s : A_{eq(\bar{x})\cdot}s \leq 0\}. \end{aligned} \tag{4.5}$$

Bemerkung: $\mathcal{Z}(\bar{x})$ nennt man die Menge der zulässigen Richtungen: Tatsächlich gibt es, wie wir gleich zeigen werden, zu jedem $s \in \mathcal{Z}(\bar{x})$ ein $\bar{\lambda} > 0$ mit

$$\bar{x} + \lambda s \;\text{ zulässig für (LP)} \;\; \forall\, \lambda \in [0, \bar{\lambda}].$$

Damit bedeutet (4.5): in jede zulässige Richtung nimmt die Zielfunktion nicht ab.

Beweis. Sei $\bar{x}$ optimale Lösung von (LP). Dann gilt $A\bar{x} \leq b$. Für jedes s mit $A_{\mathrm{eq}(\bar{x})}s \leq 0$ ist $\bar{x} + \lambda s$ für kleines $\lambda > 0$ zulässig für (LP) (inaktive Nebenbedingungen bleiben inaktiv). Daher gilt

$$0 \leq c^T(\bar{x} + \lambda s) - c^T\bar{x} = \lambda c^T s$$

und somit gilt (4.5).
Ist $\bar{x}$ keine optimale Lösung von (LP), dann ist entweder $\bar{x}$ nicht zulässig, also $A\bar{x} \not\leq b$, oder es existiert x mit

$$Ax \leq b, \quad c^T x < c^T \bar{x}.$$

Also gilt $A_{\mathrm{eq}(\bar{x})\cdot}x \leq b_{\mathrm{eq}(\bar{x})} = A_{\mathrm{eq}(\bar{x})\cdot}\bar{x}$ und daher erfüllt $s = x - \bar{x}$ zwar $A_{\mathrm{eq}(\bar{x})}s \leq 0$, aber $c^T s < 0$ und damit gilt (4.5) nicht. □

Um diese Optimalitätsbedingung in eine bequemere Form zu überführen, benötigen wir das Lemma von Farkas.

4.2 Das Farkas-Lemma

Satz 4.5 (Farkas-Lemma) *Für eine gegebene Matrix $A \in \mathbb{R}^{m\times n}$ und einen Vektor $b \in \mathbb{R}^m$ hat genau eines der folgenden Systeme eine Lösung:*

$$\begin{array}{c} Ax = b \\ x \geq 0 \end{array} \qquad \dot{\bigvee} \qquad \begin{array}{c} y^T A \geq 0 \\ y^T b < 0 \end{array}$$

Geometrische Interpretation:

Entweder gilt:	$\begin{array}{rcl} Ax &=& b\,, \\ x &\geq& 0 \end{array}$	d.h.	b liegt im Kegel, der von $A_{\cdot 1}, \ldots, A_{\cdot n}$ aufgespannt wird,
oder es gilt	$\begin{array}{rcl} y^T A &\geq& 0\,, \\ y^T b &<& 0 \end{array}$	d.h.	es gibt eine Hyperebene, die b von $A_{\cdot 1}, \ldots, A_{\cdot n}$ trennt,

aber es kann nicht beides gleichzeitig gelten.

Zum Beweis benötigen wir das nichttriviale

Lemma 4.6 *Zu beliebigem $A \in \mathbb{R}^{m\times n}$ ist die Menge $\{Ax \ : \ x \geq 0\}$ abgeschlossen.*

Wir bringen den Beweis dieses Lemmas nach dem Beweis von Satz 4.5.
Beweis. (von Satz 4.5) Es können nicht beide Fälle eintreten, denn aus $x \geq 0$ und $y^T A \geq 0$ folgt

$$0 \leq (y^T A)x = y^T(Ax) \leq y^T b < 0,$$

ein Widerspruch. Betrachte nun die Menge

$$\mathcal{K} = \{Ax : x \geq 0\}.$$

Dann ist $\mathcal{K}$ ein Kegel (d.h. aus $p \in \mathcal{K}$ folgt $\lambda p \in \mathcal{K}$ für alle $\lambda \geq 0$), denn mit $\lambda \geq 0$ und $p = Ax \in \mathcal{K}$ ist auch $\lambda p = A \underbrace{(\lambda x)}_{\geq 0} \in \mathcal{K}$.
Ausserdem ist $\mathcal{K}$ konvex, denn für $\lambda \in [0,1]$ und $p_1 = Ax_1 \in \mathcal{K}$ und $p_2 = Ax_2 \in \mathcal{K}$ ist

$$\lambda p_1 + (1-\lambda)p_2 = \lambda A x_1 + (1-\lambda)Ax_2 = A \underbrace{(\lambda x_1 + (1-\lambda)x_2)}_{\geq 0} \in \mathcal{K}.$$

Schließlich ist $\mathcal{K}$ nach Lemma 4.6 abgeschlossen.
Wir unterscheiden nun zwei Fälle:
$\boxed{b \in \mathcal{K}:}$ Dann gilt $Ax = b$ für ein $x \geq 0$, d.h. das System $Ax = b, x \geq 0$ hat eine Lösung.
$\boxed{b \notin \mathcal{K}:}$ $\mathcal{K}$ ist nichtleer, konvex und abgeschlossen. Nach dem strikten Trennungssatz 2.19 existiert also eine Hyperbene $H = \{x : \ y^T x = \delta\}$, die $\mathcal{K}$ und $\{b\}$ strikt trennt, also

$$y^T z \geq \delta \quad \forall\, z \in \mathcal{K}, \quad y^T b < \delta. \tag{4.6}$$

Wegen $A_{\cdot i} \in \mathcal{K}$ gilt dann $y^T A_{\cdot i} \geq \delta$ und somit

$$y^T A \geq \delta, \quad y^T b < \delta. \tag{4.7}$$

Es bleibt zu zeigen, dass wir $\delta = 0$ in (4.6) und somit auch in (4.7) wählen können. Jedenfalls muss $\delta \leq 0$ sein, da wegen $0 \in \mathcal{K}$ gilt $0 = y^T 0 \geq \delta$. Also gilt (4.6) mit einem $\delta \leq 0$. Dann ist (4.6) aber auch für $\delta = 0$ richtig. Falls nicht, dann gäbe es $z \in \mathcal{K}$ mit

$$y^T z < 0, \quad \text{also} \quad \lim_{\lambda\to\infty} y^T(\lambda z) = -\infty$$

im Widerspruch zu $y^T(\lambda z) \geq \delta$ wegen $\lambda z \in \mathcal{K}$. Damit gilt (4.7) mit $\delta = 0$ wie behauptet. $\square$

Beweis. (von Lemma 4.6) Zum Nachweis der Abgeschlossenheit von $\mathcal{K}$ zeigen wir, dass gilt

$$\mathcal{K} = \bigcup_{I \subset \{1,\ldots,n\}, \quad A_{\cdot I} \text{ hat linear unabhängige Spalten}} \mathcal{K}_I \tag{4.8}$$

mit $\mathcal{K}_I = \{A_{\cdot I} x_I : \quad x_I \in \mathbb{R}^{|I|}, \quad x_I \geq 0\}$.
Da jedes der $\mathcal{K}_I$ abgeschlossen ist, ist dann auch $\mathcal{K}$ als endliche Vereinigung abgeschlossener Mengen abgeschlossen.
Jedes $\mathcal{K}_I$ ist abgeschlossen: Sei $(y_k) \subset \mathcal{K}_I$ eine beliebige konvergente Folge mit Grenzwert y. Wir müssen $y \in \mathcal{K}_I$ nachweisen. Es existiert eine Folge $(x_{I,k}) \subset \mathbb{R}_+^{|I|}$ mit $y_k = A_{\cdot I} x_{I,k}$. Da $A_{\cdot I}$ linear unabhängige Spalten hat, ist die Folge $(x_{I,k})$ eindeutig bestimmt, nämlich $x_{I,k} = (A_{\cdot I}^T A_{\cdot I})^{-1} A_{\cdot I}^T y_k$. Somit existiert $\lim_{k\to\infty} x_{I,k} = x_I \in \mathbb{R}_+^{|I|}$ und es gilt

$$y = \lim_{k\to\infty} y_k = \lim_{k\to\infty} A_{\cdot I} x_{I,k} = A_{\cdot I} x_I \in \mathcal{K}_I.$$

Nachweis von (4.8): "$\supset$": Ist wegen $\mathcal{K}_I \subset \mathcal{K}$ klar.

"$\subset$": Sei $y \in \mathcal{K}$ beliebig. Dann existiert $x \geq 0$ mit

$$y = Ax = \sum_{i=1}^n A_{\cdot i} x_i = \sum_{i=1}^n \frac{1}{n+1}(n+1) A_{\cdot i} x_i + \frac{1}{n+1} 0.$$

Damit ist y eine Konvexkombination der Punkte $y_1 := (n+1)A_{\cdot 1}x_1, \ldots, y_n := (n+1)A_{\cdot n}, y_{n+1} := 0 \in \mathcal{K}$. Genau wie im Beweis des Satzes von Carathéodory (Satz 2.9) kann man aus der Konvexkombination so lange jeweils einen der Punkte $y_1, \ldots, y_n$ entfernen, bis für die verbleibenden Punkte y_i, $i \in I$, die Vektoren $y_i - y_{n+1} = y_i = (n+1)A_{\cdot i}$, $i \in I$, linear unabhängig sind. Diese verkürzte Konvexkombination für y liegt dann in $\mathcal{K}_I$. $\square$

Folgerung 4.7 *Seien $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Dann hat genau eines der beiden folgenden Systeme eine Lösung:*

$$Ax \leq b \qquad \dot{\bigvee} \qquad \begin{array}{rcl} y^T A & = & 0 \\ y & \geq & 0 \\ y^T b & < & 0. \end{array}$$

Beweis. Übung. □

Folgerung 4.8 *Für dimensionsverträgliche Matrizen A, B, C und D sowie Vektoren a, b, u und v hat genau eines der beiden folgenden Systeme eine Lösung:*

$$\begin{array}{rcrcl} Ax & + & By & \leq & a \\ Cx & + & Dy & = & b \\ x & & & \geq & 0 \end{array} \qquad \dot{\bigvee} \qquad \begin{array}{rcrcl} u^TA & + & v^TC & \geq & 0 \\ u^TB & + & v^TD & = & 0 \\ u & & & \geq & 0 \\ u^Ta & + & v^Tb & < & 0. \end{array}$$

Beweis. Übung. □

4.3 Optimalitätsbedingungen und der starke Dualitätssatz der Linearen Optimierung

Um die Notationen kompakt zu halten, werden wir im Folgenden alle Resultate für das Paar von primalem und dualem Problem

$$(P) \quad \begin{array}{ll} \min & c^Tx \\ \text{s.t.} & Ax \leq b \end{array} \qquad\qquad (D) \quad \begin{array}{ll} \max & -b^Ty \\ \text{s.t.} & A^Ty = -c \\ & y \geq 0 \end{array}$$

beweisen. Liegt ein beliebiges LP vor, dann kann es in die Form (P) gebracht, das zugehörige Problem (D) hergeleitet und unter Umständen vereinfacht werden.
Der Bequemlichkeit halber geben wir die Form des dualen Problems für ein primales LP in ganz allgemeiner Form an.

Satz 4.9 *Gegeben seien dimensionsverträgliche Matrizen A, B, C, D und Vektoren a, b, c, d.*
Betrachte ein primales lineares Problem (PA) der Form

$$(PA) \quad \begin{array}{ll} \min & c^Tx + d^Ty \\ \text{s.t.} & Ax + By \geq a \\ & Cx + Dy = b \\ & x \geq 0 \end{array}$$

Dann kann das duale lineare Problem zu (PA) geschrieben werden in der Form

$$(DA) \qquad \begin{array}{ll} \max & a^T u + b^T v \\ \text{s.t.} & A^T u + C^T v \leq c \\ & B^T u + D^T v = d \\ & u \geq 0 \end{array}$$

Beweis. (PA) kann folgendermaßen in der Form (P) geschrieben werden:

$$(P) \qquad \begin{array}{ll} \min & \begin{pmatrix} c \\ d \end{pmatrix}^T \begin{pmatrix} x \\ y \end{pmatrix} \\ \text{s.t.} & \begin{pmatrix} -A & -B \\ C & D \\ -C & -D \\ -I & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} -a \\ b \\ -b \\ 0 \end{pmatrix} \end{array} \tag{4.9}$$

Das duale Problem hierzu ist

$$(D) \qquad \begin{array}{ll} \max & -\begin{pmatrix} -a \\ b \\ -b \\ 0 \end{pmatrix}^T \begin{pmatrix} u \\ s \\ w \\ z \end{pmatrix} \\ \text{s.t.} & \begin{pmatrix} -A^T & C^T & -C^T & -I \\ -B^T & D^T & -D^T & 0 \end{pmatrix} \begin{pmatrix} u \\ s \\ w \\ z \end{pmatrix} = -\begin{pmatrix} c \\ d \end{pmatrix}, \quad u, s, w, z \geq 0. \end{array} \tag{4.10}$$

Setzen wir $v = w - s$, dann können wir (D) in der Form schreiben

$$\begin{array}{ll} \max & a^T u + b^T v \\ \text{s.t.} & A^T u + C^T v + z = c \\ & B^T u + D^T v = d \\ & u, z \geq 0 \end{array}$$

und dieses Problem ist offensichtlich äquivalent zu (DA), da z lediglich eine Schlupfvariable ist. $\square$

Bemerkung 4.10 *Das duale Problem von (DA) ist wieder (PA).*

Beweis. Übung. □

Daher macht es keinen Unterschied, ob man (PA) oder (DA) als das primale Problem betrachtet. Insbesondere ist es unerheblich, ob das primale LP ein Minimierungs- oder ein Maximierungsproblem ist.
(PA) und (DA) weisen eine schöne Symmetrie auf. Für ein beliebiges LP kann man durch Weglassen überflüssiger Variablen in (PA) leicht das zugehörige duale Problem (DA) bestimmen.
In der folgenden Tabelle 4.1 sind einige Merkregeln zusammengefasst, wie man aus einem primalen LP das zugehörige duale LP ableitet und umgekehrt.

Primal	Dual
Ungleichung	vorzeichenbeschränkte Variable
Gleichung	vorzeichenunbeschränkte Variable
vorzeichenbeschränkte Variable	Ungleichung
vorzeichenunbeschränkte Variable	Gleichung

Tabelle 4.1: Regeln für das Dualisieren von LPs

In Tabelle 4.2 sind einige Beispiele zu finden, abgeleitet aus Satz 4.9 bzw. Bemerkung 4.10.

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

Tabelle 4.2: Paare zueinander gehörender primaler und dualer LPs

Bemerkung 4.11 *Auch für die allgemeine Form von primal-dualen LPs (wie in Satz 4.9) gilt, da sie äquivalent zu einem primal-dualen Paar (P) und (D) sind, natürlich* schwache Dualität*, das heißt:*
Ist $(\bar{x}, \bar{y})$ zulässig für (PA) und $(\bar{u}, \bar{v})$ zulässig für (DA), so gilt

$$c^T \bar{x} + d^T \bar{y} \geq a^T \bar{u} + b^T \bar{v}.$$

Beweis. Einfache Übung. □

Mit Hilfe des Lemmas von Farkas können wir nun handliche Optimalitätsbedingungen angeben.

Satz 4.12 (Karush-Kuhn-Tucker-Bedingungen) *Betrachte das Paar von primalem Problem* (P) *und dualem Problem* (D)

$$(P)\quad \begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax \le b \end{array} \qquad\qquad (D)\quad \begin{array}{ll} \max & -b^T y \\ \text{s.t.} & A^T y = -c \\ & y \ge 0 \end{array}$$

mit $A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$. *Dann ist* $x \in \mathbb{R}^n$ *genau dann Optimallösung von (P), wenn ein* $y \in \mathbb{R}^m$ *existiert, so dass gilt*

$$\text{(KKT)}\quad \begin{array}{ll} Ax \le b & \text{(primale Zulässigkeit)}, \\ A^T y = -c, \quad y \ge 0 & \text{(duale Zulässigkeit)}, \\ y_i = 0 \ \text{ oder } \ b_i - A_{i\cdot} x = 0, \quad i = 1, \ldots, m & \text{(Komplementarität)}. \end{array}$$

y ist dann Optimallösung von (D) mit demselben Optimalwert wie (P).
Umgekehrt ist $y \in \mathbb{R}^m$ *Optimallösung von (D) genau dann, wenn* $x \in \mathbb{R}^n$ *existiert, so dass (KKT) gilt.* x *ist dann Optimallösung von (P) mit demselben Optimalwert wie (D).*

Beweis. Sei x Optimallösung von (P) und setze $I = \text{eq}(x)$. Nach Satz 4.4 hat dann das System

$$A_{I\cdot} s \le 0, \quad c^T s < 0$$

keine Lösung. Daher existiert nach dem Farkas-Lemma (Satz 4.5) ein z mit

$$-A_{I\cdot}^T z = c, \quad z \ge 0.$$

Definieren wir nun $y \in \mathbb{R}^m$ durch $y_I = z$, $y_{\{1,\ldots,m\}\setminus I} = 0$, dann gilt

$$A^T y = A_{I\cdot}^T y_I = -c, \quad y \ge 0$$

und zudem $y_i = 0$ oder $A_{i\cdot} x - b_i = 0$, $1 \le i \le m$. Also gilt (KKT).
Gilt umgekehrt (KKT), dann ist x primal zulässig und y dual zulässig. Weiter ist wegen $A_{I\cdot} x = b_I$

$$c^T x - (-y^T b) = y^T (A x - b) = 0,$$

nach Korollar 4.2 ist also x Optimallösung von (P) und y Optimallösung von (D) mit denselben Optimalwerten.
Sei nun y Optimallösung von (D). Der Nachweis, dass dann (KKT) mit geeignetem x gilt, folgt durch entsprechende Argumentation mit (D) statt (P) oder folgt alternativ aus dem nachfolgend bewiesenen starken Dualitätssatz 4.13 iii) und iv). □

Die Komplementaritätsbedingung bedeutet also nichts anderes, als dass primale und duale Zielfunktion übereinstimmen. Sie besagt: ist die primale Nebenbedingung $A_{i\cdot}x \leq b_i$ inaktiv, dann verschwindet die zugehörige Dualvariable y_i.
Unser nächstes Ziel ist es nun zu zeigen, dass es primal und dual zulässige Punkte gibt, die die Ungleichung des schwachen Dualitätssatzes mit Gleichheit erfüllen, falls beide Probleme zulässige Punkte besitzen. Dieser sogenannte starke Dualitätssatz und seine Konsequenzen haben eine außerordentliche Tragweite in der linearen Optimierung.

Satz 4.13 (Starker Dualitätssatz) *Folgende Aussagen sind äquivalent:*

(i) (P) und (D) haben beide zulässige Lösungen.

(ii) (P) hat eine endliche Optimallösung.

(iii) (P) und (D) haben Optimallösungen, deren Zielfunktionswerte übereinstimmen.

(iv) (D) hat eine endliche Optimallösung.

(v) (P) oder (D) hat eine endliche Optimallösung.

(vi) (P) hat zulässige Lösungen und ist nach unten beschränkt oder (D) hat zulässige Lösungen und ist nach oben beschränkt.

Der Satz gilt auch für (PA) und (DA).

Beweis. Wir zeigen den Satz für das Paar primal/dualer Probleme in der Form

$$(P) \quad \begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax \leq b \end{array} \qquad\qquad (D) \quad \begin{array}{ll} \max & -b^T y \\ \text{s.t.} & A^T y = -c \\ & y \geq 0 \end{array}$$

Für Probleme in allgemeiner Form (vgl. Satz 4.9) folgt die Aussage dann unmittelbar, da (PA), (DA) wie in Satz Satz 4.9 in der Form (P), (D) geschrieben werden können.

$\boxed{\text{(i)} \Longrightarrow \text{(ii):}}$ Sei x zulässig für (P) und y zulässig für (D). Aus dem schwachen Dualitätssatz wissen wir bereits, dass $c^T x \geq -b^T y$ gilt. Damit hat (P) einen zulässigen Punkt und ist wegen $p^* := \inf\{c^T x : \quad Ax \leq b\} \geq -b^T y$ nach unten beschränkt. Wir zeigen, dass (P) dann eine optimale Lösung hat. Falls nicht, dann hat das System

$$\begin{pmatrix} A \\ c^T \end{pmatrix} x \leq \begin{pmatrix} b \\ p^* \end{pmatrix} \tag{4.11}$$

keine Lösung. Nach Folgerung 4.7 zum Farkas-Lemma existiert also $\binom{y}{\gamma} \in \mathbb{R}^{m+1}$ mit

$$(y^T, \gamma) \begin{pmatrix} A \\ c^T \end{pmatrix} = 0, \quad y, \gamma \geq 0, \tag{4.12}$$

$$y^T b + \gamma p^* < 0. \tag{4.13}$$

Multipliziert man (4.12) von rechts mit einem primal zulässigen x (existiert nach (i)) und setzt (4.13) ein, dann erhält man wegen $Ax \leq b$ und $y \geq 0$

$$0 = y^T Ax + \gamma c^T x \leq y^T b + \gamma c^T x < -\gamma p^* + \gamma c^T x.$$

Aus $c^T x \geq p^*$ folgt nun $\gamma > 0$ und daher ergibt (4.12), (4.13) mit $\bar{y} = (1/\gamma) y$

$$A^T \bar{y} = -c, \quad \bar{y} \geq 0, \quad -b^T \bar{y} > p^*.$$

Damit ist $\bar{y}$ dual zulässig und es existiert $\bar{x}$ primal zulässig mit $-b^T \bar{y} > c^T \bar{x}$ im Widerspruch zum schwachen Dualitätssatz. Somit hat (4.11) doch eine Lösung und diese ist optimal für (P).

$\boxed{\text{(ii)} \Longrightarrow \text{(iii):}}$ Sei $\bar{x}$ Optimallösung von (P). Nach Satz 4.12 erfüllt dann $\bar{x}$ mit geeignetem $y \in \mathbb{R}^m$ (KKT) und y ist Optimallösung von (D) mit gleichem Optimalwert wie (P).

$\boxed{\text{(iii)} \Longrightarrow \text{(iv):}}$ Ist klar.

$\boxed{\text{(iv)} \Longrightarrow \text{(v):}}$ Ist trivial.

$\boxed{\text{(v)} \Longrightarrow \text{(vi):}}$ Ist klar.

$\boxed{\text{(vi)} \Longrightarrow \text{(i):}}$ Wir nehmen an, (P) habe einen zulässigen Punkt $\bar{x}$ und sei nach unten beschränkt, aber (D) sei unzulässig. Die Unzulässigkeit von $A^T y = -c,\ y \geq 0$ impliziert nach dem Farkas-Lemma (Satz 4.5), dass

$$\begin{array}{rcl} u^T A^T & \geq & 0 \\ -u^T c & < & 0 \end{array} \quad \Longleftrightarrow \quad \begin{array}{rcl} Au & \geq & 0 \\ c^T u & > & 0 \end{array}$$

eine Lösung $\bar{u}$ hat. Damit ist für alle $\lambda < 0$ der Punkt $\bar{x} + \lambda\bar{u}$ zulässig für (P), denn

$$A(\bar{x} + \lambda\bar{u}) = A\bar{x} + \underbrace{\lambda \underbrace{A\bar{u}}_{\geq 0}}_{\leq 0} \leq b.$$

Für die Zielfunktionswerte gilt

$$c^T(\bar{x} + \lambda\bar{u}) = c^T\bar{x} + \underbrace{\lambda \underbrace{c^T\bar{u}}_{>0}}_{<0} \to -\infty \quad \text{für } \lambda \to -\infty$$

was ein Widerspruch zur Beschränktheit von (P) nach unten ist.
Der zweite Fall, dass (D) zulässig und nach oben beschränkt ist, lässt sich analog behandeln, siehe Übung. □

Wir ziehen nun einige wichtige Folgerungen.

Folgerung 4.14 *Betrachte folgendes Paar primaler und dualer Probleme (P) und (D):*

$$(P) \quad \begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax \leq b \end{array} \qquad \textit{und} \qquad (D) \quad \begin{array}{ll} \max & -b^T y \\ \text{s.t.} & A^T y = -c \\ & y \geq 0 \end{array}$$

Seien $\mathcal{P}$ bzw. $\mathcal{D}$ die zulässigen Mengen von (P) bzw. (D). Definiere nun

$$\inf(P) = \begin{cases} +\infty & \textit{falls } \mathcal{P} = \emptyset, \\ -\infty & \textit{falls (P) unbeschränkt,} \\ \min\{c^T x : x \in \mathcal{P}\} & \textit{sonst,} \end{cases}$$

$$\sup(D) = \begin{cases} -\infty & \textit{falls } \mathcal{D} = \emptyset, \\ +\infty & \textit{falls (D) unbeschränkt,} \\ \max\{-b^T y : y \in \mathcal{D}\} & \textit{sonst.} \end{cases}$$

Dann gilt:

$$
\begin{array}{llcll}
(a) & \inf(P) = -\infty & \Longrightarrow & \mathcal{D} = \emptyset. & \\
(b) & \sup(D) = +\infty & \Longrightarrow & \mathcal{P} = \emptyset. & \\
(c) & \mathcal{P} = \emptyset & \Longrightarrow & \mathcal{D} = \emptyset \quad \textit{oder} & \sup(D) = +\infty. \\
(d) & \mathcal{D} = \emptyset & \Longrightarrow & \mathcal{P} = \emptyset \quad \textit{oder} & \inf(P) = -\infty.
\end{array}
$$

Beweis. (a) Ist $y \in \mathcal{D} \neq \emptyset$ so folgt aus dem schwachen Dualitätssatz 4.1, dass $\inf(P) \geq b^T y$ gilt, was ein Widerspruch ist.
(b) Analog zu (a).
(c) Ist $\mathcal{D} \neq \emptyset$ und $\sup(D) \neq +\infty$ dann gilt Satz 4.13 (vi), somit auch (i), also $\mathcal{P} \neq \emptyset$.
(d) Analog zu (c). □

Bemerkung 4.15 *Die Aussagen von Folgerung 4.14 gelten analog für das Paar dualer Probleme aus Satz 4.9.*

Beispiel 4.16 *Seien $\mathcal{P}$ und $\mathcal{D}$ zwei Polyeder:*

$$
\mathcal{P} = \left\{ x \in \mathbb{R}^2 \;\middle|\; \begin{array}{rcr} x_1 - x_2 & \leq & 0 \\ -x_1 + x_2 & \leq & -1 \end{array} \right\},
$$

$$
\mathcal{D} = \left\{ y \in \mathbb{R}^2 \;\middle|\; \begin{array}{r} y_1 - y_2 = 1 \\ -y_1 + y_2 = 1 \end{array},\; y \geq 0 \right\}.
$$

Dann sind die linearen Probleme

$$
\begin{array}{ll} \min & -x_1 - x_2 \\ \text{s.t.} & x \in \mathcal{P} \end{array} \qquad \textit{und} \qquad \begin{array}{ll} \max & y_2 \\ \text{s.t.} & y \in \mathcal{D} \end{array}
$$

dual zueinander und es gilt:

(a) $\mathcal{P} = \emptyset$, da die Addition der Nebenbedingungen zu der ungültigen Ungleichung $0 \leq -1$ führt.

(b) $\mathcal{D} = \emptyset$, da

Es kann also der Fall auftreten, dass sowohl (P) als auch (D) unzulässig ist.

Aus den Dualitätsstzen (oder alternativ aus den KKT-Bedingungen) ergibt sich unmittelbar der folgende Satz zur Charakterisierung von Optimallösungen:

Satz 4.17 (Satz vom schwachen komplementären Schlupf) *Betrachte das Paar (P) und (D) aus primalem und dualem Problem wie in Satz 4.12. Die Vektoren $\bar{x}$ und $\bar{y}$ seien zulässig für (P) bzw. (D). Dann sind folgende Aussagen äquivalent:*

(a) $\bar{x}$ *ist optimal für (P) und* $\bar{y}$ *ist optimal für (D).*

(b) $(b - A\bar{x})^T\bar{y} = 0.$

(c) $\bar{y}_i = 0$ *oder* $b_i - A_{i\cdot}\bar{x} = 0, \quad i = 1, \ldots, m.$

Beweis. Wir stellen zunächst fest, dass bei primal-dualer Zulässigkeit (b) und (c) äquivalent sind.
Gilt (b) dann stimmen nach dem schwachen Dualitätssatz 4.1 und Korollar 4.2 primaler und dualer Zielfunktionswert überein und $\bar{x}$, $\bar{y}$ sind optimal für (P) bzw. (D), also gilt (a).
Gilt (a), dann existieren nach dem starken Dualitätssatz 4.13 Optimallösungen von (P) und (D) mit gleichem Zielfunktionswert. Damit haben auch $\bar{x}$ und $\bar{y}$ als Optimallösungen gleichen Zielfunktionswert, es gilt also $0 = c^T\bar{x} - (-b^T\bar{y}) = (b - A\bar{x})^T\bar{y}$ und dies ist (b). □

Bemerkung 4.18 *In (c) können beide Komponenten gleichzeitig verschwinden. Es gibt jedoch immer Paare von Optimallösungen, bei denen immer genau eine der Komponenten* $b_i - A_{i\cdot}\bar{x}$ *oder* $\bar{y}_i$ *verschwindet. Man spricht dann von strikter Komplementarität.*

Merkregel: Die Aussage des Satzes vom komplementären Schlupf ist folgende: Ein primal-dual zulässiges Paar ist genau dann optimal, wenn für jede nicht-bindende Nichtnegativitätsbedingung die komplementäre Ungleichung bindend ist.
Transformation der allgemeinen Form (PA) auf (P) wie in Satz 4.9 liefert folgendes Resultat:

Satz 4.19 (Satz vom schwachen komplementären Schlupf) *Gegeben seien dimensionsverträgliche Matrizen A, B, C, D und Vektoren a, b, c, d. Betrachte das zueinander gehörende Paar dualer Probleme*

$$(PA) \qquad \begin{array}{ll} \min & c^T x + d^T y \\ \text{s.t.} & Ax + By \geq a \\ & Cx + Dy = b \\ & x \geq 0 \end{array}$$

und

$$(DA) \qquad \begin{array}{ll} \max & a^T u + b^T v \\ \text{s.t.} & A^T u + C^T v \leq c \\ & B^T u + D^T v = d \\ & u \geq 0 \end{array}$$

Die Vektoren $\binom{\bar{x}}{\bar{y}}$ *und* $\binom{\bar{u}}{\bar{v}}$ *seien zulässig für* (PA) *bzw.* (DA). *Dann sind folgende Aussagen äquivalent:*

(a) $\binom{\bar{x}}{\bar{y}}$ *ist optimal für* (P) *und* $\binom{\bar{u}}{\bar{v}}$ *ist optimal für* (D).

(b) $\big(c - (A^T\bar{u} + C^T\bar{v})\big)^T \bar{x} - \bar{u}^T\big(a - (A\bar{x} + B\bar{y})\big) = 0.$

(c) Für alle Komponenten $\bar{u}_i$ *der Duallösung gilt:*

$$\bar{u}_i = 0 \quad \textit{oder} \quad (A\,\bar{x} + B\,\bar{y})_i = a_i.$$

Für alle Komponenten $\bar{x}_j$ *der Primallösung gilt:*

$$\bar{x}_j = 0 \quad \textit{oder} \quad (A^T\bar{u} + C^T\bar{v})_j = c_j.$$

Beweis. Wie in Satz 4.9 kann man (PA) in der Form (P) schreiben, siehe (4.9), das duale Problem (4.10) kann dann mit $v = w - s$ und $c - A^T u - C^T v = z$ zu (DA) vereinfacht werden. Wendet man nun Satz 4.17 auf (4.9), (4.10) an und nutzt $Cx + Dy = b$, $c - A^T u - C^T v = z$, dann ist die Optimalität nach Satz 4.9 (b) äquivalent zu

$$\begin{aligned} 0 &= (\bar{u}^T, \bar{s}^T, \bar{w}^T, \bar{z}^T) \left(\begin{pmatrix} -a \\ b \\ -b \\ 0 \end{pmatrix} - \begin{pmatrix} -A & -B \\ C & D \\ -C & -D \\ -I & 0 \end{pmatrix} \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} \right) \\ &= \bar{u}^T(A\bar{x} + B\bar{y} - a) + (c - A^T\bar{u} - C^T\bar{v})^T \bar{x} \end{aligned}$$

und da alle Summanden dasselbe Vorzeichen haben, ist dies wiederum äquivalent zu (c). □

Satz 4.20 (Satz vom starken komplementären Schlupf, Satz von Goldman-Tucker)
Betrachte das Paar dualer linearer Probleme

$$(P) \quad \begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax \le b \end{array} \qquad \textit{und} \qquad (D) \quad \begin{array}{ll} \max & -b^T y \\ \text{s.t.} & A^T y = -c \\ & y \ge 0 \end{array}$$

Besitzen beide Probleme zulässige Lösungen, so existieren optimale Lösungen $\bar{x}, \bar{y}$, *so dass für alle Zeilenindizes gilt:*

$$\begin{array}{ccc} \bar{y}_i > 0 & \Longleftrightarrow & A_{i\cdot}\bar{x} = b_i, \\ \bar{y}_i = 0 & \Longleftrightarrow & A_{i\cdot}\bar{x} < b_i\,. \end{array}$$

Beweis. (Für Interessierte) Definiere die Mengen

$$\begin{aligned} N &:= \{i \,:\, \text{Es existiert Optimallösung } x \text{ von (P) mit } A_{i\cdot}x < b_i\}, \\ B &:= \{i \,:\, \text{Es existiert Optimallösung } y \text{ von (D) mit } y_i > 0\}. \end{aligned}$$

Dann existieren Optimallösungen $\bar{x}, \bar{y}$ mit

$$A_{N\cdot}\bar{x} < b_N, \quad \bar{y}_B > 0 \tag{4.14}$$

$$A_{B\cdot}\bar{x} = b_B, \quad \bar{y}_N = 0 \tag{4.15}$$

und es gilt $B \cap N = \emptyset$.
Tatsächlich existiert zu jedem $i \in N$ eine Optimallösung $\bar{x}^{(i)}$ von (P) mit $A_{i\cdot}\bar{x}^{(i)} < b_i$ und zu jedem $i \in B$ eine Optimallösung $\bar{y}^{(i)}$ von (D) mit $\bar{y}_i^{(i)} > 0$. Die Konvexkombinationen $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} \bar{x}^{(i)}$ sowie $\bar{y} = \frac{1}{m}\sum_{i=1}^{m} \bar{y}^{(i)}$ sind dann wieder Optimallösungen und erfüllen (4.14). Nun folgt (4.15) aus Satz 4.19 (c) und zusammen mit (4.14) ergibt sich $B \cap N = \emptyset$.
Wir müssen also nur noch zeigen, dass $B \cup N = \{1, \dots, m\}$.
Dazu setze $J = \{1, \dots, m\} \setminus (B \cup N)$. Annahme, $J \neq \emptyset$. Dann gilt $\bar{y}_J = 0$, $(b - A\bar{y})_J = 0$ und es existiert $i \in J$. Wir zeigen, dass das System

$$A_{J\setminus\{i\}\cdot}s \le 0, \quad A_{B\cdot}s = 0, \quad A_{i\cdot}s < 0$$

keine Lösung haben kann: für $t > 0$ klein genug gilt sonst mit (4.14), (4.15)

$$A_{B\cdot}(\bar{x} + ts) = b_B, \quad A(\bar{x} + ts) \leq b, \quad A_{i\cdot}(\bar{x} + ts) < b_i,$$

$\bar{x} + ts$ ist also primal zulässig und auch optimal, da $\bar{x} + ts$ und $\bar{y}$ wegen $\bar{y}_{J\cup N} = 0$ die Komplementaritätsbedingung erfüllen. Es wäre also $i \in N$, was $i \in J$ widerspricht.
Nach Folgerung 4.8 zum Farkas-Lemma muss also das System (setze $C^T = -A_{J\setminus\{i\}\cdot}$, $D^T = -A_{B\cdot}$, $b^T = A_{i\cdot}$)

$$-A^T_{J\setminus\{i\}\cdot} v - A^T_{B\cdot} w = A^T_{i\cdot}, \quad v \geq 0 \tag{4.16}$$

eine Lösung haben. Definieren wir $u \in \mathbb{R}^m$ durch

$$u_{J\setminus\{i\}} = v, \quad u_i = 1, \quad u_B = w, \quad u_N = 0,$$

dann ergibt (4.16)

$$A^T u = 0, \quad u_{J\cup N} \geq 0, \quad u_i > 0.$$

Für $t > 0$ klein genug gilt daher mit (4.15)

$$A^T(\bar{y} + tu) = -c^T, \quad \bar{y} + tu \geq 0, \qquad (\bar{y} + tu)_N = 0, \quad (\bar{y} + tu)_i > 0,$$

$\bar{y} + tu$ ist also dual zulässig und auch optimal, da $\bar{x}$ und $\bar{y} + tu$ wegen $(b - A\bar{x})_{J\cup B} = 0$ die Komplementaritätsbedingung erfüllen. Es wäre also $i \in B$, was $i \in J$ widerspricht. Die Annahme $J \neq \emptyset$ war demnach falsch und der Beweis ist beendet. $\square$

Bemerkung 4.21 *Der Satz vom starken komplementären Schlupf gilt entsprechend für das Paar dualer linearer Probleme aus Definition 4.9.*

Kapitel 5

Der Simplex-Algorithmus

In diesem Kapitel wollen wir das erste Verfahren zur Lösung linearer Probleme behandeln. Wir betrachten dazu LPs in der sogenannten Standardform:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0. \end{aligned} \tag{5.1}$$

Wir nehmen $n \geq m$ an. Außerdem wollen wir in diesem Kapitel zunächst annehmen, dass die Matrix $A \in \mathbb{R}^{m \times n}$ vollen Zeilenrang hat, d.h. $\operatorname{rang}(A) = m$. Das zu (5.1) duale Problem lautet:

$$\begin{aligned} \max \quad & b^T y \\ \text{s.t.} \quad & A^T y \leq c \end{aligned} \tag{5.2}$$

In (dualer) Standardform hat das zu (5.1) gehörige duale LP die Gestalt:

$$\begin{aligned} \max \quad & b^T y \\ \text{s.t.} \quad & A^T y + z = c \\ & z \geq 0 \end{aligned} \tag{5.3}$$

5.1 Basen, Basislösungen und Degeneriertheit

Aus Korollar 2.40 wissen wir, dass eine Optimallösung, sofern sie existiert, immer an einer Ecke angenommen wird. Die zulässige Menge eines LPs in Standardform ist ein Polyeder $\mathcal{P}^=(A, b)$. Wir müssen also die Ecken eines solchen Polyeders auf effiziente Weise beschreiben.

Eine Ecke von $\mathcal{P}^{=}(A,b)$ ist durch n linear unabhängige Nebenbedingungen festgelegt. Da die m Gleichungsnebenbedingungen auch in einer Ecke erfüllt sein müssen, verbleibt also, $n-m$ der n Ungleichungen $x_i \geq 0$ so auszuwählen, dass alle Bedingungen linear unabhängig sind und die zugehörige Ecke in $\mathcal{P}^{=}(A,b)$ liegt.
Sei N die Menge der $n-m$ ausgewählten Indizes und $B = \{1,2,\ldots,n\} \setminus N$. Die Forderung, dass die Bedingungen $Ax = b$ und $x_j = 0$ für $j \in N$ linear unabhängig sind, sagt aus, dass das System $Ax = b$ eindeutig lösbar sein muss, wenn man $x_j = 0$ setzt für alle $j \in N$, d.h. $A_{\cdot B}$ muss invertierbar sein.

Definition 5.1 *Gegeben sei ein LP in Standardform (5.1). Die Matrix $A \in \mathbb{R}^{m\times n}$ habe vollen Zeilenrang. Sei B ein m-Tupel mit paarweise verschiedenen Indizes aus $\{1,2,\ldots,n\}$ und sei N das $(n-m)$-Tupel der verbleibenden Indizes.*

(a) *B heißt* Basis *von (5.1), falls $A_{\cdot B}$ regulär ist.*

B heißt primal zulässig*, falls $A_{\cdot B}^{-1} b \geq 0$ ist.*

B heißt dual zulässig*, falls $c_N - A_{\cdot N}^T A_{\cdot B}^{-T} c_B \geq 0$ ist.*

Wir verwenden ab jetzt die Schreibweise A_B für $A_{\cdot B}$ bzw. A_N für $A_{\cdot N}$.

Beachte, dass die Notation $A_{\cdot I}$ nur für Mengen I definiert ist. Wir erweitern diese Definition in diesem Zusammenhang auch auf Basen und Tupel und interpretieren A_B als die Matrix, die aus der Menge der durch B induzierten Spalten besteht.

(b) *Die zu B gehörige Matrix A_B heißt* Basismatrix*, A_N* Nichtbasismatrix*.*

Die zu B gehörende Basislösung *hat die Komponenten $x_B = A_B^{-1} b$, $x_N = 0$. Die Variablen x_j, $(j \in B)$ heißen* Basisvariablen*, alle anderen* Nichtbasisvariablen*.*

(c) *Eine Basislösung x zur Basis B heißt* nicht degeneriert *(auch:* nicht entartet*), falls $x_B = A_B^{-1} b > 0$, andernfalls heißt sie* degeneriert *(auch:* entartet*).*

Eine Basislösung x heißt zulässig*, falls B primal zulässig ist.*

Eine Bestätigung des Zusammenhanges zwischen Basen und Ecken von $\mathcal{P}^{=}(A,b)$ gibt der folgende Satz.

Satz 5.2 *Sei $\mathcal{P} = \mathcal{P}^=(A, b)$ ein Polyeder mit $\operatorname{rang}(A) = m \leq n$, und sei $x \in \mathcal{P}$. Dann sind folgende Aussagen äquivalent:*

(i) x ist eine Ecke von $\mathcal{P}$.

(ii) x ist eine zulässige Basislösung (d.h. es existiert eine primal zulässige Basis B mit $x_B = A_B^{-1}b$, $x_N = 0$).

Beweis. $\boxed{\text{(i)} \Longrightarrow \text{(ii):}}$ Sei $I = \operatorname{supp}(x)$. Ist x eine Ecke von $\mathcal{P}$, sind die Spaltenvektoren $A_{\cdot j}$, $j \in I$ nach Satz 3.20 linear unabhängig.
Da $\operatorname{rang}(A) = m$ gilt, existiert eine Menge $J \subseteq \{1, 2, \ldots, n\} \setminus I$ mit $|J| = m - |I|$, so dass $A_{\cdot(I \cup J)}$ regulär ist. Setze nun $B = I \cup J$. Dann ist B eine Basis und es gilt:

$$
\begin{aligned}
x &= \begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} x_B \\ 0 \end{pmatrix} \qquad \text{(Beachte: } \operatorname{supp}(x) = I \subset B) \\
A_B^{-1}b &= A_B^{-1}Ax = A_B^{-1}\left([A_B, A_N]\begin{pmatrix} x_B \\ 0 \end{pmatrix}\right) = [I, A_B^{-1}A_N]\begin{pmatrix} x_B \\ 0 \end{pmatrix} = x_B.
\end{aligned}
$$

Da $x \geq 0$ gilt, ist B damit primal zulässig und x ist eine zulässige Basislösung.

$\boxed{\text{(ii)} \Longrightarrow \text{(i):}}$ folgt direkt aus Satz 3.20. □

Damit haben wir gezeigt, dass die Ecken von $\mathcal{P}^=(A, b)$ den Basislösungen von $Ax = b$, $x \geq 0$ entsprechen. Also gibt es zu jeder Ecke eine zulässige Basis von A. Sind zwei Ecken verschieden, so sind natürlich auch die entsprechenden Basen verschieden. Dies gilt aber nicht in umgekehrter Richtung!

$$\boxed{\text{Ecke} \overset{\text{eindeutig}}{\longleftrightarrow} \text{zulässige Basislösung} \overset{\text{nicht eindeutig}}{\longleftrightarrow} \text{zulässige Basis}}$$

Beispiel 5.3 (Degeneriertheit)

(a) Betrachte das System

$$
\begin{aligned}
x_1 + x_2 &\leq 1 \\
2x_1 + x_2 &\leq 2 \\
x_1, x_2 &\geq 0.
\end{aligned}
$$

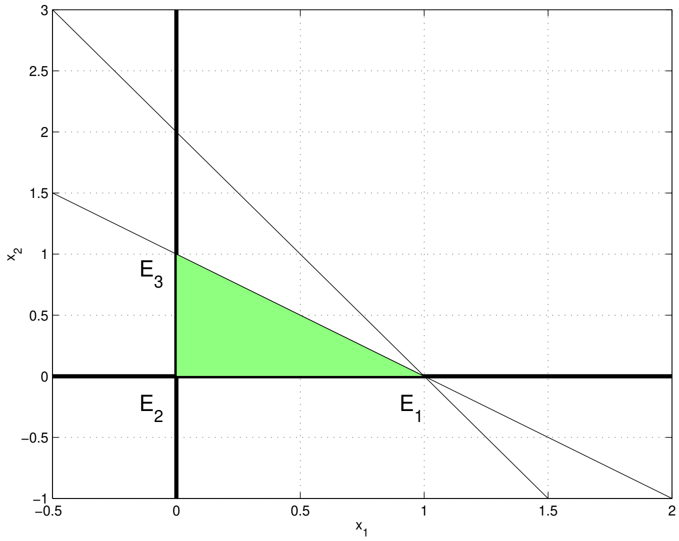

Abbildung 5.1: zum Beispiel 5.3 (a)

Das Polytop hat die drei Ecken E_1, E_2, E_3. Das System lautet in Standardform:

$$\begin{array}{rcrcrcrcl} x_1 & + & x_2 & + & s_1 & & & = & 1 \\ 2x_1 & + & x_2 & & & + & s_2 & = & 2 \\ & & & & & x_1, x_2, & s_1, s_2 & \geq & 0. \end{array}$$

Zur Ecke $(1,0)^T$ bzw. $(1,0,0,0)^T$ gehören folgende Basen:

B	A_B	x_B	x_N
$(1,2)$	$\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$	$(x_1, x_2) = (1,0)$	(s_1, s_2)
$(1,3)$	$\begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$	$(x_1, s_1) = (1,0)$	(x_2, s_2)
$(1,4)$	$\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$	$(x_1, s_2) = (1,0)$	(x_2, s_1)

Die Degeneriertheit der Basislösung resultiert aus der Redundanz der zweiten Ungleichung. Ohne diese träte hier keine Degeneriertheit auf.

Im $\mathbb{R}^2$ gibt es keine „richtige“ Degeneriertheit, d.h. jedes nicht redundante Ungleichungssystem im $\mathbb{R}^2$ impliziert, dass keine degenerierten Basislösungen existieren.

(b) *Im $\mathbb{R}^3$ tritt dagegen echte Degeneriertheit auf (Abb. 5.2).*

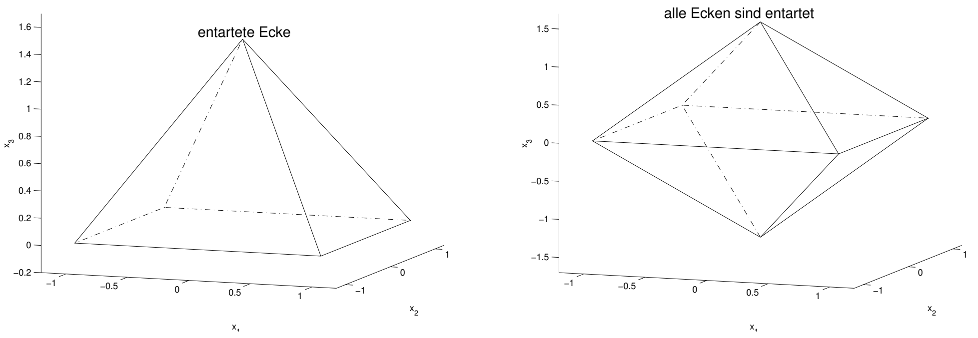

Abbildung 5.2: Degeneriertheit im $\mathbb{R}^3$

Bemerkung 5.4 *Ist x eine nichtdegenerierte Basislösung von $Ay = b$, $y \geq 0$, dann gibt es eine eindeutige zu x gehörende Basis B mit $x_B = A_B^{-1}b > 0$ und $x_N = 0$.*
Die Umkehrung gilt im Allgemeinen nicht.

Beispiel 5.5 *Sei $\mathcal{P}^=(A, b)$ gegeben durch*

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Es gibt zwei Basen $B_1 = (1, 3)$ und $B_2 = (2, 3)$ mit $x_1^T = (1, 0, 0)$ und $x_2^T = (0, 1, 0)$. Beide Basislösungen sind degeneriert, obwohl die zugehörigen Basen eindeutig sind.

5.2 Die Grundversion des Simplex-Verfahrens

Wie bereits mehrfach angedeutet, besteht die Grundidee des Simplexverfahrens darin, sich auf geschickte Weise von Ecke zu Ecke zu bewegen, bis Optimalität nachgewiesen werden kann. In diesem Abschnitt zeigen wir, wie

dieses Verfahren funktioniert und gehen davon aus, dass wir eine (primal) zulässige Basis bereits gegeben haben. Wie wir eine solche bekommen, ist Thema eines späteren Abschnitts.
Sei also $B \subseteq \{1, 2, \ldots, n\}$, $|B| = m$ mit A_B regulär und $A_B^{-1}b \geq 0$ gegeben. Die zugehörige Basislösung ist

$$x_B = A_B^{-1}b \geq 0, \qquad x_N = 0.$$

Eine Änderung dieser Lösung (x_B, x_N) ist nur dadurch erreichbar, dass eine Nichtbasisvariable erhöht wird. Die Auswirkungen auf die anderen Variablen lassen sich an folgender Beziehung ablesen:

$$\begin{aligned} Ax = b &\iff A_B x_B + A_N x_N = b \\ &\iff A_B x_B = b - A_N x_N \\ &\iff x_B = A_B^{-1}(b - A_N x_N). \end{aligned} \tag{5.4}$$

Die Auswirkung einer Erhöhung von $x_j\,(j \in N)$ auf die Zielfunktion lässt sich ebenfalls berechnen:

$$\begin{aligned} c^T x &= c_B^T x_B + c_N^T x_N \\ &\overset{(5.4)}{=} c_B^T A_B^{-1}(b - A_N x_N) + c_N^T x_N \\ &= c_B^T A_B^{-1} b + (c_N^T - c_B^T A_B^{-1} A_N) x_N. \end{aligned} \tag{5.5}$$

Eine Erhöhung eines x_j mit $j \in N$ bringt also nur dann eine mögliche Verbesserung der Zielfunktion, falls

$$c_j - c_B^T A_B^{-1} A_{\cdot j} < 0 \tag{5.6}$$

gilt für ein $j \in N$. Der Vektor $z_N = c_N - A_N^T A_B^{-T} c_B$ heißt Vektor der reduzierten Kosten (oder einfach nur reduzierte Kosten).
Wir werden im folgenden Satz zeigen, dass (x_B, x_N) optimal ist, falls die reduzierten Kosten nicht-negativ sind.
Nehmen wir also an, wir haben einen Index $j \in N$, der (5.6) erfüllt. Wir wollen x_j nun soweit wie möglich erhöhen, d.h. der resultierende Vektor soll gerade noch zulässig bleiben. Auskunft über den Einfluss von x_j auf die anderen Variablen gibt (5.4):
Alle Indizes in $N \setminus \{j\}$ bleiben von der Aktion unberührt. Wegen (5.4) gilt:

$$x_B = A_B^{-1}b - A_B^{-1}A_N x_N = A_B^{-1}b - A_B^{-1}A_{\cdot j}x_j - A_B^{-1}A_{N\setminus\{j\}}\underbrace{x_{N\setminus\{j\}}}_{=0}.$$

Der Einfluss der Erhöhung von x_j auf x_B wird also durch den Vektor

$$w = A_B^{-1} A_{\cdot j}$$

gesteuert. Gilt $w \leq 0$, so können wir x_j beliebig erhöhen und bleiben immer zulässig (d.h. $x \geq 0$). Da zusätzlich (5.6) gilt, wird dadurch auch die Zielfunktion beliebig verbessert, d.h. das LP (5.1) ist unbeschränkt.

Nehmen wir deshalb $w \not\leq 0$ an und betrachten diejenigen Indizes mit $w_i > 0$. Wir können x_j maximal so weit erhöhen, bis eine der Variablen aus x_B die Null erreicht. Damit x zulässig bleibt, muss $x_B \geq 0$ bleiben:

$$x_B = A_B^{-1} b - x_j w \geq 0.$$

Für alle Indizes $i \in \{1, \ldots, m\}$ mit $w_i > 0$ muss also gelten

$$x_j w_i \leq (A_B^{-1} b)_i \qquad \Longleftrightarrow \qquad x_j \leq \frac{(A_B^{-1} b)_i}{w_i}.$$

Die Erhöhung γ von x_j lässt sich daher in folgender Form schreiben:

$$\gamma = \min \left\{ \frac{(A_B^{-1} b)_i}{w_i} \ : \ w_i > 0,\, i \in \{1, \ldots, m\} \right\}.$$

Genau eine Variable, die als erste den Wert Null erreicht, wird nun die Basis verlassen (sie ist an der unteren Schranke angekommen) und x_j wird in die Basis aufgenommen. Danach beginnt dieser Prozess von vorne. Zusammengefasst sieht das Verfahren wie folgt aus:

Algorithmus 5.6 Der Simplex-Algorithmus, Grundversion.

Input: *Eine primal zulässige Basis* B, $\bar{x}_B = A_B^{-1}b$.

Output: (i) *Eine Optimallösung* $\bar{x}$ *für das LP (5.1)*
oder
(ii) *Die Meldung „Das LP (5.1) ist unbeschränkt".*

(1) BTRAN *(Backward Transformation)*
Löse $\bar{y}^T A_B = c_B^T$.

(2) Pricing
Berechne Vektor der reduzierten Kosten: $\bar{z}_N = c_N - A_N^T \bar{y}$.

Falls $\bar{z}_N \geq 0$, *dann ist* B *optimal,* **Stop**.

Andernfalls wähle ein $j \in N$ *mit* $\bar{z}_j < 0$. x_j *heißt die* in die Basis eintretende Variable (engl. entering).

(3) FTRAN *(Forward Transformation)*
Löse $A_B w = A_{\cdot j}$.

(4) Ratio-Test
Falls $w \leq 0$, *dann ist das LP (5.1) unbeschränkt,* **Stop**.

Andernfalls berechne

$$\gamma = \frac{\bar{x}_{B_i}}{w_i} = \min\left\{ \frac{\bar{x}_{B_k}}{w_k} \ : \ w_k > 0, \ k = 1, 2, \dots, m \right\},$$

wobei $i \in \{1, 2, \dots, m\}$ *und* $w_i > 0$ *ist.* $\bar{x}_{B_i}$ *heißt die* die Basis verlassende Variable (engl. leaving).

(5) Update
Setze

$$\begin{aligned} \bar{x}_B &:= \bar{x}_B - \gamma w, \\ N &:= (N \setminus \{j\}) \cup \{B_i\}, \\ B_i &:= j, \\ \bar{x}_j &:= \gamma. \end{aligned}$$

Gehe zu (1).

Beispiel 5.7 *Betrachten wir folgendes lineare Problem:*

$$\begin{array}{llllll}
\min & -3x_1 & -2x_2 & -2x_3 & & \\
\text{s.t.} & x_1 & & +x_3 & \leq & 8 \\
 & x_1 & +x_2 & & \leq & 7 \\
 & x_1 & +2x_2 & & \leq & 12 \\
 & & & x_1, x_2, x_3 & \geq & 0.
\end{array}$$

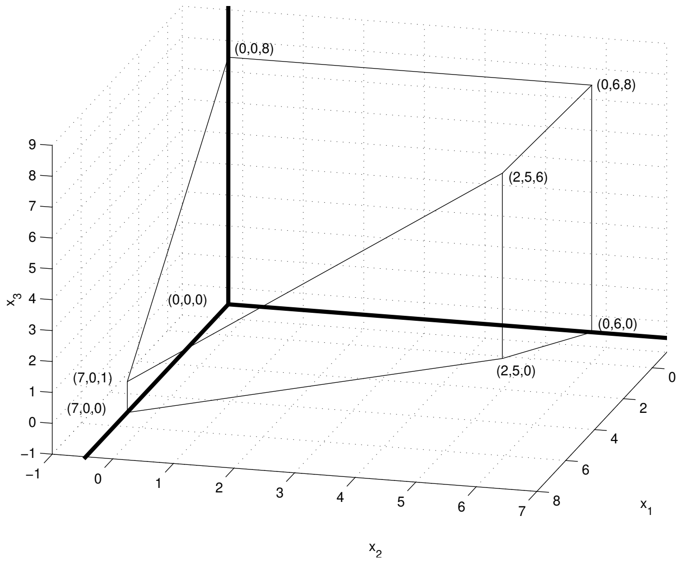

Abbildung 5.3: Grafische Darstellung zu Beispiel 5.7.

Wir bringen das LP in Standardform:

$$\begin{array}{lllllllll}
\min & -3x_1 & -2x_2 & -2x_3 & & & & & \\
\text{s.t.} & x_1 & & +x_3 & +x_4 & & & = & 8 \\
 & x_1 & +x_2 & & & +x_5 & & = & 7 \\
 & x_1 & +2x_2 & & & & +x_6 & = & 12 \\
 & & & & & & x_1, \dots, x_6 & \geq & 0.
\end{array}$$

Eine primal zulässige Basis ist $B = (4, 5, 6)$, $N = (1, 2, 3)$.
Die Basislösung ist $\bar{x}_B = (8, 7, 12)^T$, $\bar{x}_1, \bar{x}_2, \bar{x}_3 = 0$, d.h. alle Schlupfvariablen sind in der Basis und der Zielfunktionswert ist Null. Die einzelnen Schritte des Algorithmus 5.6 sehen dann wie folgt aus:

(1) BTRAN

$$\bar{y}^T \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = (0,0,0) \qquad \Longrightarrow \qquad \bar{y}^T = (0,0,0)^T$$

(2) Pricing

$$\bar{z}_N = \begin{pmatrix} -3 \\ -2 \\ -2 \end{pmatrix} - \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -3 \\ -2 \\ -2 \end{pmatrix}$$

Wähle nun $j = 1$. Die Variable x_1 wird in die Basis aufgenommen.

(3) FTRAN

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \Longrightarrow \qquad w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

(4) Ratio-Test

$$\gamma = \min \left\{ \frac{8}{1}, \frac{7}{1}, \frac{12}{1} \right\} = 7$$

mit $i = 2$ und $B_2 = 5$. x_5 verlässt die Basis.

(5) Update

$$\bar{x}_B = \begin{pmatrix} \bar{x}_4 \\ \bar{x}_5 \\ \bar{x}_6 \end{pmatrix} = \begin{pmatrix} 8 \\ 7 \\ 12 \end{pmatrix} - 7 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix}$$

Daher: $N = \{1,2,3\} \setminus \{1\} \cup \{5\} = (5,2,3)$; $B = (4,1,6)$, $B_2 = 1$, $\bar{x}_1 = 7$. Der Zielfunktionswert ist -21.

Entsprechende Fortsetzung liefert folgende Ergebnisse:

Iteration	*Zielfunktionswert*	*Eintretende Var.*	*Verlassende Var.*
1	-21	x_1	x_5
2	-23	x_3	x_4
3	-28	x_2	x_6

Nach der dritten Iteration gilt $B = (3,1,2), N = \{4,5,6\}$ *und*

$$\bar{x}_B = \begin{pmatrix} \bar{x}_3 \\ \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 5 \end{pmatrix}.$$

(1) BTRAN

$$\bar{y}^T \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} = (-2,-3,-2) \qquad \Longrightarrow \qquad \bar{y} = \begin{pmatrix} -2 \\ 0 \\ -1 \end{pmatrix}$$

(2) Pricing

$$\bar{z}_N = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \geq 0$$

$\Longrightarrow$ *B ist optimal (vgl. dazu den nachfolgenden Satz).*

Satz 5.8
Terminiert der Algorithmus 5.6, so liefert er das korrekte Ergebnis.

Beweis. Wir zeigen zuerst, dass der fünfte Schritt (Update) korrekt ist. Dazu sei B die Basis vor dem Update und B^+ das Tupel der Indizes nach dem Update, d.h. B und B^+ sind identisch, außer an der Stelle i, an welcher B_i durch j ersetzt wird. Analog sei $N^+ = \{1,\ldots,n\} \backslash N$. Der Übersichtlichkeit halber bezeichnen wir den neuen Punkt mit $\bar{x}^+$. Wir müssen also zeigen:

(i) A_{B^+} ist regulär.

(ii) $\bar{x}^+_{B^+} \geq 0\,,\ \bar{x}^+_{N^+} = 0$.

(iii) $\bar{x}^+_{B^+} = A^{-1}_{B^+} b$.

$\boxed{\text{zu (i):}}$ Angenommen, A_{B^+} ist nicht regulär. Dann ist $A_{\cdot j}$ linear abhängig von den restlichen Spalten von A_{B^+}, d.h. es existiert ein eindeutiges λ mit

$$A_{\cdot B^+ \backslash \{j\}} \lambda = A_{\cdot j}.$$

Wegen FTRAN gilt $A_{\cdot j} = A_B w$, und da $B^+ \setminus \{j\} = B \setminus \{B_i\}$ erhalten wir

$$A_{\cdot B\setminus\{B_i\}}\lambda = A_B w,$$

oder gleichbedeutend

$$A_{\cdot B\setminus\{B_i\}}\lambda = w_i A_{\cdot B_i} + \sum_{k\in\{1,\dots,m\}\setminus\{i\}} w_k A_{\cdot B_k}.$$

Da $A_{\cdot B_i}$ linear unabhängig von $A_{\cdot B\setminus\{B_i\}}$ (weil A_B regulär laut Voraussetzung), folgt daraus, dass $w_i = 0$ gelten muss, ein Widerspruch zur Wahl im Ratio-Test.

Eigentlich folgen (ii) und (iii) bereits aus unserer Herleitung des Simplex-Algorithmus. Der Vollständigkeit halber geben wir dennoch den Beweis.

$\boxed{\text{zu (ii):}}$ Da $\bar{x}_j = \gamma$ und $\gamma > 0$ gilt, folgt $\bar{x}^+_{B_i^+} = \gamma \geq 0$. Weiter gilt für alle $k \in \{1, 2, \dots, m\}$

$$\bar{x}^+_{B_k} = \bar{x}_{B_k} - \gamma w_k.$$

Ist $w_k \leq 0$, dann gilt $\bar{x}^+_{B_k} \geq 0$, da $\bar{x}_{B_k} \geq 0$.

Andernfalls gilt

$$\bar{x}^+_{B_k} = \bar{x}_{B_k} - \frac{\bar{x}_{B_i}}{w_i} w_k \geq \bar{x}_{B_k} - \frac{\bar{x}_{B_k}}{w_k} w_k = 0.$$

Insbesondere gilt dann also auch $\bar{x}^+_{B_i} = 0$, d.h. $\bar{x}^+_{N^+} = 0$, da die restlichen Variablen mit den Indizes aus $N \setminus \{j\}$ unberührt bleiben.

$\boxed{\text{zu (iii):}}$ Es genügt, $A_{B^+}\bar{x}^+_{B^+} = b$ zu zeigen. Es gilt

$$\begin{aligned}
A_{B^+}\,\bar{x}^+_{B^+} &= A_B\,\bar{x}^+_B + A_{\cdot j}\bar{x}^+_j - A_{\cdot B_i}\,\underbrace{\bar{x}^+_{B_i}}_{=0} \\
&= A_B(\bar{x}_B - \gamma w) + A_{\cdot j}\bar{x}^+_j \;=\; b - \gamma A_{\cdot j} + A_{\cdot j}\,\underbrace{\bar{x}^+_j}_{=\gamma} \\
&= b
\end{aligned}$$

Aus (i) – (iii) folgt, dass B während des gesamten Algorithmus eine primal zulässige Basis ist, und $\bar{x}$ mit $\bar{x}_B = A_B^{-1} b \geq 0$ und $\bar{x}_N = 0$ ist eine zulässige Lösung.

Nehmen wir jetzt an, dass der Algorithmus im zweiten Schritt (Pricing) stoppt. Gegeben ist dann $\bar{x}_B$ und $\bar{y}$ aus dem ersten Schritt (BTRAN). Wir wissen bereits, dass $\bar{x}$ für (5.1) zulässig ist.

Wir zeigen nun, dass $\bar{y}$ für (5.2) zulässig ist und $\bar{x}, \bar{y}$ die Bedingung (c) aus dem Satz 4.19 vom schwachen komplementären Schlupf erfüllt. Jedenfalls ist $\bar{y}$ für (5.2) zulässig, da nach BTRAN gilt $A_B^T \bar{y} = c_B$ und zudem $c_N - A_N^T \bar{y} = \bar{z}_N \geq 0$ ist. Dies zeigt $A^T \bar{y} \leq c$.
Zudem gilt $\bar{x}_N = 0$ und $A_B^T \bar{y} = c_B$, also ist die Bedingung (c) aus Satz 4.19 erfüllt und somit $\bar{x}$ optimal für (5.1). (Zur Kontrolle: $c^T \bar{x} = c_B^T \bar{x}_B = (\bar{y}^T A_B) x_B = \bar{y}^T b$.)

Nehmen wir schließlich an, dass der Algorithmus 5.6 im vierten Schritt (Ratio-Test) stoppt, d.h. es gilt $w \leq 0$. Wir wissen aus (5.4) und FTRAN, dass dann die Punkte $\bar{x}^+(\gamma)$ gegeben durch

$$\bar{x}_j^+(\gamma) = \gamma, \quad \bar{x}_B^+(\gamma) = \bar{x}_B - A_B^{-1} A_{\cdot j} \gamma = \bar{x}_B - \gamma w, \quad x_{N \setminus \{j\}}^+(\gamma) = 0$$

für jedes $\gamma > 0$ primal zulässig sind. Für den Zielfunktionswert ergibt sich nach (5.5) und BTRAN

$$c^T \bar{x}^+(\gamma) = c_B^T \bar{x}_B + \bar{z}_N^T \bar{x}_N^+(\gamma) = c_B^T \bar{x}_B + \underbrace{\bar{z}_j}_{<0} \gamma \to -\infty \quad \text{für } \gamma \to \infty.$$

Ist also $w \leq 0$ in Schritt 3, dann ist (5.1) tatsächlich nicht nach unten beschränkt. □

Es bleibt zu klären, ob der Algorithmus 5.6 endlich ist. Dazu machen wir zunächst eine einfache Beobachtung.

Satz 5.9 *Betrachte ein lineares Problem (5.1), so dass $\mathcal{P}^=(A, b) \neq \emptyset$ und alle zulässigen Basislösungen nicht degeneriert sind. Dann ist der Algorithmus 5.6 endlich.*

Beweis. Der Algortihmus generiert lauter primal zulässige Basen, jede ist nach Voraussetzung nichtdegeneriert. Sei B die aktuelle Basis. Es gilt dann $\bar{x}_B > 0$, $\bar{x}_N = 0$ (Nichtdegeneriertheit). Entweder terminiert der Algorithmus in der aktuellen Iteration, oder es gilt $\bar{z}_N \not\geq 0$ und $w \not\leq 0$. Er wählt dann $j \in N$ mit $\bar{z}_j < 0$ und bestimmt

$$\gamma = \frac{\bar{x}_{B_i}}{w_i} = \min\left\{\frac{\bar{x}_{B_k}}{w_k} \;:\; w_k > 0,\ k = 1, 2, \ldots, m\right\},$$

wobei $i \in \{1, 2, \ldots, m\}$ und $w_i > 0$ ist. Wegen $\bar{x}_B > 0$ folgt $\gamma > 0$. Der neue Zielfunktionswert ist nun nach (5.5) und BTRAN

$$c^T\bar{x}^+ = c_B^T\bar{x}_B + \underbrace{\bar{z}_j}_{<0}\underbrace{\gamma}_{>0} < c_B^T\bar{x}_B = c^T\bar{x}.$$

Ist nun $B_1, B_2, \ldots$ eine Folge von Basen, die im Algorithmus 5.6 erzeugt werden, so folgt daraus, dass keine Basis innerhalb der Folge doppelt auftreten kann, da die Zielfunktionswerte streng monoton fallen. Da es nur endlich viele Basen gibt, muss der Algorithmus irgendwann terminieren. □

Leider ist die Grundversion des Simplex-Algorithmus im degenerierten Fall im Allgemeinen nicht endlich.

Beispiel 5.10 (Kreiseln) *Siehe Übung.*

Es bleibt die Frage zu klären, ob es im zweiten Schritt (Pricing) und im dritten Schritt (Ratio-Test) Auswahlregeln gibt, die ein Kreiseln auch im degenerierten Fall verhindern. Der folgende Satz gibt eine solche Regel an.

Satz 5.11 *Unter Verwendung einer der beiden folgenden Auswahlregeln ist der Algorithmus 5.6 auch im degenerierten Fall endlich.*

(a) Bland (siehe [Bl77]).

Pricing: *Wähle den kleinsten Index $j \in N$ mit $\bar{z}_j < 0$.*

Ratio Test: *Wähle unter allen Indizes $i \in \{1, 2, \ldots, m\}$ mit $\frac{\bar{x}_{B_i}}{w_i} = \gamma$ denjenigen mit kleinstem Index B_i.*

(b) Lexikografische Regel (Dantzig, Orden, Wolfe, siehe [Da55]).

Pricing: *Wähle einen Index $j \in N$ mit $\bar{z}_j < 0$ beliebig.*

Ratio Test: *Wähle* $i \in \{1, 2, \ldots, m\}$ *mit* $\frac{\bar{x}_{B_i}}{w_i} = \gamma$ *so, dass* $\frac{(A_B^{-1}A)_{i\cdot}}{w_i}$ *lexikographisch minimal ist.*

Ein Vektor λ *sei dabei lexikografisch kleiner als ein Vektor* μ *gleicher Dimension, falls* $\lambda \neq \mu$ *und für den kleinsten Index* i *mit* $\lambda_i \neq \mu_i$ *gilt* $\lambda_i < \mu_i$.

Beweise dieser Aussagen findet man z.B. in Schrijver [Sch86] oder bei Chvátal [Ch83].

Bemerkung 5.12 *Man kann zeigen, dass 5.11(b) äquivalent zur Perturbationsmethode ist, d.h. man betrachtet*

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b + \varepsilon \\ & x \geq 0 \end{array} \tag{5.7}$$

für ein $\varepsilon \in \mathbb{R}^m$, $\varepsilon > 0$. *Man kann zeigen, dass bei geeigneter Wahl von* ε *das lineare Problem (5.7) keine degenerierten Ecken enthält und jede zulässige Basis von (5.7) eine zulässige Basis von (5.1) ist (vgl. Chvátal [Ch83]).*

Aus praktischer Sicht stellt sich natürlich nicht nur die Frage der Endlichkeit, sondern man benötigt auch Aussagen über das Worst-Case Verhalten des Simplex-Algorithmus. Weiter stellt sich die Frage, ob es eine Auswahlregel für die Schritte zwei und vier gibt, so dass der Simplex-Algorithmus polynomiale Laufzeit hat. Bislang ist eine solche Regel nicht bekannt.

Beispiel 5.13 *Betrachte für ein* ε *mit* $0 < \varepsilon < \frac{1}{2}$ *das Problem*

$$\begin{array}{lrll} \min & -x_n & & \\ \text{s.t.} & \varepsilon & \leq x_1 \leq & 1 \\ & \varepsilon x_{j-1} & \leq x_j \leq & 1 - \varepsilon x_{j-1} \qquad \text{für } j = 2, 3, \ldots, n. \end{array}$$

Dieses Beispiel ist in der Literatur unter dem Namen Klee-Minty-Würfel *bekannt (benannt nach seinen Entdeckern). Man kann zeigen, dass der Simplex-Algorithmus* 2^n *Iterationen zur Lösung des obigen LPs benötigt, demnach also alle Ecken abläuft (zum Beweis siehe beispielsweise Papadimitriou & Steiglitz [PS82]).*

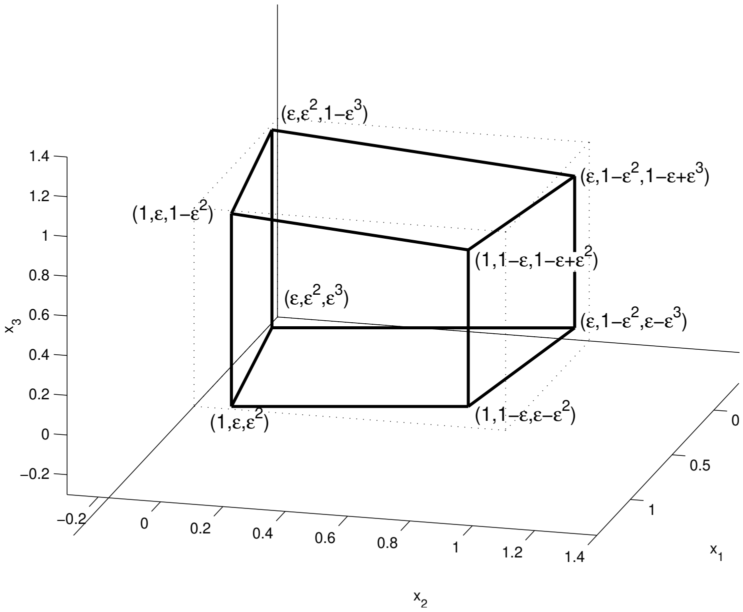

Abbildung 5.4: Klee-Minty-Würfel (Beispiel 5.13)

5.3 Phase I des Simplex-Algorithmus: Das Finden einer zulässigen Basis

In diesem Kapitel wollen wir das Problem lösen, für (5.1)

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0. \end{array}$$

eine (primal) zulässige Basis zu finden oder aber festzustellen, dass $\mathcal{P}^=(A,b) = \emptyset$ gilt. Wir nehmen ohne Einschränkung an, dass $b \geq 0$ ist (sonst multipliziere die Gleichungen mit -1, für die $b_i < 0$ ist).

Wir verzichten in diesem Kapitel zunächst auf die Annahme, dass A vollen Zeilenrang hat, d.h. wir lassen $\operatorname{rang}(A) \leq m$ zu. Wir betrachten nun folgendes lineare Problem ($\mathbb{1}$ sei dabei der Vektor, dessen Komponenten alle gleich

1 sind):

$$\begin{array}{ll} \min & \mathbb{1}^T y \\ \text{s.t.} & Ax + y = b \\ & x, y \geq 0. \end{array} \tag{5.8}$$

(5.8) ist in Standardform gegeben und die Matrix $D = (A, I)$ hat offensichtlich vollen Zeilenrang. Die Variablen des Vektors y werden als künstliche Variablen bezeichnet. Seien $\{1, 2, \ldots, n, n+1, \ldots, n+m\}$ die Spaltenindizes von D und $u = \binom{x}{y}$. Dann ist $B = (n+1, n+2, \ldots, n+m)$ eine primal zulässige Basis mit $u_B = y = b \geq 0$ und $u_N = x = 0$.
Es gilt nun folgender Satz.

Satz 5.14 *In dem linearen Programm (5.1) sei $b \geq 0$. Dann gelten für das lineare Programm (5.8):*

(i) $B = (n+1, n+2, \ldots, n+m)$ *ist eine primal zulässige Basis mit zulässiger Basislösung* $u = \binom{x}{y} = \binom{0}{b}$.

(ii) (5.8) ist lösbar.

(iii) Sei $\binom{x^*}{y^*}$ *optimale Lösung von (5.8) nach Anwendung des Simplex-Verfahrens 5.6.*

Ist dann $y^* \neq 0$, dann hat (5.1) keinen zulässigen Punkt.

Ist dagegen $y^* = 0$ und gilt $\operatorname{rang}(A) = m$, *dann ist x^* zulässige Basislösung von (5.8).*

Beweis. zu (i): Ist wegen $D_B = I$, $Du = y = b$ und $u_B = y \geq 0$, $u_N = x = 0$ klar.
zu (ii): Die Zielfunktion ist durch 0 nach unten beschränkt und nach (i) existiert ein zulässiger Punkt. Nach dem starken Dualitätssatz existiert also eine Lösung.
zu (iii): Ist $y^* \neq 0$, dann ist der Optimalwert $\mathbb{1}^T y^* > 0$. In diesem Fall kann (5.1) keinen zulässigen Punkt x besitzen, denn sonst wäre $\binom{x}{0}$ zulässig für (5.8) mit kleinerem Zielfunktionswert 0.
Ist $y^* = 0$ dann ist x^* zulässig für (5.1).
$\binom{x^*}{0}$ ist als Ergebnis des Simplex-Verfahrens eine zulässige Basislösung von (5.8). Sei B die zugehörige Basis und setze

$$B_x = \{1, \ldots, n\} \cap B, \quad B_y = B \setminus B_x, \quad N_x = \{1, \ldots, n\} \cap N, \quad N_y = N \setminus N_x.$$

Dann gilt $x^*_{B_x} \geq 0$, $x^*_{N_x} = 0$ und A_{B_x} hat linear unabhängige Spalten. Gilt nun $|B_x| = m$, dann ist B_x bereits primal zulässige Basis von (5.1). Andernfalls kann im Fall $\operatorname{rang}(A) = m$ die Matrix A_{B_x} durch $m - |B_x|$ geeignete Spalten aus A_{N_x} zu einer invertierbaren Basismatrix ergänzt werden und x^* ist die zugehörige Basislösung. $\square$

Phase I mit dem dem in (iii) angesprochenen Ergänzungsprozess einer unvollständigen Basis kann zum Beispiel wie folgt durchgeführt werden:

Algorithmus 5.15 Phase I des Simplex-Algorithmus.

Input: *A, b, c für Problem (5.1) mit $b \geq 0$.*

Output:
(i) *Eine zulässige Basis B für (5.1) mit zugehöriger Basislösung $\bar{x}$*
oder
(ii) *Die Meldung „Das LP (5.1) ist unzulässig".*
oder
(iii) *Die Meldung „$\operatorname{rang}(A) < m$ ".*

(1) *Wende den Simplex-Algorithmus 5.6 auf (5.8) mit Startbasis $B = (n+1, \dots, n+m)$ und Basislösung $\binom{0}{b}$ an. Sei $\binom{\bar{x}}{\bar{y}}$ die erhaltene Lösung mit zugehöriger Basis B.*

(2) *Ist $\bar{y} \neq 0$: STOP, das LP (5.1) ist unzulässig.*

(3) *Andernfalls setze*

$$\begin{aligned} B_x &= \{1, \dots, n\} \cap B, \quad B_y = B \setminus B_x, \\ N_x &= \{1, \dots, n\} \cap N, \quad N_y = N \setminus N_x. \end{aligned}$$

While $|B_x| < m$:

Bestimme $j \in N_x$, so dass für die Lösung w von

$$(I_{B_y - n}, A_{B_x})w = A_{\cdot j}$$

gilt $w_i \neq 0$ für ein $i \leq m - |B_x|$.

Existiert kein solches j: STOP, $\operatorname{rang}(A) < m$.

Sonst setze $B_x := B_x \cup \{j\}$, $B_y := B_y \setminus \{(B_y)_i\}$, $N_x = N_x \setminus \{j\}$.

End while

STOP, $B := B_x$ ist zulässige Basis für (5.1) mit Basislösung $\bar{x}$.

Wir zeigen noch, dass Schritt (3) die Basisergänzung durchführt:
Zu Beginn ist $(I_{B_y - n}, A_{B_x})$ als Basismatrix invertierbar. Gilt für die Lösung w von

$$(I_{B_y - n}, A_{B_x})w = A_{\cdot j} \tag{5.9}$$

$w_{\{1,\dots,m-|B_x|\}}=0$, dann ist $A_{\cdot j}$ linear abhängig von den Spalten von A_{B_x}. Im Fall rang $(A) = m$ kann dies nicht für alle $j \in N_x$ auftreten.
Mit $B_y^+ := B_y \setminus \{(B_y)_i\}$ ist nun die neue Basismatrix $(I_{B_y^+-n}, A_{B_x}, A_{\cdot j})$ invertierbar. Denn sonst ließe sich $A_{\cdot j}$ als Linearkombination der übrigen linear unabhängigen Spalten darstellen, in der eindeutigen Lösung von (5.9) wäre also $w_i = 0$, was nicht der Fall ist.

Bemerkung 5.16 *Die Anwendung von Algorithmus 5.6 auf Problem (5.8) wird häufig als* Phase I *des Simplex-Verfahrens bezeichnet.*

5.4 Implementierung des Simplex-Verfahrens

Im folgenden wollen wir noch einige Anmerkungen dazu machen, wie das Simplex-Verfahren in der Praxis implementiert wird.

Lösen der linearen Gleichungssysteme

Das Lösen von linearen Gleichungssystemen kann sehr hohe Laufzeiten verursachen. Das Bilden einer Inversen benötigt mindestens $\mathcal{O}(n^{2+\varepsilon})$ Schritte (hierbei gilt $\varepsilon \in (0,1]$). Hinzu kommt, dass die Inverse dicht besetzt sein kann (d.h. viele Nicht-Null-Einträge), obwohl die Ausgangsmatrix dünn besetzt ist. Hier ein paar Anmerkungen, wie man diese Probleme vermeiden kann.
Anstelle der Basisinversen A_B^{-1} kann man z.B. eine LU-Faktorisierung von A_B bestimmen, d.h. man erzeugt eine untere Dreiecksmatrix L und eine obere Dreiecksmatrix U, so dass $L \cdot U = PA_BQ$ gilt, wobei P, Q Permutationsmatrizen sind, die von Zeilen- und Spaltenvertauschungen bei der Pivotwahl herrühren. Das System $A_Bx = b$ kann dann in zwei Schritten gelöst werden:

$$b = A_Bx \quad\iff\quad Pb = PA_BQz = L\underbrace{Uz}_{=:y},\ x = Qz \quad\iff\quad \begin{array}{rcl} Ly &=& Pb \\ Uz &=& y \\ x &=& Qz. \end{array}$$

Durch die Dreiecksgestalt der Matrizen L und U können *beide* Gleichungssysteme durch einfaches Vorwärts-, bzw. Rückwärtseinsetzen gelöst werden. Wichtig ist bei der LU-Zerlegung, L und U dünn besetzt zu halten. Dazu gibt es eine Reihe von Pivot-Strategien bei der Bestimmung von L und U. Eine

sehr gute Implementierung einer LU-Zerlegung für dünn besetzte Matrizen ist z.B. in Suhl & Suhl [SS90] zu finden.
Dennoch wäre es immer noch sehr teuer, in jeder Iteration des Simplex-Algorithmus eine neue LU-Faktorisierung zu bestimmen. Glücklicherweise gibt es Update-Formeln, um die LU-Faktorisierung einer Basismatrix auch in späteren Iterationen nutzen zu können. Dies führt auf folgende Strategie:

- Faktorisiere eine Basismatrix A_{B_0}.
- Verwende Update-Formeln, um für nachfolgende Basen $B_1, B_2, \ldots, B_k$ die Lösung von Gleichungssystemen mit den Matrizen A_{B_k}, $A_{B_k}^T$ auf ein Sytem mit der Matrix A_{B_0} bzw. $A_{B_0}^T$ zurückzuführen, das effizient mit Hilfe der vorhandenen LU-Faktorisierung gelöst werden kann.

Die Basis bildet der folgende Satz. Mit den Bezeichnungen aus Algorithmus 5.6 gilt:

Satz 5.17 *Sei*

$$\eta_j = \begin{cases} \dfrac{1}{w_i}, & \textit{falls } j = i \\ -\dfrac{w_j}{w_i}, & \textit{sonst.} \end{cases}$$

sowie die Matrizen F und E wie folgt gegeben:

$$F = \begin{array}{c} \begin{pmatrix} 1 & & & w_1 & & & \\ & \ddots & & \vdots & & 0 & \\ & & 1 & \vdots & & & \\ & & & w_i & & & \\ & & & \vdots & 1 & & \\ & 0 & & \vdots & & \ddots & \\ & & & w_m & & & 1 \end{pmatrix} \\ \uparrow \\ i \end{array} \qquad E = \begin{array}{c} \begin{pmatrix} 1 & & & \eta_1 & & & \\ & \ddots & & \vdots & & 0 & \\ & & 1 & \vdots & & & \\ & & & \eta_i & & & \\ & & & \vdots & 1 & & \\ & 0 & & \vdots & & \ddots & \\ & & & \eta_m & & & 1 \end{pmatrix} \\ \uparrow \\ i \end{array}$$

Dann gelten die Beziehungen $A_{B^+} = A_B \cdot F$ und $A_{B^+}^{-1} = E \cdot A_B^{-1}$.

Beweis. Zunächst gilt $A_{B^+} = A_B \cdot F$ (vgl. FTRAN) und durch Nachrechnen verifiziert man $F \cdot E = I$, d.h. $F^{-1} = E$. Damit folgt nun

$$A_{B^+}^{-1} = (A_B \cdot F)^{-1} = F^{-1} \cdot A_B^{-1} = E \cdot A_B^{-1}.$$

□

Satz 5.17 kann nun verwendet werden, um Gleichungssysteme in Folgeiterationen zu lösen, ohne explizit eine neue Faktorisierung berechnen zu müssen. Sei B_0 die Basis, bei der zum letzten Mal eine LU-Faktorisierung berechnet wurde, d.h. $L \cdot U = A_{B_0}$ und wir betrachten die k-te Simplexiteration danach und wollen daher ein Gleichungssystem der Form

$$A_{B_k} \bar{x}_k = b \tag{$*$}$$

lösen. Dann gilt

$$A_{B_k} = A_{B_0} \cdot F_1 \cdot F_2 \cdots F_k$$

und

$$\begin{aligned} \bar{x}_k &= F_k^{-1} \cdot F_{k-1}^{-1} \cdots F_1^{-1} x_0 \\ &= E_k \cdot E_{k-1} \cdots E_1 x_0, \end{aligned}$$

wobei x_0 die Lösung von $A_{B_0} x = b$ ist. $\bar{x}_k$ ist die Lösung von $(*)$, denn

$$\begin{aligned} A_{B_k} \bar{x}_k &= (A_{B_0} \cdot F_1 \cdot F_2 \cdots F_k) \cdot (E_k \cdot E_{k-1} \cdots E_1 x_0) \\ &= A_{B_0} \underbrace{(F_1 \cdot F_2 \cdots F_k \cdot E_k \cdot E_{k-1} \cdots E_1)}_{=I} x_0 \\ &= A_{B_0} x_0 = b. \end{aligned}$$

Beachte, dass diese Herleitung eine nochmalige Bestätigung der Korrektheit der Update-Formel im fünften Schritt (Update) des Algorithmus 5.6 darstellt. Mit der Beziehung $A_{B_k} = A_{B_0} \cdot F_1 \cdot F_2 \cdots F_k$ lassen sich entsprechend Update-Formeln für BTRAN (zur Berechnung von y) und FTRAN (zur Berechnung von w) herleiten. Mehr Informationen zur effizienten Implementierung dieser Berechnungsmöglichkeiten findet man z.B. in Forrest & Tomlin [FT72].

Abschließend sei bemerkt, dass die Update-Formeln die Gefahr bergen, dass die numerischen Fehler in den Lösungsvektoren verstärkt werden. Es empfiehlt sich daher, hin und wieder neu zu faktorisieren (aktuelle Codes tun dies etwa alle 100 Iterationen).

Pricing

In der Literatur wird eine Vielzahl von Auswahlregeln diskutiert, welcher Index $j \in N$ im Pricing gewählt werden soll. Die am häufigsten verwendeten Regeln in heutigen State-of-the-art Paketen sind:

(1) Kleinster Index.
Diese Auswahlregel ist Teil von Blands Regel zum Vermeiden des Kreiselns.

(2) Volles Pricing (Dantzigs Regel).
Berechne die reduzierten Kosten für alle Nichtbasisvariablen und wähle einen der Indizes mit den kleinsten reduzierten Kosten.

- Wird häufig benutzt in der Praxis.
- Kann sehr aufwendig sein, wenn die Anzahl der Variablen groß ist.

(3) Partial und Multiple Pricing.
Partial Pricing versucht die Nachteile des vollen Pricings zu vermeiden und betrachtet nur eine Teilmenge der Nichtbasisvariablen. Nur wenn diese Teilmenge keine negativen reduzierten Kosten enthält, wird eine neue Teilmenge betrachtet. Dies wird fortgesetzt, bis alle reduzierten Kosten nicht-negativ sind.

Multiple Pricing nutzt die Tatsache, dass Variablen mit negativen reduzierten Kosten in Folgeiterationen häufig wiederum negative reduzierte Kosten haben. Deshalb werden Variablen, die in früheren Iterationen bereits negative reduzierte Kosten hatten, zuerst betrachtet.

Multiple Partial Pricing kombiniert diese beiden Ideen.

(4) Steepest Edge Pricing (Kuhn & Quandt [KQ63], Wolfe & Cutler [WC63]).

Die Idee besteht hierbei darin, eine Variable zu wählen (geometrisch gesehen eine Kante), die am steilsten bzgl. der Zielfunktion ist, die also pro Einheit „Variablenerhöhung“ den größten Fortschritt in der Zielfunktion bewirkt. In Formeln sieht das folgendermaßen aus:

Der Update im fünften Schritt von Algorithmus 5.6 lautet

$$\begin{pmatrix} \bar{x}_B \\ \bar{x}_N \end{pmatrix} = \begin{pmatrix} \bar{x}_B \\ \bar{x}_N \end{pmatrix} + \gamma\, \eta_j \qquad \text{mit} \qquad \eta_j = \begin{pmatrix} -A_B^{-1} A_{\cdot j} \\ (e_j)_N \end{pmatrix}.$$

Mit dieser Definition von η gilt für die reduzierten Kosten

$$\bar{z}_j = c_j - c_B^T A_B^{-1} A_{\cdot j} = c^T \eta_j.$$

Im Gegensatz zum vollen Pricing, wo ein $j \in N$ gewählt wird mit

$$c^T \eta_j = \min_{l \in N} c^T \eta_l,$$

wählen wir nun ein $j \in N$ mit

$$\frac{c^T \eta_j}{\|\eta_j\|} = \min_{l \in N} \frac{c^T \eta_l}{\|\eta_l\|}.$$

Vorteil: Deutlich weniger Simplex-Iterationen nötig.

Nachteil: Rechenaufwendig, es muss in jedem Schritt ein Gleichungssystem zur Berechnung der η_j gelöst werden. Abhilfe schaffen die Update-Formeln von Goldfarb und Reid (siehe [GR77]).

Alternative: Approximation der Normen (siehe [Ha73]). Bekannt als Devex Pricing.

Ratio-Test

Ein ernsthaftes Problem im Ratio-Test ist es, dass das Minimum γ u.U. für einen Index i angenommen wird, für den der Nenner w_i sehr klein ist. Dies führt zu numerischen Instabilitäten. Eine Idee, um dieses Problem zu lösen, geht auf Harris (siehe [Ha73]) zurück. Die Berechnung erfolgt in mehreren Schritten.

(1) Bestimme $r_k = \begin{cases} \dfrac{\bar{x}_{B_k}}{w_k} & \text{, falls } w_k > 0 \\ +\infty & \text{sonst} \end{cases}$ für $k = 1, 2, \ldots, m$.

(2) Berechne

$$t = \min \left\{ r_k + \frac{\varepsilon}{w_k} \;\middle|\; k = 1, 2, \ldots, m \right\},$$

wobei $\varepsilon > 0$ eine vorgegebene Toleranz bezeichnen soll (z.B. $\varepsilon = 10^{-6}$).

(3) Der verlassende Index i ist nun

$$i = \text{argmax}\,\{w_k : r_k \leq t\,,\ k = 1, 2, \ldots, m\}.$$

Beachte, dass $\bar{x}_B$ negativ werden kann, da $\varepsilon > 0$ gilt, und daraus eine unzulässige Basis resultiert. In diesem Fall wird die untere Schranke Null auf mindestens $-\varepsilon$ gesetzt (engl.: shifting) und die Berechnungen der Schritte (1) – (3) wiederholt.

Wie man Variablen mit unteren Schranken, deren Wert ungleich Null ist, behandelt, werden wir im nächsten Abschnitt sehen.

Die unzulässigen Schranken werden erst am Ende des Algorithmus entfernt. Dies geschieht durch einen Aufruf der Phase I mit einem anschliessenden erneuten Durchlauf der Phase II des Simplex-Algorithmus. Die Hoffnung dabei ist, dass am Ende nur wenige Variablen echt kleiner als Null sind und damit die beiden Phasen schnell ablaufen. Mehr Details dazu findet man in Gill, Murray, Saunders & Wright [GMSW89].

5.5 Varianten des Simplex-Algorithmus

In diesem Abschnitt wollen wir uns noch mit zwei Varianten / Erweiterungen des Algorithmus 5.6 beschäftigen. Zum einen ist dies die Behandlung von unteren und oberen Schranken, zum anderen ist es der duale Simplex-Algorithmus. Wir beginnen mit letzterem.

5.5.1 Der duale Simplex-Algorithmus

Die Grundidee des dualen Simplex-Algorithmus ist es, den primalen Simplex-Algorithmus 5.6 auf das duale lineare Problem in Standardform (5.3) anzuwenden, ohne (5.1) explizit zu dualisieren. Historisch war die Entwicklung jedoch anders: Man dachte zunächst, man hätte ein neues Verfahren gefunden, ehe man den obigen Zusammenhang erkannte. Betrachten wir nochmals (5.1)

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array}$$

und das dazu duale Problem in Standardform (5.3)

$$\begin{array}{ll} \max & b^T y \\ \text{s.t.} & A^T y + Iz = c \\ & z \geq 0. \end{array}$$

Wir gehen im Folgenden wieder davon aus, dass A vollen Zeilenrang hat, d.h. $\operatorname{rang}(A) = \operatorname{rang}(A^T) = m \leq n$. Mit $D = (A^T, I) \in \mathbb{R}^{n \times (m+n)}$ erhält (5.3) die Form

$$\begin{array}{ll} \max & b^T y \\ \text{s.t.} & D\binom{y}{z} = c \\ & z \geq 0. \end{array}$$

Eine Basis H des dualen Problems in (5.3) hat also die Mächtigkeit n mit $H \subseteq \{1, 2, \ldots, m, m+1, \ldots, m+n\}$. Wir machen zunächst folgende Beobachtung:

Satz 5.18 *Sei $A \in \mathbb{R}^{m \times n}$ eine Matrix mit vollem Zeilenrang und H eine zulässige Basis von (5.3). Dann ist (5.3) unbeschränkt oder es existiert eine optimale Basis H_{opt} mit $\{1, 2, \ldots, m\} \subseteq H_{opt}$.*

Beweis. Übung. □

Im Gegensatz zur Anwendung von Algorithmus 5.6 auf (5.1) haben wir hier die Besonderheit, dass nicht alle Variablen Vorzeichenbeschränkungen unterliegen. Die Variablen in y sind sogenannte freie Variablen. Satz 5.18 sagt aus, dass wir o.B.d.A. davon ausgehen können, dass alle Variablen in y in einer Basis H für (5.3) enthalten sind.

Da $|H| = n \geq m$ und y aus $m \leq n$ Variablen besteht, muss H noch $(n - m)$ Variablen aus z enthalten. Wir bezeichnen mit $N \subseteq \{1, 2, \ldots, n\}$ diese $(n-m)$ Variablen und mit $B = \{1, 2, \ldots, n\} \setminus N$ die Indizes aus z, die nicht in der Basis sind. Es gilt also

Basis: $H = \{1, \ldots, m\} \cup (N + m)$, Basismatrix: $D_H = (A^T, I_{\cdot N})$
Nichtbasis: $M = \{1, \ldots, m+n\} \setminus H = B + m$, Nichtbasismatrix: $D_M = I_{\cdot B}$.

Wir wollen nun zulässige Basen H charakterisieren: Da $D_H = (A^T, I_{\cdot N})$ regulär ist, muss auch A_B^T regulär sein, also ist A_B regulär. Nun gilt

$$
\begin{aligned}
D\begin{pmatrix} y \\ z \end{pmatrix} = c,\ z \geq 0 \quad &\Longleftrightarrow \quad A^T y + Iz = c, z \geq 0 \\
&\Longleftrightarrow \quad \left\{ \begin{array}{rcll} A_B^T y + z_B &=& c_B, & z_B \geq 0 \\ A_N^T y + z_N &=& c_N, & z_N \geq 0 \end{array} \right\} \\
&\Longleftrightarrow \quad \left\{ \begin{array}{rcl} y &=& A_B^{-T}(c_B - z_B) \\ z_N &=& c_N - A_N^T A_B^{-T}(c_B - z_B) \\ z_N, z_B &\geq& 0 \end{array} \right\} \qquad (5.10)
\end{aligned}
$$

Setzen wir die Nichtbasisvariablen zu 0, also $z_B = 0$, so ist H (primal) zulässig für (5.3), falls

$$z_N = c_N - A_N^T A_B^{-T} c_B \geq 0,$$

was äquivalent dazu ist, dass B dual zulässig ist (vgl. Definition 5.1(a)). Beachte, dass H durch N und B eindeutig festgelegt ist. Daher ist es üblich, der Definition 5.1(a) zu folgen und von einer dual zulässigen Basis B für (5.1) zu sprechen und weniger von einer (primal) zulässigen Basis H für (5.3). Betrachten wir also eine dual zulässige Basis B (Beachte: B sind die Nichtbasisvariablen im Dualen und N sind die Basisvariablen im Dualen) mit

$$
\begin{aligned}
\bar{z}_N &= c_N - A_N^T A_B^{-T} c_B \geq 0, \\
\bar{z}_B &= 0
\end{aligned}
$$

und wenden Algorithmus 5.6 auf (5.3) an.

(1) BTRAN liefert (vgl. primaler Simplex-Alg.: $A_B^T \bar{y} = c_B$)

$$
\begin{aligned}
D_H^T \bar{x} = \begin{pmatrix} b \\ 0 \end{pmatrix} \quad &\Longleftrightarrow \quad \begin{pmatrix} A \\ I_{\cdot N}^T \end{pmatrix} \bar{x} = \begin{pmatrix} b \\ 0 \end{pmatrix} \\
&\Longleftrightarrow \quad \begin{pmatrix} A_B & A_N \\ 0 & I_{NN} \end{pmatrix} \begin{pmatrix} \bar{x}_B \\ \bar{x}_N \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix} \\
&\Longleftrightarrow \quad \begin{array}{rcl} A_B \bar{x}_B + A_N \bar{x}_N &=& b \\ \bar{x}_N &=& 0 \end{array} \\
&\Longleftrightarrow \quad \bar{x}_N = 0,\ A_B \bar{x}_B = b \\
&\Longleftrightarrow \quad \bar{x}_N = 0,\ \bar{x}_B = A_B^{-1} b
\end{aligned}
$$

(2) Pricing berechnet die reduzierten Kosten (vgl. primaler Simplex-Alg.: $\bar{z}_N = c_N - A_N^T \bar{y}$) zu

$$\bar{r}_B = 0_B - D_M^T \bar{x} = -I_{\cdot B}^T \bar{x} = -I_{B\cdot} \bar{x} = -\bar{x}_B.$$

Da (5.3) ein Maximierungsproblem ist, können wir die Zielfunktion nur verbessern, falls die reduzierten Kosten > 0 sind für eine Nichtbasisvariable, also für einen Index $B_i \in B$.

Sind die reduzierten Kosten ≤ 0, gilt also

$$-\bar{x}_B \leq 0 \qquad \Longleftrightarrow \qquad \bar{x}_B \geq 0,$$

so ist H (bzw. B) dual optimal (d.h. B ist primal zulässig). Andernfalls wähle einen Index B_i mit $\bar{x}_{B_i} < 0$.

Vgl. auch die duale Zielfunktion

$$b^T y \overset{(5.10)}{=} b^T (A_B^{-T}(c_B - z_B)) = b^T A_B^{-T} c_B - \underbrace{(A_B^{-1} b)^T}_{\geq 0 \;\Rightarrow\; \text{optimal}} z_B.$$

(3) FTRAN berechnet (vgl. primaler Simplex-Alg.: $A_B w = A_{\cdot j}$)

$$\begin{aligned} D_H \begin{pmatrix} w \\ \alpha_N \end{pmatrix} = I_{\cdot B_i} \quad &\Longleftrightarrow \quad (A^T, I_{\cdot N}) \begin{pmatrix} w \\ \alpha_N \end{pmatrix} = e_{B_i} \\ &\Longleftrightarrow \quad \begin{pmatrix} A_B^T & 0 \\ A_N^T & I_N \end{pmatrix} \begin{pmatrix} w \\ \alpha_N \end{pmatrix} = \begin{pmatrix} e_i \\ 0 \end{pmatrix} \\ &\Longleftrightarrow \quad A_B^T w = e_i \quad \text{und} \quad \alpha_N = -A_N^T w. \end{aligned}$$

(4) Ratio-Test
Wir überprüfen, ob $\alpha_N \leq 0$ ist (Beachte, dass das Vorzeichen von w egal ist, da die Variablen in y freie Variablen sind). Ist dies der Fall, so ist (5.3) unbeschränkt, d.h. $\mathcal{P}^=(A, b) = \emptyset$. Andernfalls setze

$$\gamma = \frac{\bar{z}_j}{\alpha_j} = \min \left\{ \frac{\bar{z}_k}{\alpha_k} : \alpha_k > 0, k \in N \right\}$$

mit $j \in N$, $\alpha_j > 0$. Damit verlässt nun j die duale Basis H bzw. tritt in die Basis B ein.

Dies zusammen liefert

Algorithmus 5.19 Der duale Simplex-Algorithmus.

Input: *Dual zulässige Basis* B , $\bar{z}_N = c_N - A_N^T A_B^{-T} c_B \geq 0$.

Output:

(i) *Eine Optimallösung* $\bar{x}$ *für (5.1) bzw.*
eine Optimallösung $\bar{y} = A_B^{-T} c_B$ *für (5.2) bzw.*
eine Optimallösung $\bar{y}, \bar{z}_N, \bar{z}_B = 0$ *für (5.3)*

oder

(ii) *Die Meldung* $\mathcal{P}^=(A, b) = \emptyset$ *bzw. (5.3) ist unbeschränkt.*

(1) BTRAN
Löse $A_B \bar{x}_B = b$.

(2) Pricing
Falls $\bar{x}_B \geq 0$, *so ist die Basis optimal für (5.1) bzw. (5.3),* **Stop**.

Andernfalls wähle einen Index $i \in \{1, 2, \ldots, m\}$ *mit* $\bar{x}_{B_i} < 0$,
d.h. B_i *verlässt die Basis* B.

(3) FTRAN
Löse $A_B^T w = e_i$ *und berechne* $\alpha_N = -A_N^T w$.

(4) Ratio-Test
Falls $\alpha_N \leq 0$, *so ist (5.3) unbeschränkt bzw.* $\mathcal{P}^=(A, b) = \emptyset$, **Stop**.

Andernfalls berechne

$$\gamma = \frac{\bar{z}_j}{\alpha_j} = \min \left\{ \frac{\bar{z}_k}{\alpha_k} : \alpha_k > 0, k \in N \right\}$$

mit $j \in N, \alpha_j > 0$. *Die Variable* j *tritt dann in die Basis ein.*

(5) Update
Setze
$$\begin{aligned} \bar{z}_N &= \bar{z}_N - \gamma \alpha_N, \\ \bar{y} &= \bar{y} - \gamma w, \\ \bar{z}_{B_i} &= \gamma \\ N &= (N \setminus \{j\}) \cup B_i, \\ B_i &= j. \end{aligned}$$

Gehe zu (1).

Die Korrektheit des Algorithmus 5.19 und alle weiteren Konsequenzen gelten dem Abschnitt 5.2 (bzw. Abschnitt 5.3) entsprechend. Dabei wird die Tatsache genutzt, dass Algorithmus 5.19 nichts anderes ist als die Anwendung von Algorithmus 5.6 auf (5.3).

Beispiel 5.20 *Betrachte*

$$\begin{array}{llcrcl} \min & 2x_1 & + & x_2 & & \\ \text{s.t.} & -x_1 & - & x_2 & \leq & -\frac{1}{2} \\ & -4x_1 & - & x_2 & \leq & -1 \\ & & & x_1, x_2 & \geq & 0. \end{array}$$

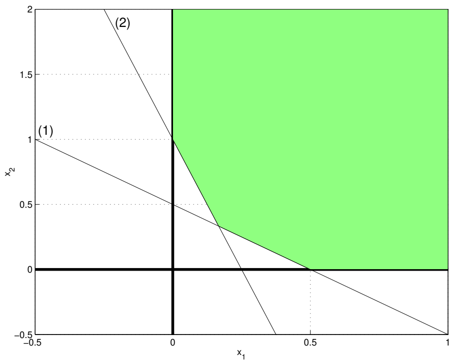

Abbildung 5.5: Grafische Darstellung zu Beispiel 5.20.

In Standardform lautet das LP

$$\begin{array}{llcrcrcrcl} \min & 2x_1 & + & x_2 & & & & & & \\ \text{s.t.} & -x_1 & - & x_2 & + & x_3 & & & = & -\frac{1}{2} \\ & -4x_1 & - & x_2 & & & + & x_4 & = & -1 \\ & & & & & & x_1, x_2, & x_3, x_4 & \geq & 0. \end{array}$$

Wir starten mit $B = (3,4)$, $N = (1,2)$. *Es gilt*

$$\bar{z}_N = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - A_N^T A_B^{-T} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \geq 0.$$

Damit ist B *dual zulässig.*

(1.1) BTRAN

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bar{x}_B = \begin{pmatrix} -\frac{1}{2} \\ -1 \end{pmatrix} \quad \Longrightarrow$$

$$\bar{x}_B = \begin{pmatrix} \bar{x}_3 \\ \bar{x}_4 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ -1 \end{pmatrix}, \qquad \bar{x}_N = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

(1.2) Pricing:
$\bar{x} \not\geq 0$. *Wähle* $i = 2$ *mit* $\bar{x}_{B_2} = \bar{x}_4 = -1 < 0$.

(1.3) FTRAN

$$A_B^T w = e_2 \quad \Leftrightarrow \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} w = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \Longrightarrow \qquad w = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

und

$$\alpha_N = -A_N^T w \quad \Longrightarrow \quad \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}.$$

(1.4) Ratio-Test

$$\gamma = \min \left\{ \frac{2}{4}, \frac{1}{1} \right\} = \frac{1}{2} \qquad \text{mit } j = 1.$$

(1.5) Update

$$\begin{aligned} \bar{z}_N &= \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix}, \\ \bar{z}_{B_2} &= \bar{z}_4 = \tfrac{1}{2} \\ N &= (4, 2), \\ B &= (3, 1). \end{aligned}$$

(2.1) BTRAN

$$\begin{pmatrix} 1 & -1 \\ 0 & -4 \end{pmatrix} \bar{x}_B = \begin{pmatrix} -1/2 \\ -1 \end{pmatrix} \qquad \Longrightarrow$$

$$\bar{x}_B = \begin{pmatrix} \bar{x}_3 \\ \bar{x}_1 \end{pmatrix} = \begin{pmatrix} -1/4 \\ 1/4 \end{pmatrix}, \qquad \bar{x}_N = \begin{pmatrix} \bar{x}_4 \\ \bar{x}_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

(2.2) Pricing $\bar{x}_B \not\geq 0$. *Wähle* $i = 1$ *mit* $\bar{x}_{B_1} = \bar{x}_3 = -\frac{1}{4} < 0$.

(2.3) FTRAN

$$A_B^T w = e_1 \quad \Leftrightarrow \quad \begin{pmatrix} 1 & 0 \\ -1 & -4 \end{pmatrix} w = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \Longrightarrow \qquad w = \begin{pmatrix} 1 \\ -1/4 \end{pmatrix}$$

und

$$\alpha_N = -A_N^T w = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1/4 \end{pmatrix} = \begin{pmatrix} 1/4 \\ 3/4 \end{pmatrix}.$$

(2.4) Ratio-Test

$$\gamma = \min\left\{\frac{1/2}{1/4}, \frac{1/2}{3/4}\right\} = \frac{2}{3} \qquad \textit{mit } j = 2.$$

(2.5) Update

$$\begin{aligned} \bar{z}_N &= \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} - \frac{2}{3}\begin{pmatrix} 1/4 \\ 3/4 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 0 \end{pmatrix}, \\ \bar{z}_{B_1} &= \bar{z}_3 = \tfrac{2}{3}, \\ N &= (4,3), \\ B &= (2,1). \end{aligned}$$

(3.1) BTRAN

$$\begin{pmatrix} -1 & -1 \\ -1 & -4 \end{pmatrix} \bar{x}_B = \begin{pmatrix} -1/2 \\ -1 \end{pmatrix} \qquad \Longrightarrow$$

$$\bar{x}_B = \begin{pmatrix} \bar{x}_2 \\ \bar{x}_1 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/6 \end{pmatrix} \qquad \bar{x}_N = \begin{pmatrix} \bar{x}_3 \\ \bar{x}_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

(3.2) Pricing $\bar{x}_B \geq 0$. Daher ist $\bar{x}_1 = \frac{1}{6}, \bar{x}_2 = \frac{1}{3}$ optimal.

5.5.2 Obere und untere Schranken

Häufig haben Variablen in linearen Problemen obere und untere Schranken, z.B. bei Relaxierungen von ganzzahligen oder 0-1 Problemen. D.h. wir haben ein LP der Form

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b \\ & l \leq x \leq u, \end{array} \tag{5.11}$$

wobei $l \in \big(\mathbb{R} \cup \{-\infty\}\big)^n$ und $u \in \big(\mathbb{R} \cup \{+\infty\}\big)^n$ ist. Gilt $l_i = -\infty$ und $u_i = +\infty$ für ein $i \in \{1, 2, \ldots, n\}$, so heißt die zugehörige Variable x_i freie Variable. Freie Variable werden, sofern sie nicht zur Basis gehören, auf den Wert Null gesetzt. Sobald sie einmal in die Basis eintreten, werden sie diese nie wieder verlassen, vgl. auch Satz 5.18.

Betrachten wir also den Fall, dass $l_i \neq -\infty$ oder $u_i \neq +\infty$ gilt für alle $i \in \{1, 2, \ldots, n\}$. Dann kann man durch eine Variablensubstitution (5.11) immer in die Form

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b \\ & 0 \leq x \leq u, \end{array} \tag{5.12}$$

bringen mit $u \in \big(\mathbb{R}_+ \cup \{+\infty\}\big)^n$.

Eine Möglichkeit, (5.12) zu lösen besteht darin, die oberen Schranken explizit als Ungleichungen in die Nebenbedingungsmatrix aufzunehmen. Dies würde das System jedoch von der Größe $(m \times n)$ auf $(m + n) \times (2n)$ erweitern. Das ist sehr unpraktikabel, denn jede Basismatrix hat somit die Größe $(m + n) \times (m + n)$ anstatt $(m \times m)$ zuvor. Im Folgenden werden wir sehen, wie die oberen Schranken implizit im Algorithmus 5.6 berücksichtigt werden können, ohne die Größe des Problems und damit die Komplexität des Algorithmus zu erhöhen.

In der Grundversion des Simplex-Algorithmus waren die Basisvariablen diejenigen, deren Wert durch das Gleichungssystem bestimmt wurde, wohingegen die Nichtbasisvariablen auf den Wert der unteren Schranke (also auf Null) gesetzt waren. Da wir jetzt auch mit oberen Schranken zu tun haben, gibt es Nichtbasisvariable, die den Wert der unteren Schranke haben und solche, die den Wert der oberen Schranke haben.
Dementsprechend unterteilen wir die Menge N der Nichtbasisvariablen in zwei Mengen N_l und N_u. In N_l sind alle Nichtbasisvariablen, die in der gerade durchgeführten Iteration den Wert der unteren Schranke annehmen, d.h.

$$N_l = \{j \in N : \bar{x}_j = 0\},$$

und analog dazu sind in der Menge

$$N_u = \{j \in N : \bar{x}_j = u_j\}$$

alle Nichtbasisvariablen, die den Wert der oberen Schranke haben. Ist B eine Basis für (5.1), so wissen wir aus (5.4), dass

$$x_B = A_B^{-1} b - A_B^{-1} A_N x_N = A_B^{-1} b - A_B^{-1} A_{N_u} x_{Nu} - A_B^{-1} A_{N_l} x_{Nl}$$

gilt. Setzen wir $x_j = u_j$ für $j \in N_u$ und $x_j = 0$ für $j \in N_l$, dann ist B primal zulässig, falls

$$0 \leq x_B = A_B^{-1} b - A_B^{-1} A_{N_u} u_{N_u} \leq u_B. \tag{5.13}$$

Für die Zielfunktion gilt mit (5.5)

$$\begin{aligned} c^T x &= c_B^T A_B^{-1} b + (c_N^T - c_B^T A_B^{-1} A_N) x_N \\ &= c_B^T A_B^{-1} b + (c_{N_l}^T - c_B^T A_B^{-1} A_{N_l}) \cdot 0 + (c_{N_u}^T - c_B^T A_B^{-1} A_{N_u}) \cdot u_{N_u}. \end{aligned}$$

Eine Verbesserung der Zielfunktion kann also nur erreicht werden, falls

$$\begin{array}{llll} c_j - c_B^T A_B^{-1} A_{\cdot j} & < & 0 & \text{für ein } j \in N_l \text{ oder} \\ c_j - c_B^T A_B^{-1} A_{\cdot j} & > & 0 & \text{für ein } j \in N_u \end{array} \tag{5.14}$$

gilt. In anderen Worten, B ist dual zulässig, falls (5.14) nicht gilt, d.h. falls

$$\begin{array}{llll} c_j - c_B^T A_B^{-1} A_{\cdot j} & \geq & 0 & \text{für alle } j \in N_l \text{ und} \\ c_j - c_B^T A_B^{-1} A_{\cdot j} & \leq & 0 & \text{für alle } j \in N_u. \end{array}$$

Entsprechend drehen sich die Vorzeichen im Ratio-Test um, wobei zusätzlich beachtet werden muss, dass die ausgewählte Variable x_j nicht um mehr als u_j erhöht bzw. erniedrigt werden kann:

$\boxed{\text{Gilt } j \in N_l,}$ so wirkt sich eine Erhöhung von x_j (das ja jetzt den Wert Null hat) um den Wert $\gamma > 0$ auf die Basis wie folgt aus, vgl. (5.4):

$$\bar{x}_B \leftarrow \bar{x}_B - \gamma w, \tag{5.15}$$

d.h. γ muss so gewählt werden, dass

$$0 \leq \bar{x}_B - \gamma w \leq u_B$$

gilt. Aus der linken Ungleichung ergibt sich $\gamma w \leq \bar{x}_B$ bzw.

$$\gamma w_i \leq \bar{x}_{B_i} \qquad \forall i = 1, \ldots, m.$$

Da $\gamma > 0$, ist diese Ungleichung für $w_i \leq 0$ immer erfüllt (da $\bar{x}_B \geq 0$). Wir können uns daher auf die $w_i > 0$ beschränken und erhalten

$$\gamma \leq \frac{\bar{x}_{B_i}}{w_i} \qquad \forall i = 1, \ldots, m, w_i > 0,$$

oder anders gesagt

$$\gamma \leq \gamma_l := \min \left\{ \frac{\bar{x}_{B_i}}{w_i} : w_i > 0, i = 1, \ldots, m \right\}.$$

Aus der rechten Ungleichung von (5.15) ergibt sich analog

$$-\gamma w_i \leq u_{B_i} - \bar{x}_{B_i} \quad \forall i = 1, \ldots, m.$$

Betrachten wir jetzt $w_i < 0$, so folgt

$$\gamma \leq \gamma_u := \min \left\{ \frac{u_{B_i} - \bar{x}_{B_i}}{-w_i} : w_i < 0, i = 1, \ldots, m \right\}.$$

Insgesamt ergibt sich damit für γ die Darstellung

$$\gamma = \min\{u_j, \gamma_l, \gamma_u\}.$$

$\boxed{\text{Gilt } j \in N_u,}$ so wirkt sich eine Verringerung von x_j um $\gamma > 0$ auf die Basis wie folgt aus, vgl. (5.13):

$$\bar{x}_B \ \leftarrow \ \bar{x}_B + \gamma w,$$

d.h. γ muss so gewählt werden, dass

$$0 \leq \bar{x}_B + \gamma w \leq u_B \tag{5.16}$$

gilt. Aus der linken Ungleichung von (5.16) erhalten wir analog zum obigen Fall

$$\gamma \leq \gamma_l := \min \left\{ \frac{\bar{x}_{B_i}}{-w_i} : w_i < 0, i = 1, \ldots, m \right\},$$

aus der rechten Ungleichung von (5.16) ergibt sich

$$\gamma \leq \gamma_u := \min \left\{ \frac{u_{B_i} - \bar{x}_{B_i}}{w_i} : w_i > 0, i = 1, \ldots, m \right\},$$

womit wir insgesamt wieder die Bedingung

$$\gamma = \min\{u_j, \gamma_l, \gamma_u\}$$

erhalten. Damit können wir den Simplex-Algorithmus mit allgemeinen oberen Schranken wie folgt angeben:

Algorithmus 5.21 Der Simplex-Algorithmus mit oberen Schranken.

Input:

- *Ein lineares Problem der Form (5.12):*

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b \\ & 0 \le x \le u, \end{array}$$

- *Eine primal zulässige Basis B und Mengen N_l mit $\bar{x}_{N_l} = 0$ sowie N_u mit $\bar{x}_{N_u} = u_{N_u}$, wobei $0 \le \bar{x}_B = A_B^{-1} b - A_B^{-1} A_{N_u} u_{N_u} \le u_B$.*

Output:

- *Eine Optimallösung $\bar{x}$ für (5.12)*

  oder

- *Die Meldung (5.12) ist unbeschränkt.*

(1) BTRAN
Löse $A_B^T \bar{y} = c_B$.

(2) Pricing
Berechne $\bar{z}_N = c_N - A_N^T \bar{y}$.

Falls $\bar{z}_{N_l} \ge 0$ und $\bar{z}_{N_u} \le 0$, ist B optimal, **Stop**.

Andernfalls wähle $j \in N_l$ *mit* $\bar{z}_j < 0$
oder $j \in N_u$ *mit* $\bar{z}_j > 0$.

(3) FTRAN
Löse $A_B\, w = A_{\cdot j}$.

(4) Ratio-Test

I.Fall: *$j \in N_l$. Bestimme*

$$\begin{aligned} \gamma_l &= \min\left\{ \frac{\bar{x}_{B_i}}{w_i} : w_i > 0,\ i \in \{1, 2, \ldots, m\} \right\}, \\ \gamma_u &= \min\left\{ \frac{u_{B_i} - \bar{x}_{B_i}}{-w_i} : w_i < 0,\ i \in \{1, 2, \ldots, m\} \right\}, \\ \gamma &= \min\{u_j, \gamma_l, \gamma_u\}. \end{aligned}$$

Gilt $\gamma = \infty$, dann ist (5.12) unbeschränkt, **Stop**.
Gilt $\gamma = u_j$, dann gehe zu (5).
Andernfalls sei $i \in \{1, 2, \ldots, m\}$ ein Index, der γ liefert.

II.Fall: $j \in N_u$.
Bestimme

$$\begin{aligned}\gamma_l &= \min\left\{\frac{\bar{x}_{B_i}}{-w_i} : w_i < 0,\ i \in \{1,2,\ldots,m\}\right\},\\ \gamma_u &= \min\left\{\frac{u_{B_i} - \bar{x}_{B_i}}{w_i} : w_i > 0,\ i \in \{1,2,\ldots,m\}\right\},\\ \gamma &= \min\{u_j, \gamma_l, \gamma_u\}.\end{aligned}$$

Gilt $\gamma = \infty$, dann ist (5.12) unbeschränkt, **Stop**.
Gilt $\gamma = u_j$, dann gehe zu (5).
Andernfalls sei $i \in \{1,2,\ldots,m\}$ ein Index, der γ liefert.

(5) Update

I.Fall: $j \in N_l$.
Setze $\bar{x}_B = \bar{x}_B - \gamma w$
$N_l = N_l \setminus \{j\}$

Falls $\gamma = u_j$, setze $N_u = N_u \cup \{j\},\ \bar{x}_j = u_j$, und gehe zu (2).
Falls $\gamma = \gamma_l$, setze $N_l = N_l \cup \{B_i\},\ B_i = j,\ \bar{x}_j = \gamma$, und gehe zu (1).
Falls $\gamma = \gamma_u$, setze $N_u = N_u \cup \{B_i\},\ B_i = j,\ \bar{x}_j = \gamma$, und gehe zu (1).

II.Fall: $j \in N_u$.
Setze $\bar{x}_B = \bar{x}_B + \gamma w$
$N_u = N_u \setminus \{j\}$

Falls $\gamma = u_j$, setze $N_l = N_l \cup \{j\}, \bar{x}_j = 0$, und gehe zu (2).
Falls $\gamma = \gamma_l$, setze $N_l = N_l \cup \{B_i\}, B_i = j, \bar{x}_j = u_j - \gamma$, gehe zu (1).
Falls $\gamma = \gamma_u$, setze $N_u = N_u \cup \{B_i\}, B_i = j, \bar{x}_j = u_j - \gamma$, gehe zu (1).

Korrektheit des Simplex-Algorithmus mit oberen Schranken beweist man analog zu Satz 5.8.
Falls in Problem (5.12) $u_i = +\infty$ gilt für alle i, so ist im obigen Algorithmus durchgehend $N_u = \emptyset$, Fall II kommt daher nie vor, und man erhält die Grundversion (Algorithmus 5.6) als Spezialfall von Algorithmus 5.21.

Für eine duale Version des Simplex-Algorithmus mit oberen Schranken siehe die Übungen.

Beispiel 5.22 *Betrachten wir folgendes LP mit oberen Schranken:*

$$\begin{array}{lrcrcl} \min & -x_1 & - & 2x_2 & & \\ \text{s.t.} & x_1 & + & x_2 & \leq & 3 \\ & 2x_1 & + & x_2 & \leq & 5 \\ & 0 & \leq & x_1 & \leq & 2 \\ & 0 & \leq & x_2 & \leq & 2 \end{array}$$

Mit Schlupfvariablen lautet das LP:

$$\begin{array}{lrcrcrcrcl} \min & -x_1 & - & 2x_2 & & & & & & \\ \text{s.t.} & x_1 & + & x_2 & + & x_3 & & & = & 3 \\ & 2x_1 & + & x_2 & & & + & x_4 & = & 5 \\ & & & & 0 & \leq & x_1, & x_2 & \leq & 4 \\ & & & & & & x_3, & x_4 & \geq & 0 \end{array}$$

Wir starten mit der Basis $B = (3,4)$ und den Mengen $N_l = (1,2)$ sowie $N_u = \emptyset$. Diese Basis ist primal zulässig, da $\bar{x}_B = (3,5)^T \geq 0$.

<u>Iteration 1:</u>

BTRAN:

$$\bar{y}^T \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = (0,0) \qquad \Longrightarrow \qquad \bar{y} = (0,0)^T.$$

Pricing: $\bar{z}_N = \binom{-1}{-2} - A_N^T \binom{0}{0} = \binom{-1}{-2}$. *Wähle $j = 2 \in N_l$ mit $z_2 = -2 < 0$.*

FTRAN: $Iw = \binom{1}{1}$.

Ratio-Test: *I.Fall, $j \in N_l$.*

$$\begin{aligned} \gamma_l &= \min\left\{\frac{3}{1}, \frac{5}{1}\right\} = 3 \\ \gamma_u &= +\infty \quad \textit{(da kein } w_i < 0.\textit{)} \end{aligned}$$

Daher gilt

$$\gamma = \min\{u_2, \gamma_l\} = \min\{2,3\} = 2.$$

Update:

$$\bar{x}_B = \begin{pmatrix} 3 \\ 5 \end{pmatrix} - 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix}.$$

$N_l = (1)$, $N_u = (2)$, $x_2 = 2$, die Basis B bleibt unverändert.
Hier wurde also (anders als in der Grundversion des Simplexalgorithmus) nicht eine Nichtbasisvariable gegen eine Basisvariable getauscht, sondern eine Variable aus N_l wandert in die Menge N_u. Dadurch bleibt die Basis unverändert, daher kann in der nächsten Iteration das Lösen des Gleichungssystems in BTRAN übersprungen werden: Die Lösung $\bar{y}$ bleibt die gleiche wie in Iteration 1.

Iteration 2:
Pricing: $\bar{z}_N = \binom{-1}{-2} - A_N^T \binom{0}{0} = \binom{-1}{-2}$. *Es gilt* $z_{N_l} = -1 < 0$, *wähle daher* $j = 1 \in N_l$.
FTRAN: $Iw = \binom{1}{2}$.
Ratio-Test: *I.Fall*, $j \in N_l$.

$$\gamma_l = \min\left\{\frac{1}{1}, \frac{3}{2}\right\} = 1$$

Daher

$$\gamma = \min\{u_1, \gamma_l\} = \min\{2, 1\} = 1.$$

Update:

$$\bar{x}_B = \begin{pmatrix} 1 \\ 3 \end{pmatrix} - 1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix}.$$

$N_l = \emptyset \cup \{B_i\} = (3)$, $N_u = (2)$, $x_1 = 1$, $B = (1, 4)$.

Iteration 3:
BTRAN:

$$\bar{y}^T \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = (-1, 0) \qquad \Longrightarrow \qquad \bar{y} = (-1, 0)^T.$$

Pricing:

$$\bar{z}_N = \begin{pmatrix} 0 \\ -2 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Es gilt also

$$\begin{aligned} z_{N_l} &\geq 0 \\ z_{N_u} &\leq 0 \end{aligned}$$

Daher ist die Basis dual zulässig und daher optimal.
Die Lösung des LPs lautet also $x_1 = 1, x_2 = 2, x_3 = 0, x_4 = 1$.

5.6 Sensitivitätsanalyse

Wir sind bisher davon ausgegangen, dass die Daten A, b und c eines linearen Problems der Form (5.1) fest vorgegeben waren. Häufig ist es jedoch der Fall, dass sich diese Daten im Laufe der Zeit ändern, wenn Bedingungen und/oder Variablen hinzukommen und damit nachoptimiert werden muss. Gehen wir davon aus, dass wir ein lineares Problem der Form (5.1)

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0. \end{array}$$

bereits gelöst haben und eine optimale Basis B mit der Lösung $\bar{x}_B = A_B^{-1} b$ und $\bar{x}_N = 0$ kennen. Wir wollen nun folgende Änderungen des linearen Problems betrachten und uns überlegen, wie wir ausgehend von B möglichst schnell eine Optimallösung des veränderten Problems finden können:

(i) Änderung der Zielfunktion c.

(ii) Änderung der rechten Seite b.

(iii) Änderung eines Eintrags in der Matrix A.

(iv) Hinzufügen einer neuen Variablen.

(v) Hinzufügen einer neuen Nebenbedingung.

(i) Änderung der Zielfunktion

Der Zielfunktionsvektor c ändert sich zu $\tilde{c}$. Dann ist B weiterhin primal zulässig, da $\bar{x}_B = A_B^{-1} b$ unverändert bleibt. Wegen der Änderung der Zielfunktion ändern sich die reduzierten Kosten zu

$$\tilde{z}_N = \tilde{c}_N - A_N^T A_B^{-T} \tilde{c}_B.$$

<u>1. Fall: $\tilde{z}_N \geq 0$.</u> Die Basis bleibt demnach auch dual zulässig, somit bleibt B optimal. Der Zielfunktionswert ändert sich eventuell.
<u>2. Fall: $\exists i \in N : \tilde{z}_i < 0$.</u> Die Basis B ist nicht mehr dual zulässig, also ist B nicht mehr optimal. Der primale Simplex-Algorithmus kann mit B als primal zulässiger Basis gestartet werden.

(ii) Änderung der rechten Seite

Statt b haben wir $\tilde{b}$. Dann ist B weiterhin dual zulässig, da die reduzierten Kosten $c_N - A_N^T A_B^{-T} c_B$ unverändert bleiben. Aufgrund der Änderung der rechten Seite ändert sich die Basislösung zu

$$\tilde{x}_B = A_B^{-1}\tilde{b}.$$

1. Fall: $\tilde{x}_B \geq 0$. Die Basis bleibt demnach primal zulässig, somit bleibt B optimal. Die Basislösung und der Zielfunktionswert ändern sich.
2. Fall: $\exists i \in B : \tilde{x}_i < 0$. Die Basis B ist nicht mehr primal zulässig, also ist B nicht mehr optimal. Der duale Simplex-Algorithmus kann mit B als dual zulässiger Basis gestartet werden.

Man kann im Falle einer Änderung der rechten Seite sehr leicht zeigen, dass der Optimalwert eines LPs stetig von der rechten Seite abhängt. Betrachten wir dazu (für fix vorgegebenes $A \in \mathbb{R}^{m\times n}$ und $c \in \mathbb{R}^n$) folgende Menge

$$\mathcal{Z} := \{b \in \mathbb{R}^m : \mathcal{P}^=(A, b) \neq \emptyset\}.$$

Da $0 \in \mathcal{Z}$, ist offensichtlich $\mathcal{Z} \neq \emptyset$. Die Funktion $v : \mathcal{Z} \to \mathbb{R} \cup \{+\infty\}$,

$$v(b) := \inf\{c^T x : Ax = b, x \geq 0\}$$

heißt Perturbationsfunktion des linearen Problems

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0. \end{array}$$

Wir zeigen nun, dass v eine stetige Funktion ist: Sei

$$\operatorname{epi}(v) := \{(b, \beta) \in \mathcal{Z} \times \mathbb{R} : v(b) \leq \beta\}$$

der Epigraph von v.

Satz 5.23 $\operatorname{epi}(v)$ *ist ein polyedrischer Kegel.*

Beweis. $\operatorname{epi}(v)$ ist die Projektion des Polyeders

$$\{(x, b, \beta) \in \mathbb{R}^n \times \mathcal{Z} \times \mathbb{R} : Ax = b, x \geq 0, c^T x \leq \beta\}$$

auf die (b, β)–Koordinaten, und daher ein Polyeder. Außerdem überlegt man sich leicht, dass epi(v) ein Kegel ist. $\square$

Insbesondere ist also epi(v) eine konvexe Menge. Nach Satz 2.30 ist deshalb v eine konvexe Funktion. Wegen Satz 2.31 ist die Perturbationsfunktion also stetig im Inneren von $\mathcal{Z}$.

Der Optimalwert eines LPs ändert sich also im Inneren von $\mathcal{Z}$ stetig in der Änderung der rechten Seite.

(iii) Änderung eines Eintrags in der Matrix A in einer Nichtbasisspalte

Die Basis B bleibt primal zulässig, da $\bar{x}_B = A_B^{-1} b$ unverändert bleibt. Aufgrund der Änderung in der Nichtbasisspalte $A_{\cdot j}$ zu $\tilde{A}_{\cdot j}$ $(j \in N)$ ändern sich die reduzierten Kosten bezüglich j zu

$$\tilde{z}_j = c_j - c_B^T A_B^{-1} \tilde{A}_{\cdot j}$$

<u>1. Fall: $\tilde{z}_j \geq 0$.</u> Die Basis bleibt demnach dual zulässig, somit bleibt B optimal. Der Zielfunktionswert ändert sich nicht.
<u>2. Fall: $\tilde{z}_j < 0$.</u> Die Basis B ist nicht mehr dual zulässig, also ist B nicht mehr optimal. Der primale Simplex-Algorithmus kann mit B als primal zulässiger Basis gestartet werden.

Bei einer Änderung der Matrix in einer Basisspalte ist keine Vorhersage möglich, da sich sowohl die Basislösung $\bar{x}_B$ als auch die reduzierten Kosten ändern.

(iv) Hinzufügen einer neuen Variablen

Eine neue Variable x_{n+1} mit Vorzeichenbeschränkung $x_{n+1} \geq 0$, Zielfunktionskoeffizient $c_{n+1} \in \mathbb{R}$ und Spalte $A_{\cdot,n+1} \in \mathbb{R}^m$ wird zum Problem hinzugefügt.
Wir fügen x_{n+1} zu den Nichtbasisvariablen hinzu, also $\tilde{N} = N \cup \{n+1\}$. Die Basis bleibt primal zulässig, da $\bar{x}_B = A_B^{-1} b$ unverändert bleibt. Für x_{n+1} erhalten wir die reduzierten Kosten

$$\bar{z}_{n+1} = c_{n+1} - c_B^T A_B^{-1} A_{\cdot,n+1}.$$

1. Fall: $\bar{z}_{n+1} \geq 0$. Die Basis bleibt demnach dual zulässig, somit bleibt B optimal. Der Zielfunktionswert ändert sich nicht.
2. Fall: $\bar{z}_{n+1} < 0$. Die Basis B ist nicht mehr dual zulässig, also ist B nicht mehr optimal. Der primale Simplex-Algorithmus kann mit B als primal zulässiger Basis gestartet werden. Im ersten Schritt wird die neue Variable in die Basis aufgenommen, da die reduzierten Kosten für alle anderen Variablen unverändert (d.h. nicht negativ) bleiben.

(v) Hinzufügen einer neuen Nebenbedingung

Wir fügen dem gegebenen Problem

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array}$$

eine neue Nebenbedingung

$$A_{m+1,\cdot}x \leq b_{m+1},$$

bzw. mit neuer Schlupfvariable

$$A_{m+1,\cdot}x + x_{n+1} = b_{m+1}, \qquad x_{n+1} \geq 0$$

hinzu. Wir unterscheiden nun folgende Fälle:

1.Fall: Die Optimallösung $\bar{x}$ erfüllt auch die neue Nebenbedingung. Dann ist $\bar{x}$ auch für das erweiterte Problem eine Optimallösung. Ist $A_{m+1,\cdot}$ linear abhängig von den Zeilen von A, so ist die neue Zeile irrelevant und kann gelöscht werden.

Ist $A_{m+1,\cdot}$ linear unabhängig von den Zeilen von A, so ist $\bar{x}$ eine degenerierte Basislösung des erweiterten Systems. Die Schlupfvariable $\bar{x}_{n+1}$ kann mit dem Wert Null in die Basis aufgenommen werden. Die neue Basis hat nun die Kardinalität $m + 1$.

2.Fall: Im Allgemeinen jedoch wird die neue Nebenbedingung nicht erfüllt sein, d.h. es gilt

$$A_{m+1,\cdot}\,\bar{x} > b_{m+1}. \tag{5.17}$$

Wir erweitern die Basis um die Schlupfvariable x_{n+1}: $\tilde{B} = B \cup \{n+1\}$. Die neue Basismatrix zur Basis $\tilde{B}$ hat die Form

$$\tilde{A}_{\tilde{B}} = \begin{pmatrix} A_B & 0 \\ A_{m+1,B} & 1 \end{pmatrix} \in \mathbb{R}^{(m+1)\times(m+1)}.$$

Die zugehörige inverse Basismatrix lautet

$$\tilde{A}_{\tilde{B}}^{-1} = \begin{pmatrix} A_B^{-1} & 0 \\ -A_{m+1,B} \cdot A_B^{-1} & 1 \end{pmatrix} \in \mathbb{R}^{(m+1)\times(m+1)},$$

und es gilt

$$\begin{aligned} \bar{x}_{\tilde{B}} &= \begin{pmatrix} A_B^{-1} & 0 \\ -A_{m+1,B} \cdot A_B^{-1} & 1 \end{pmatrix} b = \begin{pmatrix} A_B^{-1} b \\ -A_{m+1,B} \cdot A_B^{-1} b + b_{m+1} \end{pmatrix} \\ &= \begin{pmatrix} \bar{x}_B \\ -A_{m+1,B}\, \bar{x}_B + b_{m+1} \end{pmatrix}, \end{aligned}$$

wobei $\bar{x}_B \geq 0$ und $-A_{m+1,B}\, \bar{x}_B + b_{m+1} < 0$ ist wegen (5.17), d.h. $\tilde{B}$ ist nicht primal zulässig, jedoch ist sie dual zulässig, da

$$\begin{aligned} c_N^T - c_{\tilde{B}}^T \tilde{A}_{\tilde{B}}^{-1} \begin{pmatrix} A_N \\ A_{m+1,N} \end{pmatrix} &= c_N^T - (c_B^T, 0) \begin{pmatrix} A_B^{-1} & 0 \\ -A_{m+1,B} \cdot A_B^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_N \\ A_{m+1,N} \end{pmatrix} \\ &= c_N^T - (c_B^T, 0) \begin{pmatrix} A_B^{-1} A_N \\ -A_{m+1,B} \cdot A_B^{-1} A_N + A_{m+1,N} \end{pmatrix} \\ &= c_N^T - c_B^T A_B^{-1} A_N \\ &\geq 0. \end{aligned}$$

Wir können daher mit dem dualen Simplex-Algorithmus starten. Die erste die Basis verlassende Variable ist x_{n+1}.

Für das neue Problem erhalten wir entweder eine Optimallösung oder die Meldung, dass

$$\mathcal{P}^{=}\left(\begin{bmatrix} A \\ A_{m+1,\cdot} \end{bmatrix}, \begin{bmatrix} b \\ b_{m+1} \end{bmatrix} \right) = \emptyset.$$

Im letzteren Fall hat der Halbraum $A_{m+1,\cdot}x \leq b_{m+1}$ einen leeren Schnitt mit $\mathcal{P}^{=}(A, b)$.

Kapitel 6

Die Ellipsoidmethode und Polynomialität für rationale LPs

Wie wir gesehen haben, kann man Beispiele linearer Probleme konstruieren, bei denen der Simplex-Algorithmus *alle* Ecken absucht. Bei Problemen vom Klee-Minty-Typ (Beispiel 5.13) beispielsweise benötigt der Simplex-Algorithmus 2^n Iterationen, was zu sehr langen Rechenzeiten führt. Man kann zeigen, dass im Durchschnitt die Laufzeit des Simplex-Algorithmus gut ist, es ist aber bis heute eine offene Frage, ob es Auswahlregeln für die Schritte 2 und 4 des Simplex-Algorithmus gibt, die dieses Phänomen beweisbar vermeiden, d.h. für die man keine Beispiele konstruieren kann, die den Simplex-Algorithmus zu “langen” Laufzeiten zwingen.
In diesem Kapitel soll untersucht werden, wie man die Laufzeit eines Algorithmus quantifizieren kann, danach wird ein Algorithmus zum Lösen linearer Probleme vorgestellt, dessen Laufzeit “polynomial” ist. Diese wichtige Erkenntnis wurde erstmals mit Hilfe der Ellipsoidmethode von Khachiyan [Kh79] gezeigt und hat eine aktive Forschungstätigkeit zur Konstruktion von polynomialen Algorithmen für LP ausgelöst, die unter anderem sehr effiziente polynomiale Verfahren für LP hervorgebracht, insbesondere Innere-Punkte-Verfahren [Wr97].

6.1 Polynomiale Algorithmen

Um das Laufzeitverhalten eines Algorithmus exakt beschreiben zu können, ist es notwendig, Maschinenmodelle, Kodierungsvorschriften, Algorithmus-

beschreibungen etc. exakt einzuführen. Wir wollen hier die wichtigsten Konzepte dieser Theorie informell für den Spezialfall der linearen Probleme einführen. Eine mathematisch exakte Darstellung der Theorie findet sich in dem grundlegenden Buch von Garey und Johnson [GJ79].
Zunächst müssen wir beachten, dass jede Zahl in unserem Maschinenmodell endlich kodierbar sein muss. Da es kein Kodierungsschema gibt, das reelle Zahlen durch endlich viele Symbole beschreibt, beschränken wir uns auf das Rechnen mit rationalen Zahlen. Haben wir eine Ungleichung

$$a^T x \leq \alpha,$$

mit $a \in \mathbb{Q}^n$, $\alpha \in \mathbb{Q}$, so können wir auf einfache Weise das kleinste gemeinsame Vielfache p der Nenner der Komponenten von a und des Nenners von α bestimmen. Die Ungleichung

$$pa^T x \leq p\alpha$$

hat dann ganzzahlige Koeffizienten und ist offenbar äquivalent zu $a^T x \leq \alpha$. Der Fall rationaler Daten lässt sich also direkt auf den Fall ganzzahliger Daten reduzieren. Es ist zwar häufig einfacher, unter der Voraussetzung ganzzahliger Daten zu rechnen, wir setzen aber dennoch für dieses Kapitel voraus:

Annahme 6.1 *Alle Daten der betrachteten linearen Probleme sind rational, d.h. für jedes LP der Form* $\max\{c^T x : Ax \leq b\}$ *gilt* $c \in \mathbb{Q}^n$, $A \in \mathbb{Q}^{m \times n}$ *und* $b \in \mathbb{Q}^m$.

Da Computer üblicherweise mit Binärcodes arbeiten, nehmen wir weiter an, dass alle vorkommenden Zahlen binär codiert sind. Die binäre Darstellung einer ganzen Zahl n benötigt $\lceil \log_2(|n|+1) \rceil$ Stellen (Bits) und eine Stelle für das Vorzeichen (für $n = 0$ ergibt sich 1 Bit, was Sinn macht, da man kein Vorzeichen braucht). Das führt zu folgender Definition:

Definition 6.2 *Für* $n \in \mathbb{Z}$ *heißt*

$$\langle n \rangle := \lceil \log_2(|n| + 1) \rceil + 1$$

die Kodierungslänge *von* n. *Für jede rationale Zahl* $r = \frac{p}{q}$ *in teilerfremder Darstellung mit* $q > 0$ *ist die Kodierungslänge*

$$\langle r \rangle = \langle p \rangle + \langle q \rangle.$$

Die Kodierungslänge einer Matrix $A = (a_{ij}) \in \mathbb{Q}^{m\times n}$ *(analog für Vektoren) ist gegeben durch*

$$\langle A\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n}\langle a_{ij}\rangle.$$

Im folgenden Lemma werden einige nützliche Eigenschaften der so definierten Kodierungslänge gezeigt.

Lemma 6.3

(a) *Für jede Zahl $r \in \mathbb{Q}$ gilt:* $|r| \le 2^{\langle r\rangle -1} - 1$.

(b) *Für je zwei Zahlen $r, s \in \mathbb{Q}$ gilt:* $\langle rs\rangle \le \langle r\rangle + \langle s\rangle$.

(c) *Für jeden Vektor $x \in \mathbb{Q}^n$ gilt:* $\|x\|_2 \le \|x\|_1 \le 2^{\langle x\rangle - n} - 1$.

(d) *Für jede Matrix $A \in \mathbb{Q}^{n\times n}$ gilt:* $|\det A| \le 2^{\langle A\rangle - n^2} - 1$.

Beweis. (a) und (b): Übung.
(c): Sei $x = (x_1, \ldots, x_n)^T \in \mathbb{Q}^n$. Dann gilt wegen (a):

$$1 + \|x\|_1 = 1 + \sum_{i=1}^{n}|x_i| \le \prod_{i=1}^{n}(1 + |x_i|) \le \prod_{i=1}^{n} 2^{\langle x_i\rangle - 1} = 2^{\langle x\rangle - n}.$$

Die Ungleichung $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2} \le \sum_{i=1}^n |x_i| = \|x\|_1$ ist trivial.
(d): Es gilt

$$1 + |\det A| \le 1 + \prod_{j=1}^{n}\|A_{\cdot j}\|_2 \le \prod_{j=1}^{n}(1 + \|A_{\cdot j}\|_2) \le \prod_{j=1}^{n} 2^{\langle A_{\cdot j}\rangle - n} = 2^{\langle A\rangle - n^2},$$

wobei die erste Ungleichung aus der so genannten Hadamard-Ungleichung[1] folgt, die zweite Ungleichung ist trivial, und die dritte Ungleichung gilt wegen (c). □

Als nächstes überlegen wir, wie man die Laufzeit eines Algorithmus messen kann. Für unsere Zwecke genügt es zunächst festzulegen, dass wir die

[1] Die Hadamard-Ungleichung lautet: $|\det A| \le \prod_{j=1}^n \|A_{\cdot j}\|_2$.

Anzahl der elementaren Rechenschritte zählen wollen, wobei elementare Rechenschritte *Addition*, *Subtraktion*, *Multiplikation*, *Division* und *Vergleich* von ganzen oder rationalen Zahlen sind. Dies reicht aber noch nicht aus, da z.B. die Multiplikation großer Zahlen länger dauert als die Multiplikation kleiner Zahlen. Wir müssen also die Zahlengrößen mitberücksichtigen, was auf folgende Definition führt:

Definition 6.4 *Die* Laufzeit eines Algorithmus A *zur Lösung eines Problems* Π *(kurz* $L_{\mathsf{A}}(\Pi)$*) ist die Anzahl der elementaren Rechenschritte, die während der Ausführung des Algorithmus durchgeführt werden, multipliziert mit der Kodierungslänge der (bezüglich ihrer Kodierungslänge) größten Zahl, die während der Ausführung des Algorithmus aufgetreten ist.*

Nun können wir definieren, was unter polynomialer Laufzeit eines Algorithmus zu verstehen ist:

Definition 6.5 *Sei* A *ein Algorithmus zur Lösung einer Klasse von Problemen (z.B. der Klasse aller linearen Optimierungsprobleme). Die Elemente dieser Klasse (z.B. spezielle LPs) bezeichnen wir mit* Π.

(a) Die Funktion $f : \mathbb{N} \to \mathbb{N}$ *definiert durch*

$$f_{\mathsf{A}}(n) := \max\{L_{\mathsf{A}}(\Pi) : \begin{array}{l} \textit{die Kodierungslänge der Daten zur} \\ \textit{Beschreibung von } \Pi \textit{ ist höchstens } n\} \end{array}$$

heißt Laufzeitfunktion *von* A.

(b) Der Algorithmus A *hat eine* polynomiale Laufzeit *(kurz:* A *ist ein* polynomialer Algorithmus*), wenn es ein Polynom* $p : \mathbb{N} \to \mathbb{N}$ *gibt, so dass*

$$f_{\mathsf{A}}(n) \leq p(n) \quad \forall n \in \mathbb{N}.$$

Polynomiale Algorithmen sind also solche Verfahren, deren Schrittzahl multipliziert mit der Kodierungslänge der größten auftretenden Zahl durch ein Polynom in der Kodierungslänge beschränkt werden kann. Für die Klasse der linearen Probleme der Form $\max\{c^T x : Ax \leq b\}$ muss also die Laufzeit eines Algorithmus durch ein Polynom in $\langle A\rangle + \langle b\rangle + \langle c\rangle$ beschränkt werden können, wenn er polynomial sein soll.
Wie das Klee-Minty-Beispiel 5.13 zeigt, ist der Simplexalgorithmus kein polynomialer Algorithmus zur Lösung linearer Optimierungsprobleme!

6.2 Reduktion von LPs auf Zulässigkeitsprobleme

Die Ellipsoidmethode ist eigentlich kein Optimierungsverfahren, sondern eine Methode, die in einem gegebenen volldimensionalen Polytop einen Punkt findet bzw. feststellt, dass das Polyeder leer ist. Wir müssen daher allgemeine lineare Optimierungsprobleme auf diesen Fall zurückführen. Das ist auf zwei Arten möglich: Entweder, indem man Dualität nutzt, oder durch die so genannte Binärsuche.

Reduktion mittels LP-Dualität

Betrachten wir ein LP der Form

$$\begin{array}{ll} \max & c^T x \\ \text{s.t.} & Ax \leq b \\ & x \geq 0. \end{array}$$

Das duale LP dazu ist

$$\begin{array}{ll} \min & b^T y \\ \text{s.t.} & A^T y \geq c \\ & y \geq 0. \end{array}$$

Aus dem starken Dualitätssatz 4.13 wissen wir, dass die beiden Probleme genau dann Optimallösungen mit gleichem Zielfunktionswert haben, wenn beide Probleme zulässige Punkte besitzen.
Daraus folgt, dass jedes Element $\binom{x}{y}$ des Polyeders $\mathcal{P}$, das durch

$$\begin{pmatrix} -c^T & b^T \\ A & 0 \\ 0 & -A^T \\ -I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} 0 \\ b \\ -c \\ 0 \\ 0 \end{pmatrix} \tag{6.1}$$

definiert ist, eine Optimallösung x des primalen und eine Optimallösung y des dualen Problems bestimmt. Zur Lösung eines LPs genügt es also, einen Punkt in (6.1) zu finden.
Der Vorteil ist die Einfachheit des Vorgehens. Der Nachteil ist, dass das entstehende Polyeder i.A. nicht volldimensional ist, auch wenn der Zulässigkeitsbereich des primalen Problems volldimensional ist.

Die Ellipsoidmethode benötigt aber volldimensionale Polyeder. Wir werden in Kürze sehen, wie man ein äquivalentes volldimensionales Zulässigkeitsproblem erhalten kann.

Reduktion mittels Binärsuche

Der obige Trick, primales und duales LP zusammenzufassen, bläht das Ungleichungssystem auf. Eine alternative Methode, die diese Dimensionsvergrößerung vermeidet und Volldimensionalität erhält, ist die binäre Suche, die wesentlich auf der Abschätzung der "Größe" der Ecken von Polyedern beruht.

Satz 6.6 *Für jede Ecke $v = (v_1, \ldots, v_n)^T$ eines Polyeders $\mathcal{P}$ der Form $\mathcal{P}(A, b)$, $\mathcal{P}^=(A, b)$ oder $\mathcal{P} = \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}$ mit $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$ gilt:*

(a) Nach vollständigem Kürzen ist der Absolutbetrag des Zählers von v_i höchstens $2^{3\langle A\rangle+2\langle b\rangle-3n^2}$ und der Absolutbetrag des Nenners von v_i höchstens $2^{3\langle A\rangle+\langle b\rangle-3n^2}$.

(b) $|v_i| \leq 2^{2\langle A\rangle+\langle b\rangle-2n^2}$ (für alle $i = 1, \ldots, n$).

(c) Falls $A \in \mathbb{Z}^{m \times n}$, so gilt $|v_i| \leq 2^{\langle A\rangle+\langle b\rangle-n^2}$ (für alle $i = 1, \ldots, n$).

Beweis. Wir betrachten den Fall $\mathcal{P} = \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}$.
Ist v eine Ecke von $\mathcal{P}$, so gibt es eine reguläre $(n \times n)$-Teilmatrix D von $\binom{A}{-I}$ und einen entsprechenden Teilvektor d von $\binom{b}{0}$, so dass v die eindeutige Lösung des Gleichungssystems $Dx = d$ ist. Nach der Cramerschen Regel gilt dann für $i = 1, \ldots, n$:

$$v_i = \frac{\det D_i}{\det D},$$

wobei D_i aus D dadurch entsteht, dass die i-te Spalte von D durch den Vektor d ersetzt wird.
Hat D Zeilen, die negative Einheitsvektoren sind, so bezeichnen wir mit $\overline{D}$ die Matrix, die durch Streichen dieser Zeilen und der Spalten, in denen sich die Elemente -1 befinden, entsteht. Aufgrund des Determinantenentwicklungssatzes gilt $|\det D| = |\det \overline{D}|$. Außerdem ist $\overline{D}$ eine Teilmatrix von A. Daraus folgt mit Lemma 6.3 (d):

$$|\det D| = |\det \overline{D}| \leq 2^{\langle \overline{D}\rangle-n^2} \leq 2^{\langle A\rangle-n^2}.$$

Analog folgt für den Zähler:

$$|\det D_i| \leq 2^{\langle A \rangle + \langle b \rangle - n^2}.$$

$\boxed{\text{(c):}}$ Ist $|\det D| \geq 1$, und das gilt z.B. wenn $A \in \mathbb{Z}^{m \times n}$, dann gilt nach dem bisher gezeigten:

$$|v_i| \leq |\det D_i| \leq 2^{\langle A \rangle + \langle b \rangle - n^2}.$$

$\boxed{\text{(b):}}$ Ist $|\det D| < 1$, so müssen wir $|\det D|$ nach unten abschätzen. Sei $q = \prod_{i,j} q_{ij}$ das Produkt der (positiven) Nenner der Elemente $d_{ij} = \frac{p_{ij}}{q_{ij}}$ von D. Dann gilt $q|\det D| \in \mathbb{N}$, also $q|\det D| \geq 1$ und somit

$$\frac{1}{|\det D|} \leq q.$$

Weiter kommen alle nichttrivialen Nenner $q_{ij} > 1$ auch als Nenner in A vor und daher gilt $\langle q \rangle \leq \langle A \rangle - n^2$ (die n^2 Zähler von A mit Vorzeichen sind bei A zusätzlich zu codieren). Unter Verwendung von Lemma 6.3 (a) und (b) folgt

$$q \leq 2^{\langle q \rangle - 1} \leq 2^{\langle A \rangle - n^2}.$$

Daraus ergibt sich

$$|v_i| = \frac{|\det D_i|}{|\det D|} \leq q|\det D_i| \leq 2^{2\langle A \rangle + \langle b \rangle - 2n^2}.$$

$\boxed{\text{(a):}}$ Sei q wie in (b) und sei analog p das Produkt der (positiven) Nenner der Elemente von A und b. Dann enthält p auch die Nenner aller Elemente von D_i (die Einträge -1 haben trivialen Nenner 1). Dann ist wieder $p|\det D_i| \in \mathbb{Z}$ und wir erhalten $\langle p \rangle \leq \langle A \rangle + \langle b \rangle - n^2$, also

$$|p| \leq 2^{\langle p \rangle - 1} \leq 2^{\langle A \rangle + \langle b \rangle - n^2}.$$

Nun gilt $p_i := pq \det D_i \in \mathbb{Z}$, $d := pq \det D \in \mathbb{Z}$ sowie

$$v_i = \frac{\det D_i}{\det D} = \frac{pq \det D_i}{pq \det D} = \frac{p_i}{d}.$$

Der Zähler ist beschränkt durch

$$|p_i| \leq |p||q||\det D_i| \leq 2^{2\langle A \rangle + \langle b \rangle - 2n^2} 2^{\langle A \rangle + \langle b \rangle - n^2} \leq 2^{3\langle A \rangle + 2\langle b \rangle - 3n^2},$$

der Nenner durch

$$|d| \leq |p||q|| \det D| \leq 2^{2\langle A\rangle+\langle b\rangle-2n^2} 2^{\langle A\rangle-n^2} \leq 2^{3\langle A\rangle+\langle b\rangle-3n^2},$$

□

Aus dem letzten Satz folgt sofort:

Satz 6.7 *Das lineare Problem*

$$\begin{array}{ll} \max & c^T x \\ \text{s.t.} & Ax \leq b \\ & x \geq 0 \end{array} \tag{6.2}$$

hat eine Optimallösung genau dann, wenn die beiden linearen Probleme

$$\begin{array}{ll} \max & c^T x \\ \text{s.t.} & Ax \leq b \\ & x \geq 0 \\ & x_i \leq 2^{2\langle A\rangle+\langle b\rangle-2n^2}, \quad i = 1, \dots, n \end{array} \tag{6.3}$$

und

$$\begin{array}{ll} \max & c^T x \\ \text{s.t.} & Ax \leq b \\ & x \geq 0 \\ & x_i \leq 2^{2\langle A\rangle+\langle b\rangle-2n^2} + 1, \quad i = 1, \dots, n \end{array} \tag{6.4}$$

eine Optimallösung haben und die Werte der Optimallösungen übereinstimmen. Die Zielfunktionswerte von (6.3) und (6.4) stimmen genau dann nicht überein, wenn (6.2) unbeschränkt ist. (6.2) ist genau dann unzulässig, wenn (6.3) oder (6.4) unzulässig ist.

Wir haben damit das lineare Problem (6.2) über einem allgemeinen Polyeder auf das Lösen zweier LPs über Polytopen zurückgeführt. Wir müssen also im Prinzip nur zeigen, wie man LPs über Polytopen löst.

Satz 6.8 *Ist $\mathcal{P} \neq \emptyset$ ein Polytop der Form $\mathcal{P}(A, b)$, $\mathcal{P}^=(A, b)$ oder $\mathcal{P} = \{x : Ax \leq b, x \geq 0\}$ mit $A \in \mathbb{Q}^{m\times n}, b \in \mathbb{Q}^m$, so kann für $c \in \mathbb{Q}^n$ das lineare Problem*

$$\begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in \mathcal{P} \end{array}$$

nur Optimalwerte in der endlichen Menge

$$\mathcal{S} = \left\{ \tfrac{p}{q} \in \mathbb{Q} : |p| \leq n 2^{3\langle A\rangle + 2\langle b\rangle + 2\langle c\rangle - 3n^2 - n},\ 1 \leq q \leq 2^{3\langle A\rangle + \langle b\rangle + \langle c\rangle - 3n^2 - n} \right\} \tag{6.5}$$

annehmen.

Beweis. Es genügt, die Werte $c^T v$ für Ecken v von $\mathcal{P}$ abzuschätzen. Sei also $v = (v_1, \ldots, v_n)^T$ eine Ecke von $\mathcal{P}$. Wie im Beweis von Satz 6.6, a) haben die Komponenten v_i von v eine Darstellung der Form

$$v_i = \frac{\det D_i}{\det D} =: \frac{p_i}{d}, \quad p_i, d \in Z,$$

wobei wie im Beweis von Satz 6.6 gilt

$$|p_i| \leq 2^{3\langle A\rangle + 2\langle b\rangle - 3n^2}, \quad |d| \leq 2^{3\langle A\rangle + \langle b\rangle - 3n^2}.$$

Sei nun $c = (\frac{s_1}{t_1}, \ldots, \frac{s_n}{t_n})^T \in \mathbb{Q}^n$ eine teilerfremde Darstellung der Zielfunktion. Dann gilt mit $t := t_1 t_2 \cdots t_n$ und $\bar{t}_i := \frac{t}{t_i}$:

$$c^T v = \sum_{i=1}^{n} \frac{s_i p_i}{t_i d} = \frac{1}{dt} \sum_{i=1}^{n} s_i p_i \bar{t}_i.$$

Aus Lemma 6.3 (a) folgt $t \leq 2^{\langle t_1\rangle - 1} \cdots 2^{\langle t_n\rangle - 1} = 2^{\sum(\langle t_i\rangle - 1)} \leq 2^{\langle c\rangle - n}$ und somit

$$q := dt \leq 2^{3\langle A\rangle + \langle b\rangle + \langle c\rangle - 3n^2 - n}.$$

Analog erhält man

$$p := \sum_{i=1}^{n} s_i p_i \bar{t}_i \leq \sum_{i=1}^{n} 2^{\langle s_i\rangle - 1} 2^{3\langle A\rangle + 2\langle b\rangle - 3n^2} 2^{\langle c\rangle - n} \leq n 2^{3\langle A\rangle + 2\langle b\rangle + 2\langle c\rangle - 3n^2 - n}.$$

□

Der letzte Satz gibt uns nun die Möglichkeit, das Verfahren der binären Suche zur Lösung eines LPs anzuwenden. Diese Methode funktioniert mit der in (6.5) definierten Menge $\mathcal{S}$ wie folgt:

Algorithmus 6.9 (Binäre Suche)

Input: *Ein LP der Form* $\max\{c^T x : Ax \leq b, x \geq 0\}$ *mit* $A \in \mathbb{Q}^{m \times n}, b \in \mathbb{Q}^m$, $c \in \mathbb{Q}^n$ *und beschränktem zulässigem Bereich.*

Output: *Eine Optimallösung des LPs oder die Feststellung, dass die Menge* $\{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}$ *leer ist.*

(0) *Überprüfe, ob*

$$\mathcal{P} = \{x : Ax \leq b, x \geq 0\} = \emptyset.$$

Ist dies der Fall: STOP, LP unzulässig.

(1) *Wähle ein Element* $s \in \mathcal{S}$*, so dass für* $\mathcal{S}' := \{t \in \mathcal{S} : t < s\}$ *und* $\mathcal{S}'' := \{t \in \mathcal{S} : t \geq s\}$ *gilt:*

$$|\mathcal{S}'| \leq |\mathcal{S}''| \leq |\mathcal{S}'| + 1.$$

(2) *Überprüfe, ob*

$$\mathcal{P}_s = \{x : Ax \leq b, x \geq 0, c^T x \geq s\} = \emptyset.$$

(3) *Ist* $\mathcal{P}_s = \emptyset$*, so setze* $\mathcal{S} \leftarrow \mathcal{S}'$*, anderenfalls setze* $\mathcal{S} \leftarrow \mathcal{S}''$

(4) *Ist* $|\mathcal{S}| = 1$*, so gilt für* $s \in \mathcal{S}$*:*

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

und jeder Punkt in $\mathcal{P}_s$ *ist eine Optimallösung von* $\max\{c^T x : x \in \mathcal{P}\}$.

Anderenfalls gehe zu (1).

Die Korrektheit dieses Verfahrens ist offensichtlich. Ist die Binärsuche effizient?
Da in jedem Schritt die Kardinalität $|\mathcal{S}|$ von $\mathcal{S}$ fast halbiert wird, ist klar, dass höchstens

$$\overline{\overline{N}} := \lceil \log_2(|\mathcal{S}| + 1) \rceil$$

Aufspaltungen von $\mathcal{S}$ in zwei Teile notwendig sind, um eine einelementige Menge zu erhalten. Also muss Schritt (2) der binären Suche $\overline{\overline{N}}$ mal ausgeführt werden. Wegen Satz 6.8 gilt

$$|\mathcal{S}| \leq 2n2^{6\langle A \rangle + 3\langle b \rangle + 3\langle c \rangle - 6n^2 - 2n} + 1.$$

Daraus folgt, dass Schritt (2) der binären Suche höchstens

$$\overline{N} := 6\langle A\rangle + 3\langle b\rangle + 3\langle c\rangle - 6n^2 - 2n + \log_2 n + 2$$

mal durchgeführt wird. Ist Schritt (2) in einer Zeit ausführbar, die polynomial in $\langle A\rangle + \langle b\rangle + \langle c\rangle$ ist, dann ist die binäre Suche ein polynomialer Algorithmus, da ein polynomialer Algorithmus für (2) nur $\overline{N}$ mal, also polynomial oft, ausgeführt werden muss.
Zusammen liefern Satz 6.7, Satz 6.8 und Algortihmus 6.9:

Satz 6.10 *Es gibt einen polynomialen Algorithmus zur Lösung linearer Probleme mit rationalen Daten genau dann, wenn es einen Algorithmus gibt, der in polynomialer Zeit entscheidet, ob ein Polytop $\mathcal{P}$ mit rationalen Daten leer ist oder nicht, und der, falls $\mathcal{P} \neq \emptyset$, einen Punkt in $\mathcal{P}$ findet.*

Damit haben wir das lineare Optimierungsproblem reduziert auf die Frage der Lösbarkeit von Ungleichungssystemen, deren Lösungsmenge beschränkt ist.
Wie wir sehen werden, kann die Ellipsoidmethode für $A \in \mathbb{Q}^{m\times n}$, $b \in \mathbb{Q}^m$ mit $\mathcal{P} := \{x : \quad Ax \leq b\}$ nur folgendes relaxierte Zulässigkeitsproblem lösen:

$$\begin{aligned} &\text{Finde } x^* \text{ mit } x^* \in \mathcal{P} = \{x : \quad Ax \leq b\} \\ &\text{oder stelle fest, dass } \overset{\circ}{\mathcal{P}} := \{x : \quad Ax < b\} = \emptyset. \end{aligned} \tag{ZUL*}$$

Es kann also nur für volldimensionale Polyeder zulässige Punkte finden oder feststellen, dass das Polyeder nicht volldimensional ist. Dies läßt sich aber durch folgenden Satz reparieren:

Satz 6.11 *Seien $A \in \mathbb{Q}^{m\times n}$ und $b \in \mathbb{Q}^m$. Dann hat das Ungleichungssystem*

$$Ax \leq b$$

genau dann eine Lösung, wenn das strikte Ungleichungssystem

$$Ax < b + 2^{-2\langle A\rangle - \langle b\rangle}\mathbb{1}$$

eine Lösung hat. Ferner kann man aus einer Lösung des strikten Ungleichungssystems (oder auch des zugehörigen nicht strikten Ungleichungssystems) in polynomialer Zeit eine Lösung von $Ax \leq b$ konstruieren.

Um das Zulässigkeitsproblem zu lösen, kann man also die Ellipsoidmethode auf $\mathcal{P} = \{x : \quad Ax \leq b\}$ anwenden. Stellt sie fest, dass $\overset{\circ}{\mathcal{P}} = \{x : \quad Ax < b\} = \emptyset$, dann wendet man sie nochmals an auf $\{x : \quad Ax \leq b + 2^{-2\langle A\rangle - \langle b\rangle}\mathbb{1}\}$. Erhält man wieder keine Lösung, dann ist $\mathcal{P} = \emptyset$, andernfalls kann man nach Satz 6.11 in polynomialem Aufwand daraus ein $x^* \in \mathcal{P}$ konstruieren.

6.3 Die Ellipsoidmethode

Die Ellipsoidmethode ist ein polynomialer Algorithmus, der das abgeschwächte Zulässigkeitsproblem (ZUL*) löst.
Wir setzen in diesem Abschnitt $n \geq 2$ voraus.

Definition 6.12 *Eine Menge $\mathcal{E} \subset \mathbb{R}^n$ heißt* Ellipsoid (mit Zentrum a)*, wenn es einen Punkt $a \in \mathbb{R}^n$ und eine symmetrische, positiv definite Matrix A gibt, so dass*

$$\mathcal{E} = \mathcal{E}(A, a) = \{x \in \mathbb{R}^n : (x-a)^T A^{-1}(x-a) \leq 1\}.$$

Das Ellipsoid ist also durch die (ebenfalls positiv definite) Inverse A^{-1} von A definiert. Das liegt daran, dass sich bei dieser Definition viele Eigenschaften von $\mathcal{E}(A, a)$ aus den Eigenschaften von A ableiten lassen. Zum Beispiel:

- der Durchmesser von $\mathcal{E}(A, a)$ ist gleich $2\sqrt{\lambda_{\max}}$, wobei $\lambda_{\max}$ der größte Eigenwert von A ist.
- Die Symmetrieachsen von $\mathcal{E}(A, a)$ entsprechen den Eigenvektoren von A.
- $\mathcal{E}(A, a)$ ist das Bild der Einheitskugel $\mathcal{B} = \{u \in \mathbb{R}^n : \|u\|_2 \leq 1\}$ unter der affinen Transformation $f(u) = A^{1/2}u + a$. Dabei ist $A^{1/2}$ die eindeutig bestimmte Wurzel der positiv definiten Matrix A, für die $A^{1/2}A^{1/2} = A$ gilt.

Wir wollen nun die Ellipsoidmethode zunächst geometrisch betrachten.

Geometrische Beschreibung der Ellipsoidmethode
Gegeben sei ein Polytop $\mathcal{P}$. Wir wollen einen Punkt in $\mathcal{P}$ finden oder zeigen, dass $\overset{\circ}{\mathcal{P}}$ leer ist. Die Ellipsoidmethode besteht aus sechs Schritten:

(1) Konstruiere ein Ellipsoid $\mathcal{E}_0 = \mathcal{E}(A_0, a_0)$, das $\mathcal{P}$ enthält. Setze $k = 0$.

(2) Ist das gegenwärtige Ellipsoid “zu klein”, dann STOP: $\mathcal{P}$ ist leer.

(3) Teste, ob der Mittelpunkt a_k von $\mathcal{E}_k$ in $\mathcal{P}$ enthalten ist.

(4) Gilt $a_k \in \mathcal{P}$, dann haben wir einen Punkt in $\mathcal{P}$ gefunden: STOP.

(5) Gilt $a_k \notin \mathcal{P}$, dann gibt es eine $\mathcal{P}$ definierende Ungleichung

$$c^T x \leq \gamma,$$

die von a_k verletzt wird, d.h. es gilt $c^T a_k > \gamma$. Bezeichne mit

$$\mathcal{E}'_k := \mathcal{E}_k \cap \{x : c^T x \leq c^T a_k\}$$

das Halbellipsoid von $\mathcal{E}_k$, das $\mathcal{P}$ enthält. Konstruiere das Ellipsoid kleinsten Volumens, das $\mathcal{E}'_k$ enthält, und bezeichne es mit $\mathcal{E}_{k+1}$.

(6) Setze $k \leftarrow k + 1$ und gehe zu (2).

Das Prinzip dieses Verfahrens ist klar: Man beginnt mit einem Ellipsoid, das $\mathcal{P}$ enthält. Ist das Zentrum vonn $\mathcal{E}_k$ nicht in $\mathcal{P}$, so sucht man ein kleineres Ellipsoid $\mathcal{E}_{k+1}$, das $\mathcal{P}$ enthält, und so weiter. Folgende Fragen müssen nun geklärt werden:

- Wie findet man ein Anfangsellipsoid $\mathcal{E}_0$, das $\mathcal{P}$ enthält?
- Wie kann man $\mathcal{E}_{k+1}$ aus $\mathcal{E}_k$ konstruieren?
- Wann bricht man ab, d.h. was heißt "zu klein"?
- Wie viele Iterationen sind durchzuführen?

Die folgenden Sätze beantworten diese Fragen, wobei wir nicht alle der (technischen) Beweise angeben wollen.

Das **Anfangsellipsoid** soll eine Kugel mit dem Nullpunkt als Zentrum sein. Enthalten die Ungleichungen, die das Polytop definieren, explizite obere und untere Schranken für die Variablen

$$\ell_i \leq x_i \leq u_i \qquad i = 1, \ldots, n,$$

so kann

$$R := \sqrt{\sum_{i=1}^{n} (\max\{|u_i|, |\ell_i|\})^2}$$

als Radius der Kugel gewählt werden. Anderenfalls kann man zeigen:

Lemma 6.13 *Sei $\mathcal{P}$ ein Polyeder der Form $\mathcal{P}(A,b)$, $\mathcal{P}^=(A,b)$ oder $\mathcal{P} = \{x : Ax \le b, x \ge 0\}$ mit $A \in \mathbb{Q}^{m\times n}, b \in \mathbb{Q}^m$. Dann gilt*

(a) Alle Ecken von $\mathcal{P}$ sind in der Kugel $\mathcal{B}(0,R)$ enthalten mit

$$R := \sqrt{n}2^{2\langle A\rangle+\langle b\rangle-2n^2}.$$

(b) Ist $\mathcal{P}$ ein Polytop, so gilt $\mathcal{P} \subseteq \mathcal{B}(0,R) = \mathcal{E}(R^2 I, 0)$.

Beweis. Nach Satz 6.6 gilt für jede Ecke $v = (v_1,\ldots,v_n)^T$ von $\mathcal{P}$

$$|v_i| \le 2^{2\langle A\rangle+\langle b\rangle-2n^2} \quad \text{für alle} \quad i = 1,\ldots,n,$$

und daraus folgt für die euklidische Norm von v:

$$\|v\|_2 = \sqrt{\sum_{i=1}^n v_i^2} \le \sqrt{n\max\{v_i^2\}} \le \sqrt{n}2^{2\langle A\rangle+\langle b\rangle-2n^2}.$$

Also ist jede Ecke von $\mathcal{P}$ in $\mathcal{B}(0,R)$ enthalten. Ist insbesondere $\mathcal{P}$ ein Polytop, so folgt daraus $\mathcal{P} \subseteq \mathcal{B}(0,R)$. □

Damit haben wir ein Anfangsellipsoid gefunden, mit dem die Ellipsoidmethode gestartet werden kann.

Die **Konstruktion von $\mathcal{E}_{k+1}$ aus $\mathcal{E}_k$** geschieht wie folgt:

Satz 6.14 *Sei $\mathcal{E}_k = \mathcal{E}(A_k, a_k) \subset \mathbb{R}^n$ ein Ellipsoid, $c \in \mathbb{R}^n \setminus \{0\}$ und $\mathcal{E}'_k := \mathcal{E}_k \cap \{x : c^T x \le c^T a_k\}$. Setze*

$$\begin{aligned} d &:= \frac{1}{\sqrt{c^T A_k c}} A_k c, \\ a_{k+1} &:= a_k - \frac{1}{n+1} d \\ A_{k+1} &:= \frac{n^2}{n^2-1}\left(A_k - \frac{2}{n+1} d d^T\right), \end{aligned}$$

dann ist A_{k+1} positiv definit und $\mathcal{E}_{k+1} := \mathcal{E}(A_{k+1}, a_{k+1})$ ist das eindeutig bestimmte Ellipsoid kleinsten Volumens, das $\mathcal{E}'_k$ enthält.

Beweis. Siehe Grötschel, Lovász und Schrijver [GLS88]. □

Geometrisch bedeutet dieser Satz folgendes: Ist $c \in \mathbb{R}^n \setminus \{0\}$, so kann man Maximum und Minimum von $c^T x$ über $\mathcal{E}_k$ explizit angeben. Es gilt nämlich

$$z_{\max} := a_k + d, \qquad \text{und} \qquad z_{\min} := a_k - d$$

und

$$\begin{aligned} c^T z_{\max} &= \max\{c^T x : x \in \mathcal{E}_k\} = c^T a_k + \sqrt{c^T A_k c}, \\ c^T z_{\min} &= \min\{c^T x : x \in \mathcal{E}_k\} = c^T a_k - \sqrt{c^T A_k c}. \end{aligned}$$

Daraus folgt, dass der Mittelpunkt a_{k+1} des neuen Ellipsoids $\mathcal{E}_{k+1}$ auf dem Geradenstück zwischen a_k und $z_{\min}$ liegt. Die Länge dieses Geradenstücks ist $\|d\|$, und a_{k+1} erreicht man von a_k aus, indem man einen Schritt der Länge $\frac{1}{n+1}\|d\|$ in Richtung $-d$ macht.
Der Durchschnitt des Randes des Ellipsoids $\mathcal{E}_{k+1}$ mit dem Rand von $\mathcal{E}'_k$ wird gebildet durch den Punkt $z_{\min}$ und $\mathcal{E}''_k := \{x : (x - a_k)^T A_k^{-1}(x - a_k) = 1\} \cap \{x : c^T x = c^T a_k\}$. $\mathcal{E}''_k$ ist der Rand eines $(n-1)$-dimensionalen Ellipsoids im $\mathbb{R}^n$.

Das **Stoppkriterium** der Ellipsoidmethode beruht aus einem Volumenargument. Nach Konstruktion ist klar, dass das Volumen von $\mathcal{E}_{k+1}$ (bezeichnet mit $\mathrm{vol}(\mathcal{E}_{k+1})$) kleiner ist als das von $\mathcal{E}_k$. Man kann den Volumenschrumpfungsfaktor explizit berechnen. Er hängt nur von der Dimension n des Raumes $\mathbb{R}^n$ ab und nicht vom Updatevektor c.

Lemma 6.15 *Es gilt*

$$\frac{\mathrm{vol}(\mathcal{E}_{k+1})}{\mathrm{vol}(\mathcal{E}_k)} = \left(\left(\frac{n}{n+1}\right)^{n+1} \left(\frac{n}{n-1}\right)^{n-1} \right)^{\frac{1}{2}} \le e^{-\frac{1}{2n}} < 1.$$

Beweis. Siehe Grötschel, Lovász und Schrijver [GLS88]. □

Mit den Formeln aus Satz 6.14 kann also eine Folge von Ellipsoiden konstruiert werden, so dass jedes Ellipsoid $\mathcal{E}_{k+1}$ das Halbellipsoid $\mathcal{E}'_k$ und somit das Polytop $\mathcal{P}$ enthält und dass die Volumina der Ellipsoide schrumpfen.

Die Folge der Volumina konvergiert gegen Null, daher muss einmal das Volumen von $\mathcal{P}$, falls $\mathcal{P}$ ein positives Volumen hat, unterschritten werden. Daher muss nach endlich vielen Schritten der Mittelpunkt eines der Ellipsoide in $\mathcal{P}$ sein, falls $\mathcal{P} \neq \emptyset$. Das folgende Lemma hilft bei der Berechnung, nach wie vielen Schritten dies der Fall ist:

Lemma 6.16 *Seien $\mathcal{P} = \mathcal{P}(A, b) \subseteq \mathbb{R}^n$, sei $\overset{\circ}{\mathcal{P}} = \{x \in \mathbb{R}^n : Ax < b\}$ und sei $R := \sqrt{n}2^{2\langle A\rangle+\langle b\rangle-2n^2}$.*

(a) Entweder gilt $\overset{\circ}{\mathcal{P}} = \emptyset$ oder

$$\mathrm{vol}(\overset{\circ}{\mathcal{P}} \cap \mathcal{B}(0, R)) \geq 2^{-(n+1)(\langle A\rangle+\langle b\rangle-n^2)}.$$

(b) Ist $\mathcal{P}$ volldimensional, d.h. $\dim\mathcal{P} = n$, dann gilt

$$\mathrm{vol}(\mathcal{P} \cap \mathcal{B}(0, R)) = \mathrm{vol}(\overset{\circ}{\mathcal{P}} \cap \mathcal{B}(0, R)) \geq 2^{-(n+1)(\langle A\rangle+\langle b\rangle-n^2)} > 0.$$

Beweis. Siehe Grötschel, Lovász und Schrijver [GLS88]. □

Lemma 6.16 und Lemma 6.13 implizieren folgendes:
Ist für ein Polyeder $\mathcal{P}$ die Menge $\overset{\circ}{\mathcal{P}} \neq \emptyset$, dann ist auch $\overset{\circ}{\mathcal{P}} \cap \mathcal{B}(0, R) \neq \emptyset$ und es gilt

$$\mathrm{vol}(\overset{\circ}{\mathcal{P}} \cap \mathcal{B}(0, R)) \geq 2^{-(n+1)(\langle A\rangle+\langle b\rangle-n^2)}.$$

Um $\overset{\circ}{\mathcal{P}} \neq \emptyset$ zu prüfen, reicht es also aus zu überprüfen, ob gilt

$$\mathrm{vol}(\overset{\circ}{\mathcal{P}} \cap \mathcal{B}(0, R)) \geq 2^{-(n+1)(\langle A\rangle+\langle b\rangle-n^2)},$$

sonst muss $\overset{\circ}{\mathcal{P}} = \emptyset$ sein.
Wir konstruieren mit der Ellipsoidmethode Ellipsoide $\mathcal{E}_k \supset \overset{\circ}{\mathcal{P}} \cap \mathcal{B}(0, R)$, deren Volumen in jedem Schritt mindestens um $e^{-1/(2n)}$ abnimmt, solange der Mittelpunkt a_k nicht in $\mathcal{P}$ liegt. Irgendwann muss also $a_k \in \mathcal{P}$ gelten, oder wir erreichen ein N mit

$$\mathrm{vol}(\mathcal{E}_N) < 2^{-(n+1)(\langle A\rangle+\langle b\rangle-n^2)}.$$

Wegen $\mathcal{E}_N \supset \mathring{\mathcal{P}} \cap \mathcal{B}(0,R)$ bleibt dann nur $\mathring{\mathcal{P}} = \emptyset$ übrig, da das Volumen von $\mathrm{vol}(\mathcal{E}_N)$ zu klein geworden ist.

Damit können wir nun die Ellipsoidmethode zur Lösung von (ZUL*) formulieren:

Algorithmus 6.17 (Die Ellipsoidmethode)

Input: *$A \in \mathbb{Q}^{m\times n}$ und $b \in \mathbb{Q}^m$.*

Output: *Ein Punkt $x^* \in \mathbb{Q}^n$ mit $Ax \le b$ oder die Feststellung, dass die Menge $\{x \in \mathbb{R}^n : Ax < b\}$ leer ist.*

Schritt 1: Initialisierung

Setze $R := \sqrt{n}2^{2\langle A\rangle+\langle b\rangle-2n^2}$ oder kleiner, wenn Vorinformation vorhanden.

Setze

$$\begin{aligned} A_0 &:= R^2 I \\ a_0 &:= 0 \\ k &:= 0 \\ N &:= 2n\big((3n+1)\langle A\rangle + (2n+1)\langle b\rangle - n^3\big). \end{aligned}$$

(Das Anfangsellipsoid ist $\mathcal{E}_0 := \mathcal{E}(A_0, a_0)$.)

Schritt 2: Abbruchkriterium

(2a) *Gilt $k = N$, dann STOP: $Ax < b$ hat keine Lösung.*

(2b) *Gilt $Aa_k \le b$, dann STOP: der Punkt $x^* = a_k$ ist gefunden.*

(2c) *Anderenfalls sei c^T eine Zeile von A derart, dass der Mittelpunkt a_k von $\mathcal{E}_k$ die entsprechende Ungleichung verletzt.*

Schritt 3: Update

(3a) *Setze*

$$\begin{aligned} d &:= \frac{1}{\sqrt{c^T A_k c}} A_k c, \\ a_{k+1} &:= a_k - \frac{1}{n+1} d \\ A_{k+1} &:= \frac{n^2}{n^2-1}\left(A_k - \frac{2}{n+1} dd^T\right), \end{aligned}$$

($\mathcal{E}_{k+1} := \mathcal{E}(A_{k+1}, a_{k+1})$ ist das neue Ellipsoid.)

(3b) *Setze $k \leftarrow k+1$ und gehe zu Schritt 2.*

Beispiel 6.18 *Betrachten wir das Polytop $\mathcal{P} \subseteq \mathbb{R}^2$, das durch die vier Ungleichugen*

$$\tfrac{17}{20} \leq x_1 \leq \tfrac{18}{20} \qquad \textit{und} \qquad -\tfrac{1}{5} \leq x_2 \leq \tfrac{1}{5}$$

definiert ist. Als Anfangsellipsoid sei die Kugel um den Ursprung mit Radius 1 gewählt, d.h.

$$a_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad A_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In Iteration 1 stellt man fest, dass die Ungleichung $\frac{17}{20} \leq x_1$ verletzt ist, daher ergibt sich der Vektor c als $c = (-1, 0)^T$. Damit erhalten wir:

$$\begin{aligned} d &= \frac{1}{\sqrt{(-1,0)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 0 \end{pmatrix}}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 0 \end{pmatrix} \\ &= \frac{1}{1}\begin{pmatrix} -1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} -1 \\ 0 \end{pmatrix}. \end{aligned}$$

sowie

$$a_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ 0 \end{pmatrix}$$

und

$$\begin{aligned} A_1 &= \frac{4}{3}\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{2}{3}\begin{pmatrix} -1 \\ 0 \end{pmatrix}(-1,0)\right] \\ &= \frac{4}{3}\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{2}{3}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] \\ &= \begin{pmatrix} \frac{4}{9} & 0 \\ 0 & \frac{4}{3} \end{pmatrix}. \end{aligned}$$

Man fährt genauso fort, bis in Iteration fünf der Punkt $a_5 = (\frac{211}{243}, 0)^T$ in $\mathcal{P}$ enthalten ist.

Satz 6.19 *Die Ellipsoidmethode arbeitet korrekt.*

Beweis. Gibt es ein $k < N$, so dass $Aa_k \leq b$, so ist offenbar ein Punkt aus $\mathcal{P}(A, b)$ gefunden. Bricht die Ellipsoidmethode in Schritt (2a) ab, so müssen wir zeigen, dass $\overset{\circ}{\mathcal{P}} = \{x \in \mathbb{R}^n : Ax < b\}$ kein Element besitzt.

Angenommen, $\overset{\circ}{\mathcal{P}} \neq \emptyset$. Sei $\mathcal{P}' = \overset{\circ}{\mathcal{P}} \cap \mathcal{E}_0$. Dann gilt nach Lemma 6.16, dass das Volumen von $\mathcal{P}'$ mindestens $2^{-(n+1)(\langle A\rangle + \langle b\rangle - n^2)}$ beträgt. Ist $a_k \notin \mathcal{P}(A, b)$ für $0 \leq k < N$, so wird in (2c) eine verletzte Ungleichung $c^T x \leq \gamma$ gefunden. Wegen $\gamma < c^T a_k$ enthält das Halbellipsoid

$$\mathcal{E}'_k := \mathcal{E}_k \cap \{x : c^T x \leq c^T a_k\}$$

die Menge $\mathcal{P}'$. Das durch die Formeln (3a) konstruierte Ellipsoid $\mathcal{E}_{k+1}$ ist nach Satz 6.14 das volumsmäßig kleinste Ellipsoid, das $\mathcal{E}'_k$ enthält. Wegen $\mathcal{P}' \subseteq \mathcal{E}'_k$ gilt natürlich $\mathcal{P}' \subseteq \mathcal{E}_{k+1}$. Daraus folgt

$$\mathcal{P}' \subseteq \mathcal{E}_k \quad \text{für alle} \quad 0 \leq k \leq N.$$

Das Volumen des Anfangsellipsoids $\mathcal{E}_0$ kann man berechnen:

$$\mathrm{vol}(\mathcal{E}_0) = \sqrt{\det(R^2 I)}V_n = R^n V_n,$$

wobei V_n das Volumen der Einheitskugel in $\mathbb{R}^n$ ist. Wir schätzen nun sehr grob ab: Die Einheitskugel ist im Würfel $\mathcal{W} = \{x \in \mathbb{R}^n : |x_i| \leq 1,\ i = 1, \ldots, n\}$ enthalten, dessen Volumen offensichtlich $\mathrm{vol}(\mathcal{W}) = 2^n$ ist. Daraus folgt:

$$\mathrm{vol}(\mathcal{E}_0) < R^n 2^n = 2^{n(2\langle A\rangle + \langle b\rangle - 2n^2 + \log\sqrt{n} + 1)} < 2^{n(2\langle A\rangle + \langle b\rangle - n^2)}.$$

Nach Lemma 6.15 schrumpft in jedem Schritt der Ellipsoidmethode das Volumen um mindestens den Faktor $e^{-\frac{1}{2n}}$. Aus der Abschätzung von $\mathrm{vol}(\mathcal{E}_0)$ und der Formel für N erhalten wir somit

$$\mathrm{vol}(\mathcal{E}_N) \leq e^{-\frac{N}{2n}} \mathrm{vol}(\mathcal{E}_0) < 2^{-(n+1)(\langle A\rangle + \langle b\rangle - n^2)}.$$

Also gilt $\mathrm{vol}(\mathcal{E}_N) < \mathrm{vol}(\mathcal{P}')$. Das ist ein Widerspruch zu $\mathcal{P}' \subseteq \mathcal{E}_N$. Daraus folgt, dass $\mathcal{P}'$ und somit $\{x : Ax < b\}$ leer sind, wenn die Ellipsoidmethode nicht in (2a) abbricht. □

Folgerung 6.20 *Ist $\mathcal{P} = \mathcal{P}(A, b) \subseteq \mathbb{R}^n$ ein Polyeder, von dem wir wissen, dass es entweder volldimensional oder leer ist, dann findet die Ellipsoidmethode entweder einen Punkt in $\mathcal{P}$ oder beweist, dass $\mathcal{P}$ leer ist.*

Wir hatten bereits die Methodik angesprochen, wie zu verfahren ist, wenn das Polyeder nicht volldimensional ist, siehe Satz 6.11.
Wir fassen das Vorgehen nochmal zusammen: Leider sieht man einem durch ein Ungleichungssystem gegebenen Polyeder nicht unmittelbar an, ob es volldimensional ist oder nicht. Außerdem sind gerade die Polyeder, die durch die Transformation (6.1) linearer Probleme auf ein Zulässigkeitsproblem entstehen nicht volldimensional (es gilt immer $c^T x = b^T y$), so dass die Ellipsoidmethode zu keiner befriedigenden Antwort führt. Diesen Defekt kann man jedoch durch Satz 6.11 reparieren.
Damit ist die Beschreibung der Ellipsoidmethode bis auf die Abschätzung der Rechenzeit vollständig. Wollen wir entscheiden, ob ein Polyeder $\mathcal{P}(A, b)$ einen Punkt enthält, können wir die Ellipsoidmethode auf das Ungleichungssystem $Ax \leq b$ anwenden. Finden wir einen Punkt in $\mathcal{P}(A, b)$, dann sind wir fertig. Anderenfalls wissen wir, dass $\mathcal{P}(A, b)$ nicht volldimensional ist. In diesem Fall können wir auf Satz 6.11 zurückgreifen und starten die Ellipsoidmethode neu, und zwar mit dem ganzzahligen Ungleichungssystem

$$2^{2\langle A\rangle - \langle b\rangle} Ax \leq 2^{2\langle A\rangle - \langle b\rangle} b + \mathbb{1}. \tag{6.6}$$

Entscheidet die Ellipsoidmethode, dass das zu (6.6) gehörende strikte Ungleichungssystem keine Lösung hat, so können wir aus Satz 6.11 folgern, dass $\mathcal{P}(A, b)$ leer ist. Anderenfalls findet die Ellipsoidmethode einen Punkt x', der (6.6) erfüllt. Gilt $x' \in \mathcal{P}(A, b)$, dann haben wir das Gewünschte gefunden, falls nicht, kann man aus x' einen Punkt $x \in \mathcal{P}(A, b)$ konstruieren.

Zur Lösung von LPs kann man die in Abschnitt 6.2 erwähnten Reduktionen benutzen. Entweder man fasst das LP und sein duales zusammen und sucht wie oben angegeben im nicht volldimensionalen Polyeder, oder man verwendet die Binärsuche und wendet die Ellipsoidmethode $\overline{N}$ mal für niederdimensionale Polyeder des Typs $\mathcal{P}_s$ an.

In beiden Fällen ist das Gesamtverfahren polynomial, wenn die Ellipsoidmethode polynomial ist.

6.4 Laufzeit der Ellipsoidmethode

Wir wollen nun die Laufzeit der Ellipsoidmethode untersuchen und auf einige bisher verschwiegene Probleme bei ihrer Ausführung aufmerksam machen.

Offenbar ist die maximale Iterationszahl

$$N = 2n\big((3n+1)\langle A\rangle + (2n+1)\langle b\rangle - n^3\big)$$

polynomial in der Kodierungslänge von A und b. Also ist die Ellipsoidmethode genau dann polynomial, wenn jede Iteration in polynomialer Zeit ausgeführt werden kann.
Bei der Initialisierung besteht kein Problem. Test (2a) ist trivial, und die Schritte (2b) und (2c) führen wir dadurch aus, dass wir das Zentrum a_k in die Ungleichungen einsetzen und überprüfen, ob die Ungleichungen erfüllt sind oder nicht. Die Anzahl der dazu benötigten Rechenschritte ist linear in $\langle A\rangle$ und $\langle b\rangle$. Sie ist also polynomial, wenn die Kodierungslänge des Vektors a_k polynomial ist.
Hier beginnen die Schwierigkeiten: In der Updateformel (3a) muss eine Wurzel berechnet werden. Im Allgemeinen werden hier also irrationale Zahlen auftreten, die natürlich nicht exakt berechnet werden können. Die Zahl $\sqrt{c^T A_k c}$ muss daher zu einer rationalen Zahl gerundet werden. Dadurch wird geometrisch bewirkt, dass der Mittelpunkt des Ellipsoids $\mathcal{E}_{k+1}$ ein wenig verschoben wird. Mit Sicherheit enthält das verschobene Ellipsoid nicht mehr die Menge $\mathcal{E}'_k$, und möglicherweise ist auch $\mathcal{P}(A, b)$ nicht mehr in diesem Ellipsoid enthalten. Also bricht der gesamte Beweis der Korrektheit des Verfahrens zusammen.
Außerdem wird beim Update (3a) durch möglicherweise große Zahlen dividiert, und es ist nicht klar, dass die Kodierungslänge der Elemente von A_{k+1} bei wiederholter Anwendung von (3a) polynomial in $\langle A\rangle$ und $\langle b\rangle$ bleibt. Also müssen auch die Einträge von A_{k+1} gerundet werden. Dies kann zu folgenden Problemen führen: Die gerundete Matrix A^*_{k+1} ist nicht mehr positiv definit, und das Verfahren wird sinnlos. Oder A^*_{k+1} bleibt positiv definit, aber durch die Rundung hat sich die Form des zugehörigen Ellipsoids $\mathcal{E}^*_{k+1}$ so geändert, dass $\mathcal{E}^*_{k+1}$ das Polyeder $\mathcal{P}(A, b)$ nicht mehr enthält.

Alle diese Klippen kann man mit einem Trick umschiffen, dessen Korrektheitsbeweis allerdings recht aufwändig ist.
Die geometrische Idee hinter dem Trick ist folgende:
Man nehme die binäre Darstellung der Komponenten des in (3a) berechneten Vektors und der ebenfalls in (3a) berechneten Matrix und runde nach p Stellen hinter dem Binärkomma. Dadurch ändert man die Lage des Mittelpunktes und die Form des Ellipsoids ein wenig.

Nun bläst man das Ellipsoid ein bisschen auf, d.h. man multipliziert A_{k+1} mit einem Faktor $\xi > 1$, und zwar so, dass die Menge $\mathcal{E}'_k$ beweisbar in dem aufgeblasenen Ellipsoid enthalten ist. Durch die Vergrößerung des Ellipsoids wird die in Lemma 6.15 bestimmte Schrumpfungsrate verschlechtert, was bedeutet, dass man insgesamt mehr, sagen wir N', Iterationen durchführen muss. Daraus folgt, dass der Rundungsparameter p und der Aufblasparameter ξ so auf einander abgestimmt sein müssen, dass

- alle gerundeten und mit ξ multiplizierten Matrizen A_k $(1 \leq k \leq N')$ und alle gerundeten Mittelpunkte a_k $(1 \leq k \leq N')$ polynomial in $\langle A \rangle$ und $\langle b \rangle$ berechnet werden können,
- alle A_k positiv definit sind,
- $\mathcal{P} \cap \mathcal{B}(0, R) \subseteq \mathcal{E}_k$ für alle $1 \leq k \leq N'$ und
- die Iterationszahl N' ebenfalls polynomial in $\langle A \rangle$ und $\langle b \rangle$ ist.

Dies kann tatsächlich realisiert werden. Der Beweis soll hier allerdings nicht ausgeführt werden. Genaueres findet man wieder im Buch von Grötschel, Lovász und Schrijver [GLS88].

Eine geeignete Modifikation von Algorithmus 6.17, die alle obigen Kriterien erfüllt ist folgende:

- Wähle in Algorithmus 6.17

  $$N := 5n\big((3n+1)\langle A \rangle + (2n+1)\langle b \rangle - n^3\big).$$

  Ferner setze

  $$p = 8N, \quad \xi = 1 + \frac{n^2 - 3}{2n^4}.$$

- Modifiziere Schritt (3a) wie folgt:

  (3a) Setze

  $$\begin{aligned} a_{k+1} &:=_p a_k - \frac{1}{n+1} \frac{1}{\sqrt{c^T A_k c}} A_k c, \\ A_{k+1} &:=_p \xi \frac{n^2}{n^2 - 1} \left(A_k - \frac{2}{n+1} \frac{A_k c c^T A_k}{c^T A_k c} \right), \end{aligned}$$

  ($\mathcal{E}_{k+1} := \mathcal{E}(A_{k+1}, a_{k+1})$ ist das neue Ellipsoid.)

Hierbei bedeutet $:=_p$, dass bei der Rechnung auf p Nachkommastellen in der Binärdarstellung gerundet wird. Man beachte, dass bis auf die inexakte Rechnung und den Aufblähfaktor ξ alles unverändert ist.

Man kann zeigen, dass der Reduktionsfaktor der Ellipsoidvolumen nun $\leq e^{-1/(5n)}$ ist und $\mathcal{E}_{k+1} \supset \mathcal{E}'_k$ sichergestellt bleibt.
Die Ellipsoidmethode (mit Rundungsmodifikation) ist also ein polynomialer Algorithmus, der entscheidet, ob ein Polyeder leer ist oder nicht. Mit Hilfe der beschriebenen Reduktionen kann sie dazu benutzt werden, lineare Probleme in polynomialer Zeit zu lösen. Damit haben wir gezeigt:

LPs sind in polynomialer Zeit lösbar.

6.5 Separieren und Optimieren

Wir schließen mit der wichtigen Aussage, dass Optimieren und Separieren polynomial äquivalent sind. Die Äquivalenz von Separieren und Optimieren ist eines der bedeutendsten Resultate der polyedrischen Kombinatorik (vgl.[GLS81]). Es hat sowohl die Theorie der kombinatorischen und ganzzahligen Optimierung entscheidend beeinflusst, als auch die theoretische Rechtfertigung für Branch-and-Cut Verfahren (die derzeit erfolgreichste Methode zum Lösen $\mathcal{NP}$-schwerer praktischer Probleme der Diskreten Optimierung) ermöglicht.
Wir nehmen im folgenden an, dass die betrachteten Polyeder immer volldimensional und beschränkt sind.
Wir beginnen mit der Definition der Probleme, die wir in Beziehung setzen wollen.

Problem 6.21 (Optimierungsproblem — OPT) *Gegeben sei ein Polyeder $\mathcal{P} \subseteq \mathbb{R}^n$ und ein Vektor $c \in \mathbb{Q}^n$.*

(i) Bestätige, dass $\mathcal{P} = \emptyset$ oder

(ii) Finde $y \in \mathcal{P}$ mit $c^T y = \max\{c^T x \mid x \in \mathcal{P}\}$ oder

(iii) bestätige, dass $\max\{c^T x \mid x \in \mathcal{P}\}$ unbeschränkt ist, d.h. finde eine Extremale z von $\mathcal{P}$ mit $c^T z \geq 1$.

Unter unserer Annahme, dass $\mathcal{P}$ volldimensional und beschränkt sein soll, kommt natürlich nur Punkt (ii) in Frage, d.h. Problem 6.21 reduziert sich auf das Finden einer Optimallösung.

Problem 6.22 (Separierungsproblem — SEP) *Gegeben sei ein Polyeder $\mathcal{P} \subseteq \mathbb{R}^n$ und ein Vektor $y \in \mathbb{Q}^n$. Entscheide, ob $y \in \mathcal{P}$. Falls nicht, finde einen Vektor $c \in \mathbb{Q}^n$ mit $c^T y > \max\{c^T x \mid x \in \mathcal{P}\}$.*

Wir betrachten noch ein drittes Problem, das – wie wir noch sehen werden – dual zum Problem 6.22 ist.

Problem 6.23 (Verletztheitsproblem — VIOL) *Gegeben sei ein Polyeder $\mathcal{P} \subseteq \mathbb{R}^n$, ein Vektor $c \in \mathbb{Q}^n$ und ein Skalar $\gamma \in \mathbb{Q}$. Entscheide, ob $c^T x \leq \gamma$ für alle $x \in \mathcal{P}$. Falls nicht, finde einen Vektor $y \in \mathcal{P}$ mit $c^T y > \gamma$.*

Die Fragen, die uns in diesem Abschnitt interessieren, kann man folgerndermaßen beschreiben: Angenommen, wir kennen eine Methode (im Folgenden werden wir von einem Orakel sprechen), die eines der drei Probleme löst. Ist es dann auch möglich, die beiden anderen Probleme in orakel-polynomialer Zeit zu lösen?

Auf den ersten Blick scheint das Optimierungsproblem das schwerste zu sein. Könnten wir dies lösen, so könnten wir sicher auch das Verletztheitsproblem lösen. Aber könnten wir damit auch das Separierungsproblem lösen? Und umgekehrt, wenn wir separieren können, können wir dann auch optimieren?

In diesem Kapitel werden wir zeigen, dass alle drei Probleme in der Tat polynomial äquivalent sind. Wir werden im Folgenden die Probleme 6.21 bis 6.23 immer mit OPT, SEP und VIOL abkürzen. Schreiben wir Klammern um die Abkürzungen, so bezeichne dies ein Orakel für das jeweilige Problem, d.h. (OPT), (SEP) und (VIOL) sind Orakel für OPT, SEP und VIOL. Wir nehmen in diesem Abschnitt $n \geq 2$ an.

Satz 6.24 *Gegeben sei ein beschränktes und volldimensionales Polyeder $\mathcal{P}$ das endlich codierbar sei, also $\mathcal{P} = \mathcal{P}(A, b)$. mit $A \in \mathbb{Q}^{m,n}$, $b \in \mathbb{Q}^m$. Eine Schranke für die Codierungslänge sei bekannt. Angenommen wir hätten ein Orakel, das eines der drei Probleme SEP, OPT oder VIOL für $\mathcal{P}$ löst. Dann können die zwei verbliebenen Probleme in orakel-polynomialer Zeit gelöst werden.*

Beweis. Wir zeigen nur die ersten beiden der Schritte

- (SEP) $\rightarrow$ (OPT)
- (OPT) $\rightarrow$ (VIOL)
- (VIOL) $\rightarrow$ (SEP).

Zu zeigen ist, dass die Transformationen jeweils in orakel-polynomialer Zeit erfolgen können.

(SEP) $\rightarrow$ (OPT).

Angenommen, wir haben ein Separationsorakel (SEP), das SEP für $\mathcal{P}$ löst, wobei $\mathcal{P}$ volldimensional und beschränkt ist. Weiterhin haben wir einen Zielfunktionsvektor $c \in \mathbb{Q}^n$ und damit möchten wir das Problem $\max\{c^T x \mid x \in P\}$ lösen.
Wir wollen binäre Suche in Verbindung mit der Ellipsoidmethode durchführen. Bei der binären Suche treten die Polyeder auf $\mathcal{P}_s = \{x \in \mathcal{P} \;:\; c^T x \geq s\}$. Das Orakel (SEP) für $\mathcal{P}$ kann leicht zu einem Orakel (SEP) für $\mathcal{P}_s$ modifiziert werden (gilt $c^T x < s$, dann ist dies bereits eine trennende Ungleichung, sonst konsultiere (SEP) für $\mathcal{P}$).
Nun stellen wir fest, dass das Ellipsoidverfahren angewendet auf das Polyeder $\mathcal{P}_s$ in Schritt (2b) und (2c) nur ein Orakel (SEP) für $\mathcal{P}_s$ braucht. Insgesamt ist also (OPT) orakel-polynomial in (SEP).

(OPT) $\rightarrow$ (VIOL).

Gegeben sei ein Polyeder $\mathcal{P}$, wobei $\mathcal{P}$ beschränkt und volldimensional sei. Weiterhin seien ein Vektor $c \in \mathbb{Q}^n$ und eine Zahl $\gamma \in \mathbb{Q}$ gegeben. Wir haben nun zu entscheiden, ob die Ungleichung $c^T x \leq \gamma$ für $x \in P$ erfüllt ist. Ist dies nicht der Fall, so brauchen wir einen Vektor $y \in \mathcal{P}$ mit $c^T y > \gamma$. Als erstes fragen wir das Orakel (OPT) nach einer Lösung für das Problem $\max\{c^T x \mid x \in \mathcal{P}\}$. Das Orakel wird uns eine Optimallösung y zurückgeben, da $\mathcal{P}$ beschränkt und volldimensional ist. Gilt nun $c^T y \leq \gamma$, so haben wir einen Beweis dafür, dass die Ungleichung $c^T x \leq \gamma$ für alle $x \in \mathcal{P}$ gültig ist. Andernfalls haben wir einen Vektor y gefunden mit $c^T y > \gamma$.

(VIOL) → (SEP).

Man betrachtet das Polyeder aller gültigen Ungleichungen für $\mathcal{P}$. Der Beweis ist dann elementar. Wir verzichten auf Details. $\square$

Kapitel 7

Optimalitätsbedingungen für nichtlineare Probleme

In der Nichtlinearen Optimierung betrachtet man Optimierungsprobleme der Form

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0 \quad (i = 1, \ldots, m) \\ & h_j(x) = 0 \quad (j = 1, \ldots, k) \end{array} \tag{NLP}$$

wobei f, $g_i, h_j : \mathbb{R}^n \to \mathbb{R}$, $f, g_i, h_j \in C^1$. Man setzt hierbei nicht mehr voraus, dass die Funktionen linear oder konvex sind.
Notation: Die zulässige Menge eines nichtlinearen Problems bezeichnen wir wieder mit $\mathcal{X}$, d.h.

$$\mathcal{X} = \{x \in \mathbb{R}^n : g_i(x) \leq 0 \ (i = 1, \ldots, m), h_j(x) = 0 \ (j = 1, \ldots, k)\}.$$

Wir beschränken uns hier auf den Fall linearer Nebenbedingungen, also Probleme der Form

$$\min f(x) \quad \text{s.t.} \quad Ax \leq b \tag{NP}$$

mit $A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$. Der zulässige Bereich ist dann ein Polyeder, genauer $\mathcal{X} = \mathcal{P}(A, b)$.
Dies vermeidet einige Komplikationen bei der Optimalitätstheorie für Probleme mit nichtlinearen Nebenbedingungen. Der allgemeine Fall wird ausführlich in der Vorlesung über Nichtlineare Optimierung behandelt.
Bei nichtlinearen Optimierungsproblemen fehlen viele der Eigenschaften, die LPs so gut lösbar machen:

- Ein lokaler Minimalpunkt eines LPs ist gleichzeitig schon ein globaler Minimalpunkt des LPs.
- Die zulässige Menge enthält endlich viele ausgezeichnete Randpunkte (Ecken) mit der Eigenschaft: Besitzt die Zielfunktion ein Minimum, dann wird es in einem dieser Extrempunkte angenommen.
- Mit dem Simplex-Algorithmus gibt es ein Verfahren, das in endlich vielen Schritten entweder eine Optimallösung des Problems liefert oder feststellt, dass das Problem unbeschränkt ist.

All diese Eigenschaften treffen auf nichtlineare Optimierungsprobleme im Allgemeinen nicht zu.

7.1 Optimalitätsbedingungen

Zu Optimalitätsbedingungen für nichtlineare Optimierungsprobleme gelangt man ähnlich, wie wir das in Satz 4.4 für LPs gemacht haben, durch folgende Überlegung: Ist $\bar{x} \in \mathcal{X}$ ein lokales Minimum, dann existiert $\epsilon > 0$ so dass

$$f(\bar{x} + s) \geq f(\bar{x}) \quad \forall\, s \in \mathbb{R}^n, \quad \bar{x} + s \in \mathcal{X}, \quad \|s\| \leq \epsilon.$$

7.1.1 Tangentialkegel und Linearisierungskegel

Wir definieren zunächst in einem Punkt $x^* \in \mathcal{X}$ die Menge aller Richtungen s, so dass $x^* + \lambda s$ für kleine $\lambda > 0$ wieder in $\mathcal{X}$ liegt:

Definition 7.1 *Sei x^* ein zulässiger Punkt des nichtlinearen Optimierungsproblems (NP)* $\min\{f(x) : x \in \mathcal{X}\}$. *Ein Vektor $s \in \mathbb{R}^n$ heißt* zulässige Richtung *in x^*, falls eine reelle Zahl $\bar{\lambda} > 0$ existiert, so dass*

$$x^* + \lambda s \in \mathcal{X} \;\; \textit{für alle} \;\; \lambda \in [0, \bar{\lambda}].$$

Ferner heißt

$$\mathcal{Z}(x^*) = \overline{\{s : \quad s \;\; \textit{ist zulässige Richtung in } x^*\}}$$

Tangentialkegel von $\mathcal{X}$ in x^* (oder auch Kegel der zulässigen Richtungen in x^*).

Wir erhalten die folgende erste Optimalitätsbedingung.

Satz 7.2 *Es sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar. Dann gilt für jede lokale Lösung $\bar{x}$ des (NP)*

$$\min f(x) \quad \text{s.t.} \quad x \in \mathcal{X}$$

die notwendige Optimalitätsbedingung

$$\bar{x} \in \mathcal{X} \quad \textit{und} \quad \nabla f(\bar{x})^T s \geq 0 \quad \forall\, s \in \mathcal{Z}(\bar{x}). \tag{7.1}$$

Beweis. $\bar{x} \in \mathcal{X}$ ist klar. Sei nun s eine beliebige zulässige Richtung in $\bar{x}$. Dann gibt es $\bar{\lambda} > 0$ mit $\bar{x} + \lambda s \in \mathcal{X}$ für alle $\lambda \in [0, \bar{\lambda}]$. Da $\bar{x}$ ein lokales Minimum ist, können wir $\bar{\lambda} > 0$ ohne Einschränkung so wählen, dass gilt

$$f(\bar{x} + \lambda s) \geq f(\bar{x}) \quad \forall\, \lambda \in (0, \bar{\lambda}].$$

Dies ergibt

$$0 \leq \frac{f(\bar{x} + \lambda s) - f(\bar{x})}{\lambda} \overset{\lambda \searrow 0}{\rightarrow} \nabla f(\bar{x})^T s.$$

Somit gilt (7.1) für alle zulässigen Richtungen s in $\bar{x}$ und somit auch für den Abschluß $\mathcal{Z}(\bar{x})$ all dieser Richtungen. □

Der Tangentialkegel ist unhandlich. Um die Optimalitätsbedingung (7.1) zu vereinfachen, würden wir ihn gerne bequemer darstellen.
In unserem Fall von linearen Nebenbedingungen stimmt der Tangentialkegel mit dem sogenannten Linearisierungskegel überein.
Um den Begriff Linearisierungskegel plausibel zu machen, betrachten wir zunächst das allgemeine Problem

$$\min f(x) \quad \text{s.t.} \quad g_i(x) \leq 0, \quad i = 1, \ldots, m. \tag{NLPI}$$

Wir benötigen zunächst die folgende Definition.

Definition 7.3 *Sei $\mathcal{X} = \{x \in \mathbb{R}^n : g_i(x) \leq 0\ (i = 1, \ldots, m)\}$. Sei $x^* \in \mathcal{X}$. Eine Nebenbedingungsfunktion $g_i(x)$ mit $g_i(x^*) = 0$ heißt in x^** aktiv *(oder auch* bindend*). Die Menge*

$$eq(x^*) = \{i \in \{1, \ldots, m\} : g_i(x^*) = 0\}$$

heißt die Indexmenge der in x^* aktiven Nebenbedingungen.

Beispiel 7.4 *Sei $\mathcal{X}$ definiert durch die drei Funktionen*

$$\begin{array}{rcrcl} g_1(x) & = & x_1^2 + x_2^2 - 1 & \leq & 0 \\ g_2(x) & = & -x_1 & \leq & 0 \\ g_3(x) & = & -x_2 & \leq & 0. \end{array}$$

Für $x^* = (1,0)$ gilt $eq(x^*) = \{1,3\}$, für $x^* = (0,0)$ gilt $eq(x^*) = \{2,3\}$, für $x^* = (\frac{3}{5}, \frac{4}{5})$ gilt $eq(x^*) = \{1\}$, für $x^* = (\frac{1}{2}, \frac{1}{2})$ gilt $eq(x^*) = \emptyset$.

Definition 7.5 *Seien $g_1, \ldots, g_m \in C^1$, sei $\mathcal{X} = \mathcal{X}(g) = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \ldots, m\}$ und sei $x^* \in \mathcal{X}$. Die Menge*

$$\mathcal{L}_{\mathcal{X}(g)}(x^*) = \{s \in \mathbb{R}^n : \nabla g_i(x^*)^T s \leq 0 \ \textit{ für alle } \ i \in eq(x^*)\}$$

heißt Linearisierungskegel von $\mathcal{X}$ in x^* zur Darstellung $\mathcal{X}(g)$.

Ist der zulässige Bereich $\mathcal{X}$ ein Polyeder $\mathcal{P}(A,b)$, dann ist der **Linearisierungskegel** in x^* gegeben durch

$$\mathcal{L}_{\mathcal{P}(A,b)}(x^*) = \{s \ : \ A_{\mathrm{eq}(x^*)\cdot} s \leq 0\}$$

und stimmt mit dem Tangentialkegel überein, wie das folgende Lemma zeigt.

Lemma 7.6 *Es sei $\mathcal{X} = \mathcal{P}(A,b)$ mit $A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$. Ist $x^* \in \mathcal{X}$, dann gilt*

$$\mathcal{Z}(x^*) = \mathcal{L}_{\mathcal{P}(A,b)}(x^*) = \{s \ : \ A_{eq(x^*)\cdot} s \leq 0\}.$$

Beweis. "$\subset$": Sei s eine zulässige Richtung in x^*. Dann gilt mit einem $\bar{\lambda} > 0$

$$A_{\mathrm{eq}(x^*)\cdot} x^* = b_{\mathrm{eq}(x^*)}, \quad A_{\mathrm{eq}(x^*)\cdot}(x^* + \lambda s) \leq b_{\mathrm{eq}(x^*)} \quad \forall\, \lambda \in [0, \bar{\lambda}].$$

Subtraktion der ersten Gleichung von der zweiten Ungleichung ergibt

$$A_{\mathrm{eq}(x^*)\cdot} \lambda s \leq 0 \quad \forall\, \lambda \in [0, \bar{\lambda}].$$

Damit liegt jede zulässige Richtung s in der abgeschlossenen Menge $\mathcal{L}_{\mathcal{P}(A,b)}(x^*)$ und somit auch der Abschluss $\mathcal{Z}(x^*)$ der Menge dieser Richtungen.

"$\supset$": Es gelte $s \in \mathcal{L}_{\mathcal{P}(A,b)}(x^*)$. Für $\bar{\lambda} > 0$ klein genug gilt dann (inaktive Nebenbedingungen bleiben inaktiv)

$$\begin{array}{l} A_{\mathrm{eq}(x^*)\cdot}(x^* + \lambda s) = b_{\mathrm{eq}(x^*)} + \lambda A_{\mathrm{eq}(x^*)\cdot} s \leq b_{\mathrm{eq}(x^*)}, \\ A_{(\{1,\ldots,m\}\setminus \mathrm{eq}(x^*))\cdot}(x^* + \lambda s) < b_{(\{1,\ldots,m\}\setminus \mathrm{eq}(x^*))} \quad \forall\, \lambda \in [0, \bar{\lambda}]. \end{array}$$

Somit gilt $s \in \mathcal{Z}(x^*)$ □

Bemerkung 7.7 Im Fall nichtlinearer Nebenbedingungen $\mathcal{X} = \mathcal{X}(g) = \{x \in \mathbb{R}^n : g_i(x) \leq 0, Mi = 1, \ldots, m\}$ mit $g_i \in C^1$ gilt immer $\mathcal{Z}(x^*) \subset \mathcal{L}_{\mathcal{X}(g)}(x^*)$, wobei $\mathcal{Z}(x^*)$ genau wie in Definition 7.1 definiert ist.
Die wichtige Inklusion $\mathcal{Z}(x^*) \supset \mathcal{L}_{\mathcal{X}(g)}(x^*)$, die es erlaubt, in der Optimalitätsbedingung (7.1) den Tangentialkegel $\mathcal{Z}(x^*)$ durch den Linearisierungskegel zu ersetzen, gilt jedoch im Fall nichtlinearer Nebenbedingungen nur unter zusätzlichen Bedingungen, einer sogenannten Constraint Qualification. Details werden in Optimierung 3 behandelt.

Aus der Optimalitätsbedingung (7.1) erhalten wir unmittelbar im Fall linearer Nebenbedingungen

Korollar 7.8 *Es sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar, $A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$. Ist $\bar{x}$ eine lokale Lösung des Problems*

$$\min f(x) \quad \text{s.t.} \quad x \in \mathcal{P}(A, b),$$

dann gelten die notwendige Optimalitätsbedingung

$$A\bar{x} \leq b \quad \textit{und} \quad \nabla f(\bar{x})^T s \geq 0 \quad \forall\, s \in \{s \ : \ A_{eq(\bar{x})\cdot} s \leq 0\}. \tag{7.2}$$

Beweis. Nach Lemma 7.6 folgt dies sofort aus Satz 7.2 □

7.1.2 Karush-Kuhn-Tucker-Bedingungen

Wir wenden nun das Farkas-Lemma an, um (7.2) in ein Gleichungs- und Ungleichungssystem zu verwandeln.

Satz 7.9 (Karush-Kuhn-Tucker-Bedingungen für (NP)) *Seien $f \in C^1$, $A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$. $\bar{x}$ sei ein lokaler Minimalpunkt von*

$$\min f(x) \quad \text{s.t.} \quad Ax \leq b.$$

Dann gibt es einen Vektor $\bar{u} \in \mathbb{R}^m$ mit

$$\begin{aligned} \nabla f(\bar{x}) + A^T \bar{u} = 0, &\quad \textit{(Multiplikatorregel)} \\ A\bar{x} \leq b, &\quad \textit{(Zulässigkeit)} \\ \bar{u} \geq 0, \quad (A\bar{x} - b)_i \bar{u}_i = 0, &\quad i = 1, \ldots, m, \quad \textit{(Komplementarität).} \end{aligned} \tag{7.3}$$

Beweis. Die zweite Bedingung gilt, weil $\bar{x}$ zulässig sein muss.
Setze $I = \mathrm{eq}(\bar{x})$. Nach Korollar 7.8 hat dann das System

$$A_{I\cdot}s \leq 0, \quad \nabla f(\bar{x})^T s < 0$$

keine Lösung. Daher existiert nach dem Farkas-Lemma (Satz 4.5) ein z mit

$$-A_{I\cdot}^T z = \nabla f(\bar{x}), \quad z \geq 0.$$

Definieren wir nun $\bar{u} \in \mathbb{R}^m$ durch $\bar{u}_I = z$, $\bar{u}_{\{1,\dots,m\}\setminus I} = 0$, dann gilt

$$A^T \bar{u} = A_{I\cdot}^T \bar{u}_I = -\nabla f(\bar{x}), \quad \bar{u} \geq 0,$$

Weiter ist wegen $A_{I\cdot}\bar{x} = b_I$ und $\bar{u}_{\{1,\dots,m\}\setminus I} = 0$ auch die Komplementaritätsbedingung

$$(A\bar{x} - b)_i \bar{u}_i = 0, \quad i = 1, \dots, m$$

erfüllt. □

Die Komponenten des Vektors $\bar{u}$ heißen Lagrange-Multiplikatoren. Ein Punkt $\bar{x}$, der die Karush-Kuhn-Tucker-Bedingungen erfüllt, heißt Karush-Kuhn-Tucker-Punkt, auch kurz: KKT-Punkt.

Wir können die KKT-Bedingungen geometrisch so deuten:

Minus Gradient der Zielfunktion liegt im Kegel, der von den Gradienten der in $\bar{x}$ aktiven Nebenbedingungen aufgespannt wird.

Wir zeigen jetzt, dass die KKT-Bedingungen (7.3) bei konvexer Zielfunktion hinreichend sind, und zwar für globale Minimalpunkte.

Satz 7.10 (KKT-Bedingungen im konvexen Fall) *Sei $f \in C^1$ konvex. Betrachte das Optimierungsproblem (NP). Wenn $\bar{x} \in \mathbb{R}^n$ und $\bar{u} \in \mathbb{R}^m$ die KKT-Bedingungen (7.3) erfüllen, dann ist $\bar{x}$ (lokaler = globaler) Minimalpunkt des Problems.*

Beweis. Sei $\bar{x}$ ein Punkt, der die KKT-Bedingungen erfüllt. Sei $x \in \mathcal{X}$ beliebig. Wegen $\bar{u} \geq 0$ und der Linearität der Nebenbedingungen folgt mit den KKT-Bedingungen

$$\bar{u}^T A(x - \bar{x}) \leq \bar{u}^T (Ax - b - (A\bar{x} - b))) = \bar{u}^T (Ax - b) \leq 0.$$

Zusammen mit der Konvexität von f, Satz 2.32 und (7.3), 2) ergibt sich

$$f(x) - f(\bar{x}) \geq \nabla f(\bar{x})^T (x - \bar{x}) = -\bar{u}^T A(x - \bar{x}) \geq 0.$$

□

Bemerkung 7.11 Der Satz läßt sich auf (NLPI) erweitern, falls $f, g_i \in C^1$ konvex sind. Siehe Optimierung 3.

Schließlich betrachten wir noch den Fall von Gleichungs- und Ungleichungsnebenbedingungen. Wir erhalten das folgende Resultat:

Satz 7.12 (Karush-Kuhn-Tucker-Bedingungen für (NPGU)) *Sei $f \in C^1$, $A \in \mathbb{R}^{m,n}$, $B \in \mathbb{R}^{k,n}$, $b \in \mathbb{R}^m$, $d \in \mathbb{R}^k$. $\bar{x}$ sei ein lokaler Minimalpunkt von*

$$\min f(x) \quad \text{s.t.} \quad x \in \mathcal{X} := \{x \ : Ax \leq b, \quad Bx = d\}. \tag{NPGU}$$

Dann gibt es Vektoren $\bar{u} \in \mathbb{R}^m$, $\bar{v} \in \mathbb{R}^k$ mit

$$\begin{aligned} \nabla f(\bar{x}) + A^T \bar{u} + B^T \bar{v} &= 0 \\ A\bar{x} &\leq b, \\ B\bar{x} &= d, \\ \bar{u} \geq 0, \quad (A\bar{x} - b)_i \bar{u}_i = 0, \quad i &= 1, \ldots, m. \end{aligned} \tag{7.4}$$

Beweis. Wir schreiben $Bx - d = 0$ als $\binom{Bx}{-Bx} \leq \binom{d}{-d}$. Dann liefert Satz 7.9 Multiplikatoren $\bar{u}, \bar{v}^+, \bar{v}^- \geq 0$ mit

$$\begin{aligned} &\nabla f(\bar{x}) + A^T \bar{u} + B^T (\bar{v}^+ - \bar{v}^-) = 0 \\ &\bar{u}_i (A\bar{x} - b)_i = 0, \quad i = 1, \ldots, m, \\ &\bar{v}_j^+ (B\bar{x} - d)_j = 0, \quad \bar{v}_j^- (B\bar{x} - d)_j = 0, \quad j = 1, \ldots, k, \end{aligned}$$

wobei die letzte Zeile wegen $B\bar{x} = d$ nichts neues bringt. Mit $\bar{v} = \bar{v}^+ - \bar{v}^-$ gilt also (7.4). □

Kapitel 8

Quadratische Probleme

Quadratische Probleme bilden eine wichtige Klasse von Optimierungsproblemen, die sowohl von unabhängigem Interesse sind als auch als Teilprobleme bei Lösungsverfahren für allgemeine nichtlineare Probleme auftreten, zum Beispiel beim so genannten Sequential Quadratic Programming-Verfahren (SQP).
Zunächst betrachten wir gleichheitsrestringierte quadratische Probleme, danach behandeln wir zusätzlich auftretende Ungleichungsnebenbedingungen mit der sogenannten Strategie der aktiven Menge. Grundlegend für beides sind die KKT-Bedingungen.

8.1 Probleme mit Gleichheitsrestriktionen

Wir betrachten zunächst ein quadratisches Problem mit Gleichheitsrestriktionen

$$\begin{array}{ll} \min & f(x) := \frac{1}{2}x^T Q x + c^T x \\ \text{s.t.} & b_j^T x = \beta_j \quad (j = 1, \dots, p), \end{array} \tag{8.1}$$

wobei $Q \in \mathbb{R}^{n \times n}$ symmetrisch, $c \in \mathbb{R}^n$, $b_j \in \mathbb{R}^n$ und $\beta_j \in \mathbb{R}$ $(j = 1, \dots, p)$ gegeben sind.
Sei $\bar{x}$ ein lokaler Minimalpunkt von (8.1). Alle Nebenbedingungen sind linear, daher existieren nach Satz 7.12 Lagrange-Multiplikatoren $\bar{v}_j \in \mathbb{R}$ $(j = 1, \dots, p)$, so dass das Paar $(\bar{x}, \bar{v})$ den KKT-Bedingungen

$$\begin{aligned} Qx + c + \sum_{j=1}^{p} v_j b_j &= 0, \\ b_j^T x &= \beta_j \quad (j = 1, \dots, p) \end{aligned}$$

von (8.1) genügt. Bezeichnen wir mit $B \in \mathbb{R}^{p \times n}$ die Matrix mit den Vektoren b_j^T als Zeilenvektoren und setzen $\beta := (\beta_1, \ldots, \beta_p)^T$, so lässt sich das obige Gleichungssystem formulieren als

$$\begin{aligned} Qx + B^T v &= -c, \\ Bx &= \beta. \end{aligned}$$

Dies ist ein lineares Gleichungssystem zur Berechnung eines KKT-Punktes von (8.1). Damit ist das folgende Resultat bewiesen:

Satz 8.1 *Ein Paar $(\bar{x}, \bar{v}) \in \mathbb{R}^n \times \mathbb{R}^p$ ist genau dann ein KKT-Punkt des Optimierungsproblems (8.1), wenn $(\bar{x}, \bar{v})$ Lösung des linearen Gleichungssystems*

$$\begin{pmatrix} Q & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ v \end{pmatrix} = \begin{pmatrix} -c \\ \beta \end{pmatrix} \tag{8.2}$$

ist.

Im Fall einer positiv semi-definiten Matrix Q (so dass die Zielfunktion konvex ist, siehe Satz 2.33) sind die KKT-Bedingungen von (8.1) vollständig äquivalent zum eigentlichen Optimierungsproblem (8.1), vgl. die Sätze 7.12 und 7.10. In diesem Fall lässt sich die Lösung eines gleichheitsrestringierten quadratischen Optimierungsproblems also auf die Lösung des zugehörigen linearen Gleichungssystems (8.2) reduzieren.

Für den nächsten Abschnitt ist es praktisch, den Inhalt des Satzes 8.1 noch etwas umzuformulieren: Schreiben wir $x = x^k + \Delta x$ im Gleichungssystem (8.2) mit einem für das quadratische Problem (8.1) *zulässigen* Vektor x^k sowie einem Korrekturterm $\Delta x \in \mathbb{R}^n$, so ergeben sich aus (8.2) die folgenden

Äquivalenzen:

$$
\begin{aligned}
&\begin{pmatrix} Q & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ v \end{pmatrix} = \begin{pmatrix} -c \\ \beta \end{pmatrix} \\
&\qquad \Longleftrightarrow \begin{pmatrix} Q & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x^k + \Delta x \\ v \end{pmatrix} = \begin{pmatrix} -c \\ \beta \end{pmatrix} \\
&\qquad \Longleftrightarrow \begin{pmatrix} Q & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ v \end{pmatrix} = \begin{pmatrix} -c \\ \beta \end{pmatrix} - \begin{pmatrix} Q & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x^k \\ 0 \end{pmatrix} \\
&\qquad \Longleftrightarrow \begin{pmatrix} Q & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ v \end{pmatrix} = \begin{pmatrix} -c - Qx^k \\ \beta - Bx^k \end{pmatrix} \\
&\qquad \Longleftrightarrow \begin{pmatrix} Q & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ v \end{pmatrix} = \begin{pmatrix} -\nabla f(x^k) \\ 0 \end{pmatrix},
\end{aligned}
$$

denn x^k ist zulässig für (8.1), d.h. $Bx^k = \beta$, und es ist $\nabla f(x) = Qx + c$. Somit haben wir die nachstehende Formulierung des Satzes 8.1, die bei der Behandlung quadratischer Probleme mit Gleichheits- und Ungleichheitsrestriktionen im folgenden Abschnitt hilfreich sein wird.

Satz 8.2 *Sei $x^k \in \mathbb{R}^n$ ein zulässiger Punkt für das quadratische Optimierungsproblem (8.1). Dann ist $(\bar{x}, \bar{v}) \in \mathbb{R}^n \times \mathbb{R}^p$ genau dann ein KKT-Punkt von (8.1), wenn $\bar{x} = x^k + \Delta x^*$ gilt und $(\Delta x^*, \bar{v})$ eine Lösung des linearen Gleichungssystems*

$$
\begin{pmatrix} Q & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ v \end{pmatrix} = \begin{pmatrix} -\nabla f(x^k) \\ 0 \end{pmatrix}
$$

ist.

8.2 Strategie der aktiven Menge für Ungleichungen

Im Folgenden untersuchen wir das quadratische Problem mit Gleichungs- und Ungleichungsrestriktionen

$$\begin{aligned} \min \quad & f(x) := \tfrac{1}{2}x^TQx + c^Tx \\ \text{s.t.} \quad & b_j^Tx = \beta_j \quad (j = 1, \dots, p), \\ & a_i^Tx \le \alpha_i \quad (i = 1, \dots, m) \end{aligned} \tag{8.3}$$

mit $Q \in \mathbb{R}^{n\times n}$ symmetrisch, $c \in \mathbb{R}^n$, a_i, $b_j \in \mathbb{R}^n$ und α_i, $\beta_j \in \mathbb{R}\ \forall i, j$.

Die wesentliche Idee zur Lösung von (8.3) besteht darin, dass man (unter Benutzung von Satz 8.2) eine Folge von gleichheitsrestringierten Problemen löst, welche sich dadurch ergeben, dass man in (8.3) nur die im aktuellen Iterationspunkt aktiven Restriktionen berücksichtigt.

Dazu definiert man in Iteration k eine geeignete Approximation $\mathcal{A}_k$ an die Indexmenge

$$\mathrm{eq}(x^k) := \{i \in \{1, \dots, m\} : a_i^Tx^k = \alpha_i\}$$

der in x^k aktiven Ungleichungsnebenbedingungen von (8.3) und berechnet mit Satz 8.2 einen KKT-Punkt des Hilfsproblems

$$\begin{aligned} \min \quad & \tfrac{1}{2}x^TQx + c^Tx \\ \text{s.t.} \quad & b_j^Tx = \beta_j \quad (j = 1, \dots, p), \\ & a_i^Tx = \alpha_i \quad (i \in \mathcal{A}_k) \end{aligned}$$

Ist dieser Punkt auch ein KKT-Punkt von (8.3), dann sind wir fertig, andernfalls muss x^k bzw. $\mathcal{A}_k$ geeignet modifiziert werden.

Im folgenden sei wie vorher $B \in \mathbb{R}^{p\times n}$ die Matrix mit den Zeilen b_j^T ($j = 1, \dots, p$), und $A_k \in \mathbb{R}^{|\mathcal{A}_k|\times n}$ sei die Matrix mit den Zeilen a_i^T ($i \in \mathcal{A}_k$).

Algorithmus 8.3 (Strategie der aktiven Menge für quadratische Probleme)

Input: *ein quadratisches Problem der Form (8.3)*

Output: *ein KKT-Punkt für (8.3)*

Schritt 0: *Bestimme ein für (8.3) zulässiges* $x^0 \in \mathbb{R}^n$ *sowie zugehörige Lagrange-Multiplikatoren* $u^0 \in \mathbb{R}^m$ *und* $v^0 \in \mathbb{R}^p$*, setze* $\mathcal{A}_0 := \{i : a_i^T x^0 = \alpha_i\}$ *und* $k := 0$.

Schritt 1: *Ist* (x^k, u^k, v^k) *ein KKT-Punkt von (8.3): STOP.*

Schritt 2: *Setze* $u_i^{k+1} := 0$ *für* $i \notin \mathcal{A}_k$ *und bestimme* $(\Delta x^k, u_{\mathcal{A}_k}^{k+1}, v^{k+1})$ *als Lösung des linearen Gleichungssystems*

$$\begin{pmatrix} Q & A_k^T & B^T \\ A_k & 0 & 0 \\ B & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ u_{\mathcal{A}_k} \\ v \end{pmatrix} = \begin{pmatrix} -\nabla f(x^k) \\ 0 \\ 0 \end{pmatrix} \tag{8.4}$$

Schritt 3: *Unterscheide folgende Fälle:*

(a) *Ist* $\Delta x^k = 0$ *und* $u_i^{k+1} \geq 0$ *für alle* $i \in \mathcal{A}_k$*: STOP.*

(b) *Ist* $\Delta x^k = 0$ *und* $\min\{u_i^{k+1} : i \in \mathcal{A}_k\} < 0$*, so bestimme einen Index* q *mit* $u_q^{k+1} = \min\{u_i^{k+1} : i \in \mathcal{A}_k\}$*, setze*

$$x^{k+1} := x^k, \quad \mathcal{A}_{k+1} := \mathcal{A}_k \setminus \{q\},$$

und gehe zu Schritt 4.

(c) *Ist* $\Delta x^k \neq 0$ *und* $x^k + \Delta x^k$ *zulässig für (8.3), so setze*

$$x^{k+1} := x^k + \Delta x^k, \quad \mathcal{A}_{k+1} := \mathcal{A}_k,$$

und gehe zu Schritt 4.

(d) *Ist* $\Delta x^k \neq 0$ *und* $x^k + \Delta x^k$ *nicht zulässig für (8.3), so bestimme einen Index* $r \notin \mathcal{A}_k$*, mit* $a_r^T \Delta x^k > 0$ *und*

$$t_k := \frac{\alpha_r - a_r^T x^k}{a_r^T \Delta x^k} = \min\left\{\frac{\alpha_i - a_i^T x^k}{a_i^T \Delta x^k} : i \notin A_k \ \ \textit{mit} \ \ a_i^T \Delta x^k > 0\right\},$$

setze

$$x^{k+1} := x^k + t_k \Delta x^k, \quad \mathcal{A}_{k+1} := \mathcal{A}_k \cup \{r\},$$

und gehe zu Schritt 4.

Schritt 4: *Setze* $k \longleftarrow k + 1$*, und gehe zu Schritt 1.*

Einige Erklärungen zum Algorithmus 8.3:

Zu Schritt 0: Ein zulässiger Punkt x^0 für das Problem (8.3) kann mittels Phase I des Simplex-Algorithmus gefunden werden, vgl. Abschnitt 5.3.

Zu Schritt 2: Da die aktuelle Iterierte x^k für das quadratische Problem (8.3) zulässig ist (der Startvektor x^0 ist zulässig wegen Schritt 0, und die Zulässigkeit von x^k ergibt sich induktiv in Schritt 3), folgt aus dem Satz 8.2, dass der Punkt $(x^k + \Delta x^k, u_{\mathcal{A}_k}^{k+1}, v^{k+1})$ aus Schritt 2 ein KKT-Punkt des gleichheitsrestringierten Problems

$$\begin{array}{ll} \min & \frac{1}{2}x^T Q x + c^T x \\ \text{s.t.} & b_j^T x = \beta_j \quad (j = 1, \ldots, p) \\ & a_i^T x = \alpha_i \quad (i \in \mathcal{A}_k) \end{array} \tag{QP$_k$}$$

ist.

Zu Schritt 3:
Fall (a): Im Fall $\Delta x^k = 0$ und $u_i^{k+1} \geq 0$ für alle $i \in \mathcal{A}_k$ erkennt man sofort, dass das Tripel

$$(x^k, u^{k+1}, v^{k+1}) \quad \text{mit} \quad u_i^{k+1} = 0 \text{ für } i \notin \mathcal{A}_k$$

auch ein KKT-Punkt des eigentlichen Problems (8.3) ist. Dies erklärt insbesondere das Abbruchkriterium im Schritt 3(a).

Fall (b): Ist dagegen $\Delta x^k = 0$ und $u_i^{k+1} < 0$ für ein $i \in \mathcal{A}_k$ (dies bedeutet, dass wir einerseits noch nicht in einem KKT-Punkt von (8.3) sein können, dass andererseits auf der aktuellen Restriktionsmenge die Zielfunktion aber nicht weiter verringert werden kann), so lockern wir die Restriktionen und entfernen einen Index aus der Menge $\mathcal{A}_k$.
Dieser Schritt heißt Inaktivierungsschritt. Dabei entnehmen wir einen solchen Index $q \in \mathcal{A}_k$, für den der zugehörige Lagrange-Multiplikator am stärksten negativ ist.

Fall (c): Ist $\Delta x^k \neq 0$ und $x^k + \Delta x^k$ zulässig für (8.3), so akzeptieren wir natürlich den Punkt $(x^k + \Delta x^k, u^{k+1}, v^{k+1})$ als neue Iterierte, ohne dabei die Indexmenge $\mathcal{A}_k$ zu verändern.

Fall (d): Ist $x^k + \Delta x^k$ hingegen nicht zulässig für (8.3), so ist eine der bislang strikt erfüllten Ungleichungen verletzt. Statt eines vollen Schrittes $x^k + \Delta x^k$ setzen wir daher

$$x^{k+1} := x^k + t_k \Delta x^k$$

mit einer Schrittweite $t_k > 0$, welche gerade garantiert, dass auch die Ungleichungen $i \notin \mathcal{A}_k$ im neuen Punkt x^{k+1} erfüllt sind. Dies liefert die Forderung

$$a_i^T x^{k+1} = a_i^T x^k + t_k a_i^T \Delta x^k \leq \alpha_i \quad \forall i \notin \mathcal{A}_k.$$

Da $a_i^T x^k \leq \alpha_i$ gilt, ist diese Forderung für Indizes $i \notin \mathcal{A}_k$ mit $a_i^T \Delta x^k \leq 0$ mit jedem $t_k > 0$ automatisch erfüllt.
Für Indizes $i \notin \mathcal{A}_k$ mit $a_i^T \Delta x^k > 0$ liefert die obige Forderung gerade

$$t_k \leq \frac{\alpha_i - a_i^T x^k}{a_i^T \Delta x^k},$$

also

$$t_k = \min\left\{\frac{\alpha_i - a_i^T x^k}{a_i^T \Delta x^k} : i \notin \mathcal{A}_k \text{ mit } a_i^T \Delta x^k > 0\right\}.$$

Wir nehmen dadurch eine neue aktive Nebenbedingung zur Indexmenge $\mathcal{A}_k$ hinzu. Dieser Schritt wird auch Aktivierungsschritt genannt.
Bleibt noch die Frage, ob es im Schritt 3(d) immer einen Index $i \notin \mathcal{A}_k$ mit $a_i^T \Delta x^k > 0$ gibt. Dies muss aber so sein, denn andernfalls wäre (wegen $a_i^T x^k \leq \alpha_i \ \forall i$)

$$a_i^T (x^k + \Delta x^k) \leq \alpha_i \quad \forall i \notin \mathcal{A}_k.$$

Da überdies wegen (8.4) auch $a_i^T \Delta x^k = 0$ und somit

$$a_i^T (x^k + \Delta x^k) \leq \alpha_i \quad \forall i \in \mathcal{A}_k$$

gilt, wäre $x^k + \Delta x^k$ zulässig für Problem (8.3), ein Widerspruch zur Annahme in Schritt 3(d).

Zur Illustration des Algorithmus 8.3 betrachten wir ein Beispiel:

Beispiel 8.4 *Betrachten wir das Problem*

$$\begin{array}{ll}
\min & \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + 2x_1 + x_2 \\
\text{s.t.} & \begin{array}{rcrcl}
-x_1 & - & x_2 & \leq & 0, \\
 & & x_2 & \leq & 2, \\
x_1 & + & x_2 & \leq & 5, \\
-x_1 & + & x_2 & \leq & 2, \\
x_1 & & & \leq & 5, \\
 & - & x_2 & \leq & 1.
\end{array}
\end{array}$$

Wir haben: $Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Als zulässiger Startpunkt sei $x^0 := (5,0)^T$ *gewählt. Als anfänglichen Lagrange-Multiplikator setzen wir* $u^0 := (0,0,0,0,0,0)^T$ *(dies hat auf den weiteren Verlauf der Iteration aber keinen Einfluss). Somit ist* $\mathcal{A}_0 = \{3,5\}$.

Iteration 1: Schritt 1: (x^0, u^0) *ist kein KKT-Punkt, da*

$$\nabla f(x^0) + \sum_{i=1}^{6} 0 \cdot \nabla g_i(x^0) = \binom{7}{1} \neq \binom{0}{0}.$$

Schritt 2: Löse das Gleichungssystem

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ u_3 \\ u_5 \end{pmatrix} = \begin{pmatrix} -7 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

Lösung ist $\Delta x^0 = (0,0)$, $u_3 = -1$ *und* $u_5 = -6$.
Schritt 3: Fall (b), mit $q = 5$. *Daher:*

$$x^1 := (5,0), \qquad \mathcal{A}_1 := \{3\}.$$

Iteration 2: Schritt 1: (x^1, u^1) *ist kein KKT-Punkt, da* $u \not\geq 0$.
Schritt 2: Löse das Gleichungssystem

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} -7 \\ -1 \\ 0 \end{pmatrix}.$$

Lösung ist $\Delta x^1 = (-3,3)$ *und* $u_3 = -4$.
Schritt 3: Fall (d), weil $(5,0) + (-3,3) = (2,3)$ *nicht zulässig (2. Nebenbedingung ist verletzt). Betrachte die Indizes* $\{i \notin \mathcal{A}_1 : a_i^T \Delta x^1 > 0\} = \{2,4\}$. *Daher*

$$t_1 = \min\left\{ \frac{2 - (0,1)\binom{5}{0}}{(0,1)\binom{-3}{3}}, \frac{2 - (-1,1)\binom{5}{0}}{(-1,1)\binom{-3}{3}} \right\} = \min\{\tfrac{2}{3}, \tfrac{7}{6}\} = \tfrac{2}{3},$$

angenommen bei $r = 2$. *Somit:*

$$x^2 := x^1 + \tfrac{2}{3}\Delta x^1 = (3,2), \qquad \mathcal{A}_2 := \mathcal{A}_1 \cup \{2\} = \{2,3\}.$$

Die Ergebnisse der Iterationen 3 bis 8 sind in der folgenden Tabelle dargestellt:

k	x^k	$\mathcal{A}_k$	$f(x^k)$	u^k
0	*(5,0)*	*{3,5}*	*22.50*	*(0,0,0,0,0,0)*
1	*(5,0)*	*{3}*	*22.50*	*(0,0,-1,0,-6,0)*
2	*(3,2)*	*{2,3}*	*14.50*	*(0,0,-4,0,0,0)*
3	*(3,2)*	*{2}*	*14.50*	*(0,2,-5,0,0,0)*
4	*(0,2)*	*{2,4}*	*4.00*	*(0,-3,0,0,0,0)*
5	*(0,2)*	*{4}*	*4.00*	*(0,-5,0,2,0,0)*
6	*(-1,1)*	*{1,4}*	*0.00*	*(0,0,0,-0.5,0,0)*
7	*(-1,1)*	*{1}*	*0.00*	*(1.5,0,0,-0.5,0,0)*
8	*(-0.5,0.5)*	*{1}*	*-0.25*	*(1.5,0,0,0,0,0)*

Wir zeigen nun, dass der Algorithmus 8.3 im Falle eines quadratischen Problemes (8.3) mit positiv definiter Matrix Q sowie linear unabhängigen Vektoren a_i $(i \in \mathcal{A}_0)$ und b_j $(j = 1, \ldots, p)$ wohldefiniert ist (d.h., die linearen Gleichungssysteme im Schritt 2 sind stets eindeutig lösbar). Dies folgt aus den Aussagen (a) und (b) des folgenden Satzes.

Satz 8.5 *Gegeben sei das quadratische Optimierungproblem (8.3) mit einer symmetrischen Matrix* $Q \in \mathbb{R}^{n \times n}$ *und* $c \in \mathbb{R}^n$, a_i, $b_j \in \mathbb{R}^n$ $(i = 1, \ldots, m,$ $j = 1, \ldots, p)$.

(a) Ist die Matrix Q *positiv definit und sind die Vektoren* a_i $(i \in \mathcal{A}_k)$, b_j $(j = 1, \ldots, p)$ *linear unabhängig, so ist das lineare Gleichungssystem (8.4) in Schritt 2 von Algorithmus 8.3 eindeutig lösbar.*

(b) Sind im k*-ten Schritt von Algorithmus 8.3 die Vektoren* a_i $(i \in \mathcal{A}_k)$, b_j $(j = 1, \ldots, p)$ *linear unabhängig und tritt in Schritt 3 kein Abbruch ein, so sind auch die Vektoren* a_i $(i \in \mathcal{A}_{k+1})$, b_j $(j = 1, \ldots, p)$ *linear unabhängig.*

(c) Ist die Matrix Q *positiv definit, so gilt für den in Schritt 3 berechneten Vektor* Δx^k *im Fall* $\Delta x^k \neq 0$

$$\nabla f(x^k)^T \Delta x^k < 0,$$

d.h. Δx^k *ist dann eine Abstiegsrichtung.*

Beweis. $\boxed{\textbf{(a)}}$ Wir zeigen, dass das homogene Gleichungssystem

$$\begin{pmatrix} Q & A_k^T & B^T \\ A_k & 0 & 0 \\ B & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

nur die triviale Lösung $(\Delta x, u, v) = 0$ besitzt. Der zweite und dritte Zeilenblock dieses Gleichungssystems liefert

$$A_k \Delta x = 0, \qquad B\Delta x = 0. \tag{8.5}$$

Der erste Zeilenblock lautet

$$Q\Delta x + A_k^T u + B^T v = 0. \tag{8.6}$$

Multiplikation von links mit Δx^T ergibt unter Verwendung von (8.5)

$$0 = \Delta x^T Q \Delta x + (A_k \Delta x)^T u + (B\Delta x)^T v = \Delta x^T Q \Delta x.$$

Wegen der positiven Definitheit von Q folgt

$$\Delta x = 0, \tag{8.7}$$

womit sich (8.6) reduziert auf

$$(A_k^T B^T) \begin{pmatrix} u \\ v \end{pmatrix} = 0.$$

Wegen der vorausgesetzten linearen Unabhängigkeit der Vektoren a_i $(i \in \mathcal{A}_k)$ und b_j $(j = 1, \ldots, p)$ sind die Spalten der Matrix $(A_k^T B^T)$ linear unabhängig, und es folgt

$$u = 0, \qquad v = 0.$$

Zusammen mit (8.7) ist somit $(\Delta x, u, v) = 0$.

$\boxed{\textbf{(b)}}$ Die Vektoren a_i $(i \in \mathcal{A}_k)$, b_j $(j = 1, \ldots, p)$ seien linear unabhängig. Wir zeigen: Die Vektoren a_i $(i \in \mathcal{A}_{k+1})$, b_j $(j = 1, \ldots, p)$ sind dann ebenfalls linear unabhängig.
Tritt in Schritt 3 der Fall (b) oder (c) ein, so ist $\mathcal{A}_{k+1} \subseteq \mathcal{A}_k$, und die Behauptung ist klar.

Tritt dagegen (d) ein, so existiert $r \notin \mathcal{A}_k$ mit $a_r^T \Delta x^k > 0$ und

$$t_k = \frac{\alpha_r - a_r^T x^k}{a_r^T \Delta x^k},$$

und es ist

$$\mathcal{A}_{k+1} \backslash \mathcal{A}_k = \{r\}.$$

Angenommen a_r ist linear abhängig von $a_i (i \in \mathcal{A}_k)$, b_j $(j = 1, \dots, p)$:

$$a_r = \sum_{i \in \mathcal{A}_k} \gamma_i a_i + \sum_{j=1}^{p} \delta_j bj$$

für gewisse Koeffizienten $\gamma_i (i \in \mathcal{A}_k)$ und $\delta_j (j = 1, \dots, p)$. Multiplikation mit Δx^k liefert

$$a_r^T \Delta x^k = \sum_{i \in \mathcal{A}_k} \gamma_i a_i^T \Delta x^k + \sum_{j=1}^{p} \delta_j b_j^T \Delta x^k,$$

woraus man wegen (8.4), zweiter und dritter Zeilenblock,

$$a_r^T \Delta x^k = 0$$

und damit einen Widerspruch zu $a_r^T \Delta x^k > 0$ erhält. Die Vektoren a_i $(i \in \mathcal{A}_{k+1})$, b_j $(j = 1, \dots, p)$ sind also linear unabhängig. $\Box$ (a), (b)

Literaturverzeichnis

[Bl77] R.G. Bland: *New finite pivoting rules for the simplex method.* Mathematics of Operations Research 2 (1977), 103-107.

[Ch83] V. Chvátal: *Linear Programming.* Freeman, New York (1983).

[Da55] G.B. Dantzig, A. Orden,P. Wolfe: *The generalized simplex method for minimizing a linear form under linear inequality restraints.* Pacific Journal of Mathematics 5 (1955), 183-195.

[FT72] J.J.H. Forrest and J.A. Tomlin: *Updated Triangular Factors of the Basis to Maintain Sparsity in the Product From Simplex Method.* Mathematical Programming 2 (1972), 263-278.

[GJ79] M.R. Garey and D.S. Johnson: *Computers and Intractability: A Guide to the Theory of NP-Completeness.* Freeman Verlag (1979).

[GK99] C. Geiger und *C. Kanzow*: *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben.* Springer Verlag (1999).

[GK00] C. Geiger und *C. Kanzow*: *Theorie und Numerik restringierter Optimierungsaufgaben.* Springer Verlag (2000).

[GMSW89] P.E. Gill, W. Murray, M.A. Saunders, M.H. Wright: *A Practical Anti-Cycling Procedure for Linearly Constrained Optimization.* Mathematical Programming 45 (1989), 437-474.

[GR77] D. Goldfarb and J.K. Reid: *A practical steepest-edge simplex algorithm.* Mathematical Programming 12 (1977), 361-371.

[GLS81] M. Grötschel, L. Lovász, A. Schrijver: *The ellipsoid method and its consequences in combinatorial optimization.* Combinatorica 1 (1981), 169-197.

[GLS88] M. Grötschel, L. Lovász, A. Schrijver: *Geometric Algorithms and Combinatorial Optimization.* Springer Verlag (1988).

[Ha73] P.M.J. Harris: *Pivot Selection Methods of the Devex LP-Code.* Mathematical Programming 5 (1973), 1-28.

[He03] H. Heuser: *Lehrbuch der Analysis, Teil 1 und Teil 2.* Teubner Verlag, mehrere Auflagen.

[Ho79] R. Horst: *Nichtlineare Optimierung.* Carl Hanser Verlag (1979).

[Kh79] L. Khachiyan: *A polynomial algorithm in linear programming.* Doklady Akademiia Nauk SSSR, 244:1093-1096 (translated in Soviet Mathematics Doklady 20:191-194,1979

[KQ63] H.W. Kuhn, R.E. Quandt: *An experimental study of the simplex method.* Experimental Arithmetic, High-Speed computing and Mathematics, American Mathematical Society, Providence, R.I. (1963), 107-124.

[NW99] J. Nocedal und S.J. Wright: *Numerical Optimization.* Springer (1999).

[PS82] C.H. Papademitrion and K. Steiglitz: *Combinatorical Optimization.* Prentice-Hall (1982).

[Ro70] R.T. Rockafellar: *Convex Analysis.* Princeton University Press (1970).

[Sch86] A. Schrijver: *Theory of Linear and Integer Programming.* John Wiley & Sons (1986).

[SS90] U.H. Suhl und L.M. Suhl: *Computing sparse LU Factorization for Large-Scale Linear Programming Bases.* ORSA Journal on Computing 2 (1990), 325-335.

[WC63] P. Wolfe, P and *L. Cutler*: *Experiments in linear programming.* Recent Advances in Mathematical Programming. McGrow-Hill, New York (1963), 177-200.

[Wr97] S. J. Wright: *Primal-dual interior-point methods.* SIAM, Philadelphia (1997).