
Vorlesung

Lineare und Diskrete Optimierung

gehalten im WS 2022/23

Prof. Dr. Tobias Harks

Universität Passau

Institut für Mathematik

Dr.-Hans-Kapfinger-Straße 30

94032 Passau

Email: tobias.harks@uni-passau.de

8. Oktober 2026

Vorwort

Dieses Skript ist im Laufe mehrerer Optimierungsvorlesungen des “Augsburger Optimierungszyklus” in den Jahren 2015-2022 entstanden und weiterentwickelt worden. Die Kapitel zur Simplex-Methode und Dualität basieren auf handschriftlichen Notizen zu den Vorlesungen [Lineare und Nichtlineare Optimierung](#) gehalten von Helmut Maurer im SS 2003 am Institut für Mathematik der Universität Münster. Kapitel 4 basiert weitestgehend auf einem Skript von Stefan Ulbrich (TU Darmstadt, Stand WS 2013), welches wiederum in Teilen auf ein älteres Skript von Martin Grötschel recurriert.

Ein besonderer Dank gilt Lukas Graf und Georg Kraus, die während der Veranstaltung im SS 2020 einige hilfreiche Korrekturen eingefügt haben.

Augsburg, Juli 2020

Prof. Dr. Tobias Harks

Inhaltsverzeichnis

1	Einführung	7
1.1	Motivation und Grundbegriffe	7
1.2	Beispiele und Anwendungen	7
1.3	Finite Optimierung	9
1.4	Differenzierbare Klassische Optimierung	12
1.4.1	Variationsungleichung und Unrestringierte Optimierung	12
1.4.2	Konvexe Optimierung	13
2	Konvexität und Trennbarkeit	17
2.1	Konvexe Mengen und Kegel	17
2.2	Konvexkombinationen	18
2.3	Trennbarkeit	21
3	Einführung in die Lineare Optimierung	27
3.1	Beispiele	27
3.1.1	Produktionsmodelle	27
3.1.2	Mischungsaufgaben	29
3.1.3	Transportprobleme	30
3.2	Mathematische Formulierung linearer Optimierungsprobleme	30
3.3	Rückführung anderer LP's auf Standardform	31
3.3.1	Ungleichungen	31
3.3.2	Freie Variable	32
4	Grundlagen der Polyedertheorie	35
4.1	Seitenflächen von Polyedern	35
4.2	Ecken und Extrempunkte	39
4.3	Basislösungen	46
4.4	Degeneriertheit von Basislösungen	48
5	Das Simplex Verfahren	51
5.1	Parametrisierung des Lösungsraumes	52
5.2	Basistausch	53
5.3	Zielfunktion und Reduzierte Kosten	59
5.4	Anwendung auf LPs mit Ungleichungen	63

6 | Inhaltsverzeichnis

5.5	Kreiseln des Simplex Verfahrens und Lexikographische Auswahlregel	65
5.6	Berechnung einer zulässigen Basislösung - Die Zweiphasen-Methode	67
6	Dualität für Lineare Probleme	75
6.1	Duale Programme	75
6.2	Duale Programme für Transportprobleme	78
6.3	Der Dualitätssatz	79
6.4	Das Duale Simplex-Verfahren	82
6.5	Sensitivität und Schattenpreise	86
7	Postoptimierung	89
7.1	Änderung der rechten Seite b	89
7.2	Änderung der Zielfunktion c	90
7.3	Hinzufügen einer neuen Variablen	90
7.4	Änderung einer Spalte von A	90
8	Parametrische Lineare Optimierung	95
8.1	Störung der Zielfunktion c	95
8.2	Störung der rechten Seite b	97
9	Komplexität der Linearen Optimierung	103
9.1	Klee-Minty Würfel	103
9.2	Polynomiale Algorithmen	108
9.3	Reduktion auf Zulässigkeitsprobleme – LP-Dualität	110
9.4	Reduktion mit Binärsuche	111
9.5	Geometrische Grundlagen von Ellipsoiden	115
9.6	Die Ellipsoid-Methode	120
9.7	Polynomialität der Ellipsoid-Methode	127
9.8	Nicht-Volldimensionale Polytope	128
9.9	Bestimmung eines exakten zulässigen Punktes	129
9.10	Implementation	129
10	Lineare Ganzzahlige Optimierung	131
10.1	Totale duale Ganzzahligkeit und ganzzahlige Programme	132
10.1.1	Total dual ganzzahlige Systeme	132
10.1.2	Ganzzahlige rechte Seiten und ganzzahlige Optima	133
10.1.3	Zusammenhang mit total unimodularen Matrizen	134
10.1.4	Anwendung: Bipartites Matching	134
10.2	Die Branch & Bound Methode	136
10.3	LP-Basierte Relaxierungen	137
10.4	Schnittebenenverfahren	139

Kapitel 1

Einführung

1.1 Motivation und Grundbegriffe

Allgemein betrachten wir einen reellen normierten Vektorraum $(V, \|\cdot\|)$ und $K \subset V$ nicht-leer. Ferner sei eine reellwertige Funktion

$$f : K \rightarrow \mathbb{R}$$

gegeben. Gesucht ist nun $\min\{f(x) \mid x \in K\}$. Häufig schreibt man auch

$$\begin{aligned} &\text{minimiere } f(x) \\ &\text{unter den Nebenbedingungen } x \in K. \end{aligned} \tag{1.1}$$

Eine optimale Lösung des Problems (1.1) wird auch globales Minimum genannt.

Definition 1.1. Ein Punkt $x^* \in K$ ist ein globales Minimum von f über K falls

$$f(x) \geq f(x^*) \text{ für alle } x \in K.$$

Falls

$$f(x) > f(x^*) \text{ für alle } x \in K, x \neq x^*,$$

so sprechen wir von einem strikten globalen Minimum. Globale Maxima sind analog definiert.

- infinite Optimierung (V ist unendlich dimensionaler Funktionenraum),
- finite Optimierung ($V = \mathbb{R}^n$),
- kontinuierliche Optimierung ($\text{int}(K) \neq \emptyset$),
- diskrete Optimierung ($K \subseteq \mathbb{Z}^n$).

1.2 Beispiele und Anwendungen

Beispiel 1.2 (Portfoliooptimierungsproblem nach Markowitz). Gegeben seien $j = 1, \dots, n$ mögliche Anlagen (z.B. Aktien, Fonds, Optionen, Wertpapiere). Jede Anlage wirft im nächsten Zeitintervall einen Gewinn (oder Verlust) R_j ab.

8 | Kapitel 1. Einführung

Ein mögliches Ziel ist es, den erwarteten Gewinn

$$E[R] = \sum_{j=1}^n x_j E[R_j], \quad R = \sum_{j=1}^n x_j R_j$$

zu maximieren (E bezeichnet den Erwartungswert).

Als Maß für das Risiko kann die Varianz des Gewinns dienen:

$$\text{Var}[R] = E[R - E[R]]^2 = E \left[\sum_{j=1}^n x_j (R_j - E[R_j]) \right]^2.$$

Ein Kompromiss zwischen hohem Gewinn und geringem Risiko kann durch Lösen des folgenden Optimierungsproblems erreicht werden:

$$\begin{aligned} \min & - \sum_{j=1}^n x_j E[R_j] + \alpha \cdot E \left[\sum_{j=1}^n x_j (R_j - E[R_j]) \right]^2 \\ \text{s.t. : } & \sum_{j=1}^n x_j = 1, x_j \geq 0, j = 1, \dots, n. \end{aligned}$$

Beispiel 1.3 (Optimaler Mengeneinkauf). Es soll eine bestimmte Menge M eines Gutes gekauft werden. Dazu werden Angebote von n Lieferanten eingeholt, wobei kein Lieferant die gewünschte Menge alleine liefern kann, sondern höchstens M_i Einheiten des Gutes. Die Preise $f_i(x_i)$ des i -ten Lieferanten sind i.a. abhängig von der Liefermenge x_i und beinhalten Mengenrabatte.

$$\begin{aligned} \min & \sum_{i=1}^n f_i(x_i) \\ \text{s.t. : } & \sum_{i=1}^n x_i = M \\ & 0 \leq x_i \leq M_i, i = 1, \dots, n. \end{aligned}$$

Beispiel 1.4 (Regression). Ein Experiment liefert die Messpunkte $(t_i, y_i), i = 1, \dots, m$. Der dem Experiment zu Grunde liegende Vorgang werde durch die Funktion $f(t, p)$ modelliert, die einen funktionalen Zusammenhang zwischen den Messstellen t_i und den Messwerten y_i herstellt. Daher wird versucht, die Messpunkte so gut wie möglich zu approximieren, indem

$$\sum_{i=1}^m (y_i - f(t_i, p))^2$$

bezüglich p minimiert wird. Ein allgemeines Least-Squares Problem lautet

$$\min \sum_{i=1}^q \Phi_i(x)^2$$

$$\begin{aligned} \text{s.t. : } g_j(x) &\leq 0, j = 1, \dots, r \\ h_j(x) &= 0, j = 1, \dots, s. \end{aligned}$$

Beispiel 1.5 (Optimalsteuerung). Gesucht sei eine Steuerfunktion, um ein Auto mit möglichst geringem (noch genauer zu definierendem) Aufwand in vorgegebener Zeit t_f von Punkt A (aus Ruhelage) nach Punkt B (in Ruhelage) zu steuern. Bekannt aus der Physik ist nach den Newtonschen Gesetzen, dass der Weg $s(t)$, die Geschwindigkeit $v(t)$ und die Beschleunigung $a(t)$ folgenden Gesetzmäßigkeiten unterliegen:

$$\dot{s}(t) = v(t), \dot{v}(t) = a(t).$$

Wir betrachten die Aufgabe eindimensional auf der Verbindungslinie von A nach B (der Länge d) im Zeitintervall $[0, t_f]$. Setzen wir die Koordinate der Punkt A bzw. B auf $s = 0$ bzw. $s = d$, so fordern wir gemäß der Aufgabenstellung $s(0) = 0, v(0) = 0, s(t_f) = d, v(t_f) = 0$. Der Gesamtaufwand wird über das Integral der quadrierten Steuerfunktion gemessen, d.h. wir möchten erreichen, dass

$$\int_0^{t_f} a(t)^2 dt$$

möglichst klein wird. Insgesamt lautet die Formulierung daher

$$\begin{aligned} \min \int_0^{t_f} a(t)^2 dt \\ \dot{s}(t) &= v(t) \\ \dot{v}(t) &= a(t) \\ s(0) &= 0, s(t_f) = d \\ v(0) &= 0, v(t_f) = 0. \end{aligned}$$

1.3 Finite Optimierung

In dieser Vorlesung werden wir nur finite bzw. endlich dimensionale Optimierungsprobleme betrachten und im Folgenden immer $V := \mathbb{R}^n$ annehmen.

Die Menge der Punkte aus K , in denen das Minimum angenommen wird, bezeichnen wir mit $\arg \min(f, K)$. Es gilt

$$\alpha = \min\{f(x) : x \in K\} \Leftrightarrow \arg \min(f, K) = \{x \in K \mid f(x) = \alpha\}.$$

Definition 1.6. Ein Punkt $x^* \in K$ ist ein lokales Minimum von f über K falls $\rho > 0$ existiert, so dass

$$f(x) \geq f(x^*) \text{ für alle } x \in K \cap B_\rho(x^*),$$

wobei

$$B_\rho(x) = \{y \in \mathbb{R}^n \mid \|x - y\| < \rho\}$$

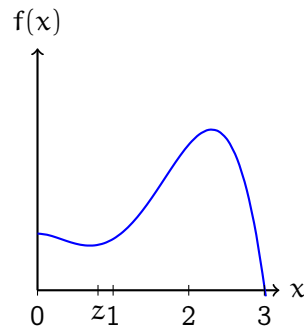


Abbildung 1.1: Die Funktion f nimmt auf $K = [0, 3]$ sein **globales** Minimum auf $x^* = 3$ an. Es gibt ein weiteres **lokales** Minimum im Punkt z .

die offene Einheitskugel mit Zentrum x und Radius $\rho > 0$ bezeichnet. Falls

$$f(x) > f(x^*) \text{ für alle } x \in K \cap B_\rho(x^*), x \neq x^*,$$

so sprechen wir von einem strikten lokalen Minimum.

In der Regel kann K durch funktionale Gleichungs- und Ungleichungsrestriktionen beschrieben werden. In diesem Fall kann K dargestellt werden als

$$K = \{x \in \mathbb{R}^n \mid h_i(x) = 0, \ i \in I_1 = \{1, \dots, m\}, g_j(x) \leq 0, \ j \in I_2 = \{1, \dots, p\}\}, \quad (1.2)$$

wobei angenommen wird dass alle auftretenden Funktionen hinreichend glatt sind, d.h., $f, h_i, g_j \in C^2$ für alle $i \in I_1, j \in I_2$.

Für obige Form von K erhalten wir folgende verfeinerten Problemstellungen (vgl. Oberle, Skript 2012, Einführung in die Optimierung):

- Unrestringierte Optimierung: $m = p = 0$
- Restringierte Optimierung: $m > 0$ oder $p > 0$
- Lineare Optimierung: f linear, g_j, h_i affin linear
- Quadratische Optimierung: $f(x) = x^T A x + b^T x + c$, g_j, h_i affin linear
- Konvexe Optimierung: f konvex, g_j konvex, h_i affin linear
- L_2 -Probleme: Die Zielfunktion f hat die spezielle Form

$$f(x) = \sum_{i=1}^n w_i (f_i(x))^2,$$

mit glatten Funktionen f_i und Gewichten $w_i > 0$.

- Minimax Probleme: Die Zielfunktion f hat die spezielle Form

$$f(x) = \max\{f_i(x), i = 1, \dots, m\}$$

mit glatten Funktionen f_i .

Die folgenden Fragen treten im Zusammenhang mit Optimierungsproblemen natürlicherweise auf:

- Existieren optimale Lösungen?
- Sind sie eindeutig?
- Lassen sich notwendige und hinreichende Optimalitätsbedingungen herleiten?
- Welche Algorithmen eignen sich zur Berechnung einer optimalen Lösung?
- Wie hängt eine optimale Lösung von den Problemparametern ab ?

Bezüglich der Existenz von Optimallösungen werden wir zunächst einen zentralen Satz der Analysis wiederholen.

Satz 1.7 (Weierstrass). Sei $K \subset \mathbb{R}^n$ nichtleer und kompakt und $f : K \rightarrow \mathbb{R}$ stetig. Dann existiert ein $x^* \in K$ mit

$$f(x^*) \leq f(x) \text{ für alle } x \in K.$$

Beweis. Da f stetig ist, ist das Bild $f(K)$ der kompakten Menge K beschränkt in $\mathbb{R}$ und das Infimum

$$A := \inf\{f(x) | x \in K\} \in \mathbb{R}$$

existiert (siehe den Satz über die Existenz von Infima/Suprema für nichtleere und beschränkte Teilmengen in $\mathbb{R}$). Weiterhin existiert eine Folge $x_n \in K, n \in \mathbb{N}$ so dass gilt

$$\lim_{n \rightarrow \infty} f(x_n) = A.$$

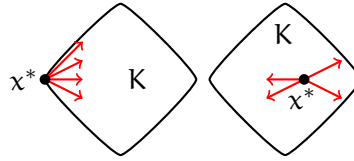
(Die Existenz einer solchen Folge sollte auch klar sein, man nehme z.B. Folgeglieder $x_n \in K$ mit $f(x_n) < A + 1/n$). Da die Folge $x_n, n \in \mathbb{N}$ beschränkt ist, besitzt sie nach dem Satz von Bolzano/Weierstrass eine konvergente Teilfolge $x_{n_k}, k \in \mathbb{N}$ mit

$$\lim_{k \rightarrow \infty} x_{n_k} =: x^* \in K.$$

Aus der Stetigkeit von f folgt

$$f(x^*) = \lim_{k \rightarrow \infty} f(x_{n_k}) = A.$$

Somit nimmt f auf K ihr Minimum bei x^* an. □

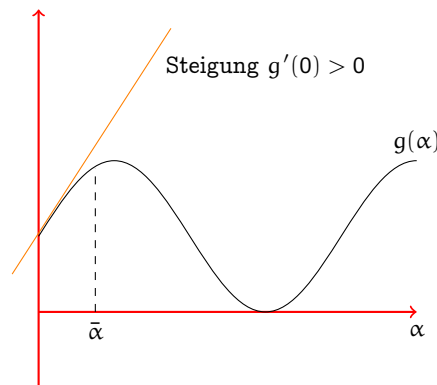
Abbildung 1.2: Die roten Pfeile repräsentieren zulässige Richtungen in $D_K(x^*)$.

1.4 Differenzierbare Klassische Optimierung

1.4.1 Variationsungleichung und Unrestringierte Optimierung

Definition 1.8 (Zulässigen Richtungen). Sei $x \in K \subseteq \mathbb{R}^n$ mit $K \neq \emptyset$. Der Vektor $d \in \mathbb{R}^n$ ist eine zulässige Richtung an x , falls ein $\bar{\alpha} > 0$ existiert, sodass $x + \alpha d \in K$ für alle $0 \leq \alpha \leq \bar{\alpha}$.

Mit $D_K(x)$ bezeichnen wir die Menge der zulässigen Richtungen an x . Man sieht leicht ein, dass $D_K(x)$ ein Kegel mit Spitze im Nullpunkt (vgl. 2.1) ist, daher wird $D_K(x)$ auch der Kegel der zulässigen Richtungen genannt.



Für stetig differenzierbare Optimierungsprobleme der Form (1.1) erhalten wir folgende notwendige Optimalitätsbedingung.

Satz 1.9 (Variationsungleichung). Sei $K \subseteq \mathbb{R}^n$ und $f : \mathbb{R}^n \rightarrow \mathbb{R}$ stetig differenzierbar. Sei x^* ein lokales Minimum von f über K und d eine zulässige Richtung an x^* . Dann gilt

$$\nabla f(x^*)^\top d \geq 0.$$

Beweis. Nach Voraussetzung existiert eine zulässige Richtung d an x^* . Somit existiert $\bar{\alpha} > 0$ sodass $x^*(\alpha) := x^* + \alpha d \in K$ für alle $0 \leq \alpha \leq \bar{\alpha}$. Wir definieren eine eindimensionale Funktion $g(\alpha) := f(x^*(\alpha))$. Für ein lokales Minimum x^* (bzgl. Umgebung $B_\rho(x^*)$) gilt

$$g(\alpha) \geq g(0) \text{ für alle } \alpha \in [0, \min\{\bar{\rho}, \bar{\alpha}\}],$$

wobei

$$\bar{\rho} := \sup\{\alpha \geq 0 \mid x^* + \alpha d \in B_{\rho/2}(x^*)\}.$$

Somit folgt

$$\lim_{\alpha \rightarrow +0} \frac{g(\alpha) - g(0)}{\alpha} \geq 0.$$

Mit der Differenzierbarkeit von f gilt nun

$$0 \leq \lim_{\alpha \rightarrow +0} \frac{g(\alpha) - g(0)}{\alpha} = g'(0) = \nabla f(x^*)^\top d.$$

□

Satz 1.10. Sei $K \subseteq \mathbb{R}^n$ und $f : K \rightarrow \mathbb{R}$ stetig differenzierbar. Sei $x^* \in \text{int}(K)$ ein lokales Minimum von f über K . Dann gilt

$$\nabla f(x^*) = 0. \quad (1.3)$$

Insbesondere gilt (1.3) für jedes lokale Minimum eines unrestringierten Optimierungsproblems.

Beweis. Mit der Variationsungleichung aus Satz 1.9 gilt für jede zulässige Richtung $d \in \mathbb{R}^n$ die Ungleichung $\nabla f(x^*)^\top d \geq 0$. Da $x^* \in \text{int}(K)$ ist für jeden Vektor $d \in \mathbb{R}^n$ auch $-d$ zulässig. Daraus folgt die Behauptung. □

Bemerkung 1.11. Es sei darauf verwiesen dass das Konzept von zulässigen Richtungen wie in Definition 1.8 nicht geeignet ist um zulässige Richtungen für gekrümmte Hyperflächen (oder nicht-volldimensionale Mengen) darzustellen. Für solche Fälle benötigt man anschaulich ein Konzept welches **infinitesimale** Schrittweiten in eine Richtung erlaubt. Diese Idee führt auf Begriffe wie **Tangentialkegel** und **Linearisierungskegel**, die wir uns für Optimierung II aufheben.

1.4.2 Konvexe Optimierung

Wir betrachten nun die Minimierung einer konvexen differenzierbaren Funktion f auf einer konvexen Menge $K \subset \mathbb{R}^n$.

Definition 1.12. Eine Menge $K \subset \mathbb{R}^n$ heißt konvex, wenn für alle $x, y \in K$ die Verbindungsstrecke zwischen x und y in K liegt, d.h.

$$\lambda x + (1 - \lambda)y \in K \text{ für alle } \lambda \in [0, 1].$$

Definition 1.13. Sei $K \subset \mathbb{R}^n$ konvex. Eine Funktion $f : K \rightarrow \mathbb{R}$ heißt konvex, wenn für alle $x, y \in K$ und $\lambda \in [0, 1]$ gilt:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y). \quad (1.4)$$

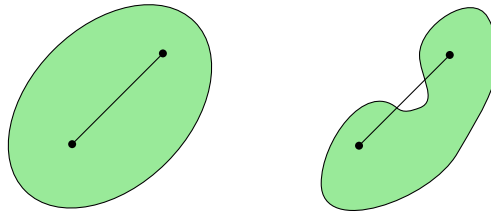


Abbildung 1.3: Links: Konvexe Menge. Rechts: Nicht-konvexe Menge

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

Satz 1.14. Sei $K \subset \mathbb{R}^n$ konvex, und seien $f_1, f_2 : K \rightarrow \mathbb{R}$ konvexe Funktionen und sei $\alpha > 0$. Dann sind auch αf_1 , $f_1 + f_2$ und $\max\{f_1, f_2\}$ konvex auf K .

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 1.15. Sei $K \subset \mathbb{R}^n$ konvex, und $f : K \rightarrow \mathbb{R}$. Die Menge

$$\text{Epi}(f) = \{(x, \alpha) \in K \times \mathbb{R} : f(x) \leq \alpha\}$$

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

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

untere Niveaumenge von f zum Niveau β .

Wir erhalten folgende Charakterisierung.

Satz 1.16. Sei $K \subseteq \mathbb{R}^n$ eine beliebige Teilmenge und $f : K \rightarrow \mathbb{R}$. Dann gilt:

1. f ist konvex $\Leftrightarrow \text{Epi}(f)$ ist konvex.
2. f ist konvex $\Rightarrow L(f, \beta)$ ist konvex für jedes $\beta \in \mathbb{R}$. Die Umkehrung gilt i.A. nicht.

Beweis. Übung. □

Für konvexe differenzierbare Funktionen erhalten wir folgende Charakterisierung.

Satz 1.17. Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ und $f \in C^1$. Dann gilt:

1. f ist genau dann konvex auf der konvexen Menge $K \subseteq \mathbb{R}^n$, wenn für alle $x, y \in K$ gilt:

$$f(y) \geq f(x) + \nabla f(x)^T(y - x). \quad (1.5)$$

2. f ist strikt konvex $\Rightarrow$ (1.5) gilt strikt für alle $x \neq y \in K$.

Beweis. Wir zeigen zunächst die Richtung $\Leftarrow$ für die erste Aussage. Es gelte (1.5) für alle $x, y \in K$. Wähle zwei beliebige Punkte $x, y \in K$ und $\lambda \in (0, 1)$. Wegen der Konvexität von K ist dann auch

$$z = \lambda x + (1 - \lambda)y \in K. \quad (1.6)$$

Mit (1.5) gilt für $x, y, z \in K$

$$f(x) \geq f(z) + (x - z)^T \nabla f(z) \quad (1.7)$$

$$f(y) \geq f(z) + (y - z)^T \nabla f(z). \quad (1.8)$$

Multiplizieren wir nun (1.7) mit λ und (1.8) mit $(1 - \lambda)$, und addieren die beiden Ungleichungen, so erhalten wir

$$\begin{aligned} \lambda f(x) + (1 - \lambda)f(y) &\geq f(z) + \left((\lambda(x - z) + (1 - \lambda)(y - z)) \right)^T \nabla f(z) \\ &= f(z) + \left(\lambda x + (1 - \lambda)y - z \right)^T \nabla f(z) \\ &= f(z). \end{aligned}$$

Wegen (1.6) ist der zweite Ausdruck auf der rechten Seite der zweiten Ungleichung 0. Somit folgt die Konvexität von f .

$\Rightarrow$: Sei f konvex. Wir wählen $x, y \in K$ und definieren die Hilfsfunktion $\psi : \mathbb{R} \rightarrow \mathbb{R}$ als

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

Mittels Konvexität von f gilt für alle $\lambda \in [0, 1]$ die Eigenschaft $\psi(\lambda) \geq 0$. Desweiteren gilt $\psi(0) = 0$. Wir berechnen nun die Ableitung von ψ an der Stelle 0 und erhalten

$$0 \leq \lim_{t \rightarrow 0+} \frac{\psi(t) - \psi(0)}{t} = \dot{\psi}(0) = -f(x) + f(y) - \nabla f(x)^T(y - x).$$

Die zweite Aussage folgt unmittelbar. □

Wir erhalten nun sogar ein hinreichendes Optimalitätskriterium für konvexe Optimierungsprobleme.

Satz 1.18. Sei $K \subset \mathbb{R}^n$ konvex und sei $f : K \rightarrow \mathbb{R}$ eine differenzierbare konvexe Funktion. Dann ist jedes lokale Minimum von f auch ein globales Minimum.

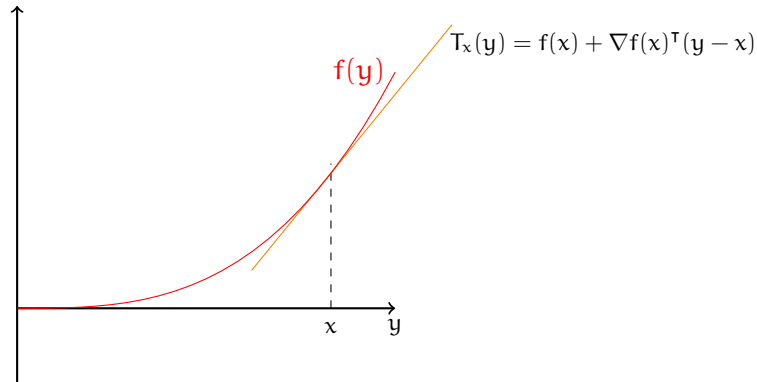


Abbildung 1.4: Illustration von Ungleichung 1.5. $T_x(y)$ repräsentiert die Tangentialebene von f im Punkt x und es gilt $T_x(y) \leq f(y)$ für alle $y \in K$.

Beweis. Sei x^* ein lokales Minimum. Dann gilt mit Satz 1.9, dass für jede zulässige Richtung $d \in \mathbb{R}^n$ an x^* gelten muss $\nabla f(x^*)^T d \geq 0$. Weil K konvex ist, gilt für beliebiges $y \in K$, dass $x^* + \lambda(y - x^*) = \lambda y + (1 - \lambda)x^* \in K$ für alle $\lambda \in [0, 1]$. Somit ist $y - x^*$ eine zulässige Richtung an x^* . Wir erhalten

$$f(y) \geq f(x^*) + \nabla f(x^*)^T(y - x^*) \geq f(x^*),$$

wobei die erste Ungleichung aus Satz 1.17 folgt und die zweite aus der Variationsungleichung. $\square$

Es folgt unmittelbar folgendes Optimalitätskriterium für unrestringierte konvexe Probleme.

Korollar 1.19. Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine differenzierbare konvexe Funktion. Dann ist jeder Punkt $x^* \in \mathbb{R}^n$ mit $\nabla f(x^*) = 0$ ein globales Minimum des zugehörigen unrestringierten Optimierungsproblems.

Kapitel 2

Konvexität und Trennbarkeit

2.1 Konvexe Mengen und Kegel

Definition 2.1. 1. Für $M \subset \mathbb{R}^n$ heißt

$$\text{co}(M) := \cap \{K \supset M \mid K \text{ konvex}\}$$

die konvexe Hülle von M . Wie wir zeigen werden ist die konvexe Hülle die kleinste konvexe Menge, welche M enthält. Für $x^0, \dots, x^k \in \mathbb{R}^n$ heißt

$$\text{co}(x^0, \dots, x^k) = \text{co}(\{x^0, \dots, x^k\})$$

das von $x^0, \dots, x^k$ aufgespannte Simplex. Wenn $x^1 - x^0, \dots, x^k - x^0$ linear unabhängig sind, so heißt das Simplex nichtentartet.

2. Eine Teilmenge $K \subset \mathbb{R}^n$ heißt Kegel (mit Spitze im Nullpunkt), wenn für alle $x \in K$ der Halbstrahl durch x in K liegt, d.h.

$$\alpha x \in K \text{ für alle } \alpha \geq 0.$$

Ein Kegel K , der konvex ist, heißt konvexer Kegel.

3. Sei $K \subset \mathbb{R}^n$ und $x \in K$. Der Kegel

$$K(x) := \{\alpha(y - x) \mid y \in K, \alpha > 0\} = \bigcup_{\alpha > 0} \alpha(K - x)$$

heißt die konische Hülle von K bzgl. x .

Zur Erläuterung dieser Definitionen dienen die folgenden Beispiele in Abbildung 2.1. Wir definieren nun wichtige Klassen von konvexen Mengen.

Definition 2.2. Sei $a \in \mathbb{R}^n \setminus \{0\}$ und $b \in \mathbb{R}$.

1. Eine Menge $H = \{x \in \mathbb{R}^n : a^\top x = b\}$ wird Hyperebene genannt.
2. Die Mengen $H^- = \{x \in \mathbb{R}^n : a^\top x \leq b\}$ und $H^+ = \{x \in \mathbb{R}^n : a^\top x \geq b\}$ werden Halbräume genannt.

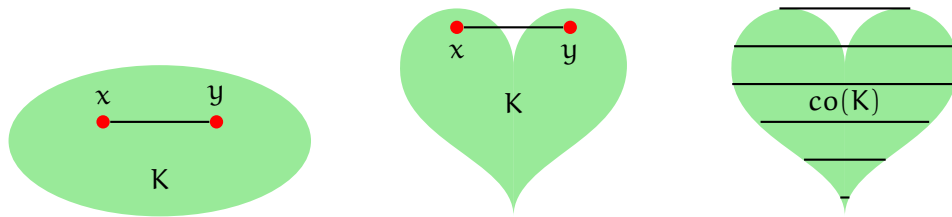


Abbildung 2.1: Die erste Menge ist konvex. Das zweite Herz ist nichtkonvex und die dritte Menge repräsentiert die konvexe Hülle des nicht-konvexen Herzens.

3. Sei nun A eine reellwertige $m \times n$ Matrix und $b \in \mathbb{R}^m$.

$K = \{x \in \mathbb{R}^n | Ax \leq b\}$ wird Polyeder und $K = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$ Polyeder in Standardform genannt.

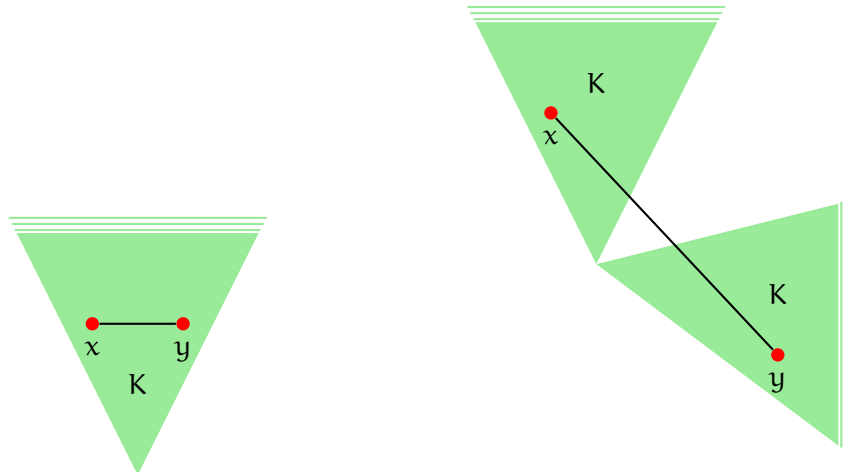


Abbildung 2.2: Links: konvexer Kegel. Rechts: nicht-konvexer Kegel.

2.2 Konvexkombinationen

Definition 2.3. Seien $x^1, \dots, x^k \in \mathbb{R}^n$ und $\lambda_1, \dots, \lambda_k \in \mathbb{R}_{\geq 0}$ mit $\lambda_1 + \dots + \lambda_k = 1$.

- Der Vektor $\sum_{i=1}^k \lambda_i x^i$ heißt Konvexkombination von $x^1, \dots, x^k$.

Wir erhalten folgenden Satz.

- Satz 2.4.**
1. Der Schnitt von konvexen Mengen ist konvex.
 2. Jedes Polyeder ist eine konvexe Menge.
 3. Eine Konvexkombination einer endlichen Anzahl von Punkten einer konvexen Menge liegt in derselben Menge.

4. Die konvexe Hülle einer Menge $K \subset \mathbb{R}^n$ ist die Menge der Konvexkombinationen von Punkten in K . Insbesondere ist die konvexe Hülle endlich vieler Punkte konvex.

Beweis. (1): Seien $X_i, i \in I$ konvexe Mengen und definiere $X := \bigcap_{i \in I} X_i$. Für $x, y \in X$ gilt $x, y \in X_i$ für alle $i \in I$, somit folgt $\lambda x + (1 - \lambda)y \in X_i$ für alle $i \in I$. Schlussendlich folgt $\lambda x + (1 - \lambda)y \in X$.

(2): Ein Polyeder ist der Schnitt von endlich vielen (konvexen) Halbräumen.

(3): Wir zeigen mit Induktion über k , dass jede Konvexkombination von k vielen Punkten aus X wieder in X liegt. Für $k = 1$ ist die Aussage trivial und für $k = 2$ folgt die Aussage aus der Konvexität von X . Für den Schritt $k - 1 \rightarrow k$ betrachte eine Konvexkombination $\mu_1 x^1 + \dots + \mu_k x^k$. Falls $\mu_i = 0$ für ein $i \in \{1, \dots, k\}$ können wir die Induktionsvoraussetzung anwenden, daher nehmen wir nun o.B.d.A. an dass $\mu_k \in (0, 1)$ gilt. Definiere

$$v_1 = \frac{\mu_1}{1 - \mu_k}, \dots, v_{k-1} = \frac{\mu_{k-1}}{1 - \mu_k} \geq 0, \sum_{l=1}^{k-1} v_l = 1.$$

Desweiteren setze

$$y := \sum_{l=1}^{k-1} v_l x^l \in X$$

nach Induktionsvoraussetzung. Nun gilt

$$x = (1 - \mu_k)y + \mu_k x^k \in X,$$

weil der Fall $k = 2$ schon gezeigt wurde.

(4): Sei L die Menge der Konvexkombinationen von Punkten in K und $\text{co}(K)$ die konvexe Hülle von K .

$L \subseteq \text{co}(K)$: Da $\text{co}(K)$ nach (1) konvex ist, liegen alle Konvexkombinationen von Punkten aus $\text{co}(K)$ wieder in $\text{co}(K)$. Da zusätzlich nach Definition der konvexen Hülle $K \subseteq \text{co}(K)$ gilt, ist $L \subseteq \text{co}(K)$.

$\text{co}(K) \subseteq L$: Seien nun $x, y \in L$ mit

$$x = \sum_{i=1}^k \alpha_i x^i \text{ und } y = \sum_{j=1}^l \beta_j y^j, \text{ wobei } \alpha_i, \beta_j \geq 0, \sum_{i=1}^k \alpha_i = 1, \sum_{j=1}^l \beta_j = 1.$$

Für ein beliebiges $\lambda \in [0, 1]$ gilt dann für

$$z = \lambda x + (1 - \lambda)y = \sum_{i=1}^k \lambda \alpha_i x^i + \sum_{j=1}^l (1 - \lambda) \beta_j y^j,$$

dass dieses Element auch in L liegt, denn

$$0 \leq \lambda \alpha_i \leq 1 \forall i, \quad 0 \leq (1 - \lambda) \beta_j \leq 1 \forall j \text{ und}$$

$$\sum_{i=1}^k \lambda \alpha_i + \sum_{j=1}^l (1-\lambda) \beta_j = \lambda \sum_{i=1}^k \alpha_i + (1-\lambda) \sum_{j=1}^l \beta_j = \lambda + 1 - \lambda = 1,$$

was nichts anderes bedeutet, dass z eine Konvexkombination von Elementen aus K ist. Damit ist L konvex. Offensichtlich ist auch $K \subseteq L$, da sich jedes Element $x^p \in K$ schreiben lässt als

$$x^p = \sum_{j \in J} \lambda_j x^j \text{ mit } \lambda_p = 1 \text{ und } \lambda_j = 0 \text{ für } j \neq p.$$

Nach der Definition der konvexen Hülle von K ist $\text{co}(K) \subseteq L$, da L nach obiger Argumentation eine der K enthaltenden konvexen Mengen ist, über die der Schnitt gebildet wird. □

Wir erhalten als Folgerung.

Korollar 2.5. Die Menge der Konvexkombinationen von $x^1, \dots, x^k \in \mathbb{R}^n$ ist die kleinste (bzgl. Inklusion) konvexe Teilmenge von $\mathbb{R}^n$, welche die Punkte $x^1, \dots, x^k$ enthält.

Beweis. Sei X die Menge der Konvexkombinationen von $x^1, \dots, x^k \in \mathbb{R}^n$. Definiere nun

$$Y := \text{co}(x^1, \dots, x^k) = \bigcap_{\substack{C \subseteq \mathbb{R}^n \\ C \text{ konvex} \\ C \text{ enthält } x^1, \dots, x^k}} C. \quad (2.1)$$

Y ist wohldefiniert weil $\mathbb{R}^n$ ein Kandidat der C Mengen ist. Nach Satz 2.4(1.) ist Y konvex als Schnitt konvexer Mengen. Y ist auch die kleinste konvexe Menge, die die Punkte $x^1, \dots, x^k$ enthält. Da X nach Satz 2.4(4.) konvex ist, erhalten wir sofort $Y \subseteq X$. Sei $x \in X$. Nach Definition von X gilt $x = \sum_{i=1}^k \lambda_i x^i$. Weil nach Voraussetzung $x^1, \dots, x^k \in Y$ gilt, folgt mit Satz 2.4(3.), dass $x \in Y$. □

Der folgende Satz sagt, dass man für die Darstellung eines beliebigen Punktes in $\text{co}(K)$ maximal $(n+1)$ Punkte benötigt, wobei n die Dimension des Raumes ist.

Satz 2.6 (Carathéodory). Für $K \subset \mathbb{R}^n$ ist $\text{co}(K)$ die Menge aller Konvexkombinationen von höchstens $(n+1)$ -elementigen Teilmengen von K .

Beweis. Sei $x \in \text{co}(K)$. Wegen Satz 2.4 (4) existieren $x^1, \dots, x^k \in K$ so dass

$$x = \sum_{i=1}^k \lambda_i x^i \text{ mit } \lambda_i \geq 0 \text{ für alle } i = 1, \dots, k \text{ und } \sum_{i=1}^k \lambda_i = 1.$$

Wenn bereits $k \leq n+1$ gilt, dann ist nichts mehr zu zeigen. Gilt $k > n+1$, dann zeigen wir, dass man für die Darstellung von x auf einen der k Punkte verzichten kann: Betrachten wir dazu die $(k-1)$ Vektoren $y^i = x^i - x^k, i = 1, \dots, k-1$. Für $k > n+1$ sind

die Punkte y^i linear abhängig, d.h. es existieren $\alpha_1, \dots, \alpha_{k-1}$ mit $\alpha_j \neq 0$ für mindestens ein $j \in \{1, \dots, k-1\}$, so dass

$$\begin{aligned} \sum_{i=1}^{k-1} \alpha_i y^i &= 0 \\ \Leftrightarrow \sum_{i=1}^{k-1} \alpha_i (x^i - x^k) &= 0 \\ \Leftrightarrow \sum_{i=1}^{k-1} \alpha_i x^i + \left(-\sum_{i=1}^{k-1} \alpha_i\right) x^k &= 0. \end{aligned}$$

Mit der Definition $\alpha_k = -\sum_{i=1}^{k-1} \alpha_i$ erhalten wir

$$\sum_{i=1}^k \alpha_i x^i = 0 \text{ und } \sum_{i=1}^k \alpha_i = 0.$$

Wegen $\alpha_j \neq 0$ für mindestens ein $j \in \{1, \dots, k-1\}$ ist folgender Wert wohldefiniert:

$$i_0 = \arg \min_{i \in \{1, \dots, k\}} \left\{ \frac{\lambda_i}{\alpha_i} \mid \alpha_i > 0 \right\} = \frac{\lambda_{i_0}}{\alpha_{i_0}}.$$

Es gilt nun

$$\lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}} \alpha_i \geq 0 \quad \forall i \quad \text{und} \quad \sum_{i=1}^k \lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}} \alpha_i = 1.$$

Desweiteren

$$x = \sum_{i=1}^k \lambda_i x^i = \sum_{i=1}^k \left(\lambda_i x^i - \frac{\lambda_{i_0}}{\alpha_{i_0}} \alpha_i x^i \right) = \sum_{i=1}^k \left(\lambda_i - \alpha_i \frac{\lambda_{i_0}}{\alpha_{i_0}} \right) x^i.$$

Dabei ist $\lambda_{i_0} - \alpha_{i_0} \frac{\lambda_{i_0}}{\alpha_{i_0}} = 0$, so dass x wie gefordert als Konvexkombination von nur $k-1$ Punkten dargestellt werden kann. $\square$

2.3 Trennbarkeit

Wir werden nun den anschaulich einleuchtenden Sachverhalt zeigen, dass man zwei disjunkte konvexe Mengen durch eine Hyperebene trennen kann, d.h., die Hyperebene unterteilt den Raum in zwei Halbräume, und beide Mengen liegen im jeweils anderen Halbraum.

Definition 2.7. Zwei Teilmengen $K_1, K_2 \subset \mathbb{R}^n$ heißen trennbar, wenn es $c \in \mathbb{R}$ und einen Zeilenvektor $\lambda \in \mathbb{R}^n, \lambda \neq 0$, gibt mit

$$\lambda x \leq c \leq \lambda y \quad \text{für alle } x \in K_1, y \in K_2.$$

Die Hyperebene $H = \{x \in \mathbb{R}^n \mid \lambda x = c\}$ heißt trennende Hyperebene (vgl. Abbildung 2.3); Die Mengen K_1, K_2 heißen echt trennbar durch die Hyperebene H , wenn $K_1 \cup K_2$ nicht

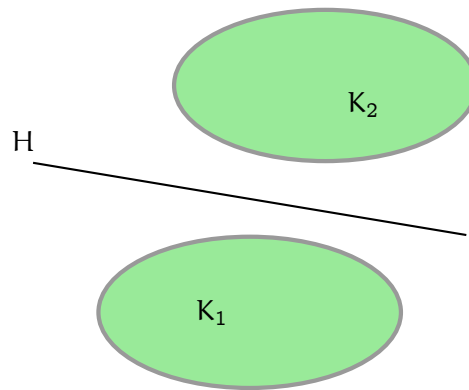


Abbildung 2.3: Illustration der (echten) Trennbarkeit von zwei disjunkten konvexen Mengen.

in H enthalten ist. Die Hyperebene H definiert die beiden Halbräume

$$H^+ = \{x \in \mathbb{R}^n | \lambda x \geq c\}, \quad H^- = \{x \in \mathbb{R}^n | \lambda x \leq c\}.$$

Die Mengen K_1, K_2 sind also trennbar, wenn entweder $K_1 \subseteq H^+, K_2 \subseteq H^-$ oder $K_1 \subseteq H^-, K_2 \subseteq H^+$.

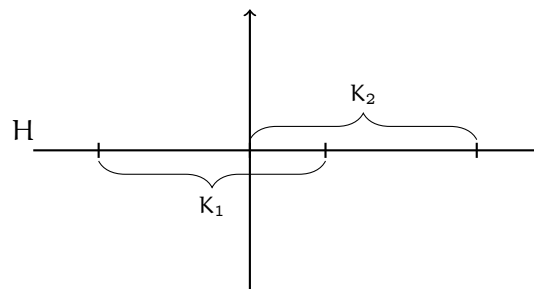


Abbildung 2.4: Zwei Mengen $K_1, K_2 \subset \mathbb{R}^2$ die trennbar, jedoch nicht echt trennbar sind. Die trennende Hyperebene ist gegeben durch $H = \{x \in \mathbb{R}^2 | x_2 = 0\}$.

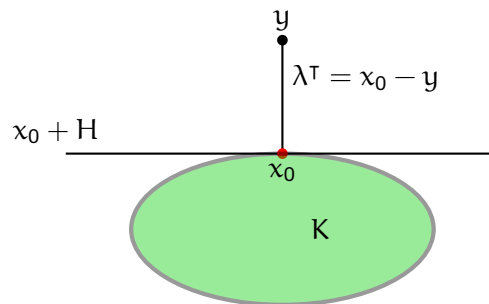


Abbildung 2.5: Illustration des Beweises von Satz 2.8. Gilt im Beispiel $x_0 = y$ dann ist $H + y$ eine trennende Stützhyperebene.

Satz 2.8. Sei $K \subset \mathbb{R}^n$ eine nichtleere, konvexe Menge und sei $y \notin \overset{\circ}{K} := \text{int}(K)$. Dann sind $\{y\}$ und K trennbar, d.h. es gibt einen Zeilenvektor $\lambda \in \mathbb{R}^n \setminus \{0\}$ mit

$$\lambda y \leq \lambda x \text{ für alle } x \in K.$$

Ist $\overset{\circ}{K} \neq \emptyset$ so sind $\{y\}$ und $\overset{\circ}{K}$ echt trennbar, und es gilt

$$\lambda y < \lambda x \text{ für alle } x \in \overset{\circ}{K}.$$

Beweis. 1. Fall: $y \notin \bar{K}$, wobei $\bar{K}$ den topologischen Abschluss von K bezeichnet. Mit $\|x\|$ werde die euklidische Norm bezeichnet. Setze

$$d := \inf_{x \in \bar{K}} \|x - y\| > 0.$$

Die Funktion $f(x) := \|y - x\|$ ist stetig und nimmt auf der kompakten Menge $\bar{K} \cap \{x \in \mathbb{R}^n \mid \|x - y\| \leq 2d\}$ ihr Minimum an (Satz von Weierstrass). Wegen der Abgeschlossenheit von $\bar{K}$ gibt es daher $x_0 \in \bar{K}$ mit $d = \|y - x_0\|$. Mit der Konvexität von $\bar{K}$ kann man ausserdem zeigen, dass der Lotpunkt x_0 eindeutig bestimmt ist (vgl. Abbildung 2.5).

Wir zeigen nun, dass $\lambda := (x_0 - y)^T \neq 0$ die Behauptung des Satzes erfüllt. Sei $x \in K$. Wegen der Konvexität von $\bar{K}$ gilt

$$x_0 + \alpha(x - x_0) \in \bar{K} \text{ für } 0 \leq \alpha \leq 1.$$

Daher ist

$$\|x_0 + \alpha(x - x_0) - y\|^2 \geq \|x_0 - y\|^2$$

und hieraus erhält man durch Ausrechnen der Skalarprodukte

$$2\alpha(x_0 - y)^T(x - x_0) + \alpha^2 \|x - x_0\|^2 \geq 0.$$

Die Division durch $\alpha > 0$ liefert im Limes $\alpha \downarrow 0$ die Beziehung

$$(x_0 - y)^T(x - x_0) \geq 0$$

und damit wegen $\lambda^T = (x_0 - y)$ und $d = \|\lambda\|$

$$\lambda x \geq \lambda x_0 = \lambda y + d^2 > \lambda y.$$

Also ist $H := \{x \in \mathbb{R}^n \mid \lambda x = \lambda x_0\}$ eine Hyperebene, die $\{y\}$ und K trennt.

2. Fall: $y \in \partial K = \bar{K} - \overset{\circ}{K}$.

Zu $y \in \partial K$ gibt es eine Folge $\{y_k\}$, $y_k \notin \bar{K}$, mit $y = \lim_{k \rightarrow \infty} y_k$. Zu y_k kann gemäß dem ersten Fall ein Zeilenvektor $\lambda_k \neq 0$ gewählt werden mit

$$\lambda_k y_k \leq \lambda_k x \text{ für alle } x \in K.$$

24 | Kapitel 2. Konvexität und Trennbarkeit

O.E. kann $\|\lambda_k\| = 1$ gesetzt werden und daher können wir annehmen, dass die beschränkte Folge $\{\lambda_k\}$ konvergent ist mit $\lambda = \lim_{k \rightarrow \infty} \lambda_k$, $\|\lambda\| = 1$ (mit dem Satz von Bolzano-Weierstraß folgt dass jede beschränkte Folge eine konvergente Teilfolge besitzt). Durch Grenzübergang in der obigen Ungleichung ergibt sich dann

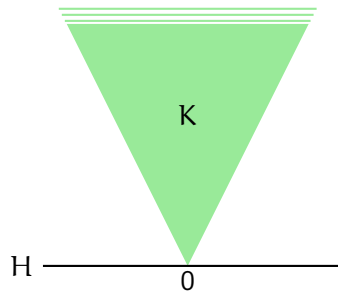
$$\lambda y \leq \lambda x \text{ für alle } x \in K.$$

Die Aussage

$$\lambda y < \lambda x \text{ für alle } x \in \overset{\circ}{K}$$

folgt unmittelbar. □

Im Fall $y \in \partial K$ heißt die trennende Hyperebene Stützhyperebene von K in y . Aus dem Trennungssatz 2.8 folgert man direkt:



Abbildungung 2.6: Trennung eines konvexen Kegels mittels Hyperebene durch 0.

Satz 2.9. Sei $K \subset \mathbb{R}^n$ ein nichtleerer konvexer, abgeschlossener Kegel und sei $y \notin K$. Dann gibt es einen Zeilenvektor $\lambda \in \mathbb{R}^n \setminus \{0\}$ mit

$$\lambda y < 0 \leq \lambda x \text{ für alle } x \in K.$$

Beweis. Wegen $\bar{K} = K$ gibt es nach dem ersten Fall von Satz 2.8 einen Zeilenvektor $\lambda \in \mathbb{R}^n \setminus \{0\}$ so dass für alle $x \in K$ gilt $\lambda y < \lambda x$. Aus $0 \in K$ folgt $\lambda y < \lambda 0 = 0$.

Angenommen es existiere ein $x \in K$ mit $\lambda x < 0$. Dann folgt

$$\lambda(\alpha x) = \alpha(\lambda x) \xrightarrow{\alpha \rightarrow \infty} -\infty,$$

im Widerspruch zu Beschränktheit von λK durch λy nach unten. Also gilt

$$\forall x \in K : \lambda y < 0 \leq \lambda x.$$

□

Hieraus gewinnt man den folgenden Alternativsatz:

Satz 2.10 (Lemma von Farkas). Sei B eine $k \times n$ Matrix und $d \in \mathbb{R}^k$. Dann gilt genau eine der folgenden Aussagen:

1. $Bx = d, x \geq 0$ hat eine Lösung $x \in \mathbb{R}^n$.
2. $\lambda B \geq 0, \lambda d < 0$ hat eine Lösung $\lambda \in \mathbb{R}^k$.

Beweis. Der Kegel

$$K := \{Bx \mid x \geq 0\} \subset \mathbb{R}^k$$

ist nichtleer, konvex und abgeschlossen. Es gilt genau eine der beiden Aussagen

(a) $d \in K$.

(b) $d \notin K$.

Die Aussage (a) ist genau die Aussage (1) des Satzes. Im Falle (b) gibt es nach dem Trennungssatz 2.8 ein $\lambda \in \mathbb{R}^k$ mit

$$\lambda d < 0 \leq \lambda z \text{ für alle } z \in K.$$

Also $\lambda Bx \geq 0$ für alle $x \geq 0$, d.h., $\lambda B \geq 0$. Dies ist die Aussage (2) des Satzes. Man beachte dass die jeweiligen Aussagen (a) und (1) bzw. (b) und (2) äquivalent sind und sich (1) und (2) somit ausschließen. $\square$

Kapitel 3

Einführung in die Lineare Optimierung

3.1 Beispiele

3.1.1 Produktionsmodelle

Ein Unternehmer produziert n Produkte $P_1, \dots, P_n$, zu deren Herstellung m Aktivitäten $A_1, \dots, A_m$ (Rohstoffe, Arbeitskräfte, Arbeitsstunden o.ä.) benötigt werden. Das Produkt P_j enthalte a_{ij} Anteile der Aktivität A_i und möge beim Verkauf pro Einheit einen Reingewinn von c_j Zahlungseinheiten erzielen. Von der Aktivität A_i sei die Menge b_i verfügbar. Die Produktionsmenge x_j des Produktes P_j soll nun bestimmt werden, so dass der Gewinn maximal wird. Gesucht ist daher das Maximum von

$$z(x) = \sum_{j=1}^n c_j x_j$$

unter den Nebenbedingungen

$$\sum_{j=1}^n a_{ij} x_j \leq b_i, \text{ für alle } i = 1, \dots, m$$
$$x_j \geq 0, \text{ für alle } j = 1, \dots, n.$$

Beispiel 3.1. Der Besitzer einer kleinen Schuhfabrik will je ein Modell eines Damen- und Herrenschuhs herstellen. Er verfügt über 40 Angestellte und über einen Maschinenpark von 10 Maschinen. Die pro Monat zur Verfügung stehende Arbeitszeit und Ledermenge ist in der folgenden Tabelle 3.1 enthalten: Mit den Optimierungsvariablen

x_1 : Zahl der produzierten Damenschuhe

x_2 : Zahl der produzierten Herrenschuhe

lautet die Optimierungsaufgabe:

Tabelle 3.1: Angaben zum Produktionsproblem

	Damenschuh	Herrenschuh	verfügbar
Herstellungszeit [h]	20	10	8000
Maschinenbearbeitung [h]	4	5	2000
Lederbedarf [dm ²]	6	15	4500
Reingewinn [EUR]	16	32	—

$$\text{maximiere } z(x_1, x_2) = 16x_1 + 32x_2$$

unter d.N.

$$20x_1 + 10x_2 \leq 8000$$

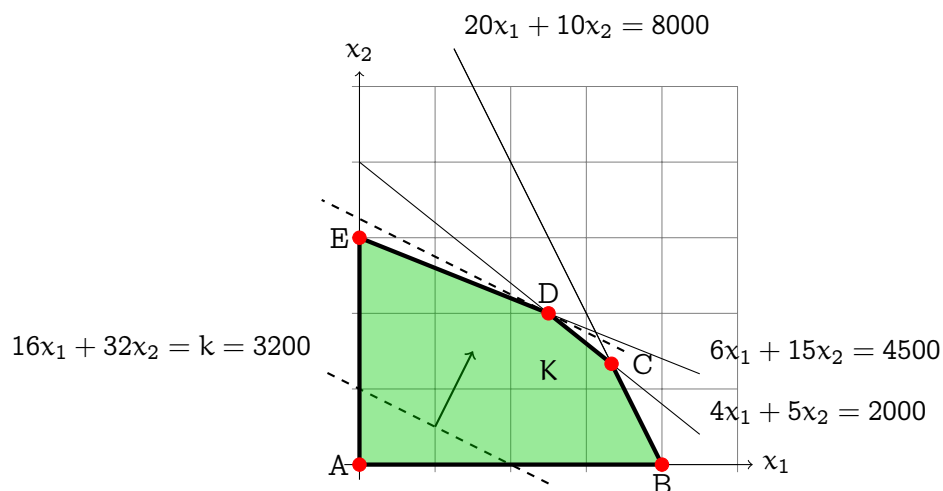
$$4x_1 + 5x_2 \leq 2000$$

$$6x_1 + 15x_2 \leq 4500$$

$$x_1 \geq 0$$

$$x_2 \geq 0$$

Die obigen Ungleichungen beschreiben Halbebenen im $\mathbb{R}^2$, die durch Geraden begrenzt werden. Der Durchschnitt aller dieser Halbebenen beschreibt den zulässigen Bereich aller Punktepaare (x_1, x_2) , die den Ungleichungen genügen. Die Niveaulinien $z(x_1, x_2) = 16x_1 + 32x_2 = k$, mit variablem k , beschreibt eine Schar von parallelen Geraden. Es ist ein maximales k_m derart zu bestimmen, dass (mindestens) ein Punkt des zulässigen Bereichs auf der Geraden $z(x_1, x_2) = k_m$ liegt.



Offenbar nimmt die Zielfunktion z ihren maximalen Wert in der Ecke D des zulässigen Bereichs K an. Die Ecken von K sind A, B, C, D, E ; z.B. berechnet man die Ecken B, C, D

mit zugehörigem Wert z :

Dabei bezeichnen $y_1 = 8000 - 20x_1 - 10x_2$, $y_2 = 2000 - 4x_1 - 5x_2$ und $y_3 = 4500 - 6x_1 - 15x_2$ die ungenutzten Ressourcen. Eine Gittereinheit in der Abbildung entspricht 100 Produktionseinheiten.

Tabelle 3.2: Angaben zum Produktionsproblem

Ecke	z	x_1	x_2	y_1	y_2	y_3
A	0	0	0	8000	2000	4500
B	6400	400	0	0	400	2100
C	9600	1000/3	400/3	0	0	500
D	10400	250	200	1000	0	0
E	9600	0	300	5000	500	0

In diesem Beispiel haben in jeder Ecke genau zwei der fünf Variablen den Wert 0 (nicht-entartete Basislösungen). Bei Durchlaufen der Ecken $B \rightarrow C \rightarrow D$ vergrößert sich der Wert der Zielfunktion. Das Optimum in der Ecke D ist:

$$x_1 = 250, x_2 = 200, z(x_1, x_2) = 10400.$$

3.1.2 Mischungsaufgaben

Es seien n Mischungen gegeben, die aus m Stoffen zusammengesetzt sind. Eine Einheit der Mischung M_j enthalte a_{ij} Einheiten des Stoffes S_i . Aus den Mischungen $M_1, \dots, M_n$ soll eine neue Mischung hergestellt werden. Mit x_j wird der Anteil der Mischung M_j bei der neuen Mischung bezeichnet. Seien I, J Teilmengen von $\{1, \dots, m\}$. Eine Einheit der neuen Mischung enthalte von jedem Stoff S_i für $i \in I$ mindestens b_i Einheiten und von jedem Stoff s_i für $i \in J$ höchstens b'_i Einheiten. Die Kosten pro Einheit der Mischung M_j seien c_j . Wir erhalten die Aufgabe:

$$\begin{aligned}
&\text{minimiere } z(x_1, \dots, x_n) = \sum_{j=1}^n c_j x_j \\
&\text{unter d.N.} \\
&\sum_{j=1}^n a_{ij} x_j \geq b_i, i \in I \\
&\sum_{j=1}^n a_{ij} x_j \leq b'_i, i \in J \\
&\sum_{j=1}^n x_j = 1 \\
&x_j \geq 0, j = 1, \dots, n
\end{aligned}$$

3.1.3 Transportprobleme

In m Fabriken werden jeweils die Warenmengen $a_i, 1 \leq i \leq m$ erzeugt. Diese Waren sollen an n Abnehmer, die jeweils die Warenmengen $b_j, 1 \leq j \leq n$ benötigen, so versandt werden, dass die Gesamtkosten minimal werden. Bezeichnen wir mit c_{ij} die Transportkosten pro Wareneinheit von Fabrik i zum Abnehmer j , so lautet das zugehörige Modell, in dem mit x_{ij} die tatsächlich von i nach j versandte Menge bezeichnet wird:

$$\begin{aligned}
&\text{minimiere } z(x_1, \dots, x_n) = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \\
&\text{unter d.N.} \\
&\sum_{j=1}^n x_{ij} = a_i, 1 \leq i \leq m, \\
&\sum_{i=1}^m x_{ij} = b_j, 1 \leq j \leq n, \\
&x_{ij} \geq 0, i = 1, \dots, m, j = 1, \dots, n.
\end{aligned}$$

Ist $n = m$, und $a_i = b_j = 1$ für alle i, j , so geht das Transportproblem in ein Zuordnungsproblem über (Personaleinsatzplanung).

3.2 Mathematische Formulierung linearer Optimierungsprobleme

Das lineare Optimierungsproblem (LP) lautet in Standardform

$$\begin{aligned}
&\text{minimiere } z(x_1, \dots, x_n) = \sum_{j=1}^n c_j x_j \\
&\text{unter d.N.} \\
&\sum_{j=1}^n a_{ij} x_j = b_i, 1 \leq i \leq m \\
&x_j \geq 0, j = 1, \dots, n
\end{aligned}$$

In Vektor-Schreibweise mit

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, c = (c_1, \dots, c_n) \in \mathbb{R}^n, b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m, A = (a_{ik}) \text{ } m \times n \text{ Matrix}$$

lautet die Standardform

$$\begin{aligned}
&\text{minimiere } z(x) = cx \\
&\text{unter d.N.} \\
&Ax = b \\
&x \geq 0.
\end{aligned} \tag{3.1}$$

3.3 Rückführung anderer LP's auf Standardform

Haben wir ein lineares Optimierungsproblem, welches nicht in Standardform gegeben ist (weil es Ungleichungsnebenbedingungen oder freie Variablen enthält), so können wir dieses wie folgt auf Standardform bringen.

3.3.1 Ungleichungen

Sei A eine $m \times n$ Matrix (nicht notwendig $m < n$). Das Problem lautet

$$\begin{aligned}
&\text{minimiere } z(x) = cx \\
&\text{unter d.N.} \\
&Ax \leq b \\
&x \geq 0.
\end{aligned} \tag{3.2}$$

Definiere Schlupf-Variablen

$$y := b - Ax \in \mathbb{R}^m.$$

Dann sind die Ungleichungsnebenbedingungen $Ax \leq b$ äquivalent zu

$$Ax + y = b, \quad y \geq 0.$$

Mit $\tilde{c} := (c, 0) \in \mathbb{R}^{n+m}$, $\tilde{x} := \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n+m}$, I der $m \times m$ Einheitsmatrix und der $m \times (n + m)$ -Matrix $\tilde{A} := (A|I)$ ist (3.2) äquivalent zur Standardform

$$\begin{aligned} &\text{minimiere } \tilde{z}(\tilde{x}) = \tilde{c}\tilde{x} \\ &\text{unter d.N.} \\ &\quad \tilde{A}\tilde{x} = b \\ &\quad \tilde{x} \geq 0. \end{aligned} \tag{3.3}$$

3.3.2 Freie Variable

In der Standardform (3.1) seien einige x_i ohne Vorzeichenbedingung, z.B. x_1 . Das Problem lautet also

$$\begin{aligned} &\text{minimiere } z(x) = cx \\ &\text{unter d.N.} \\ &\quad Ax = b \\ &\quad x_2 \geq 0, \dots, x_n \geq 0. \end{aligned} \tag{3.4}$$

1. Methode: Elimination von x_1

Falls die erste Spalte von A nicht null ist, wähle einen Index i mit $a_{i1} \neq 0$ und eliminiere aus

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i$$

die Variable x_1 als Linearkombination von $x_2, \dots, x_n$. Diesen Ausdruck setzt man auch in die Zielfunktion ein; die entstehende Konstante muss bei der Berechnung des ursprünglichen Optimalwerts berücksichtigt werden. Ist die erste Spalte null, kann x_1 bei $c_1 = 0$ weggelassen werden. Bei $c_1 \neq 0$ ist das Problem unbeschränkt, sofern die übrigen Nebenbedingungen zulässig sind. Die zweite Methode funktioniert in allen Fällen. Dies ergibt bei erfolgreicher Elimination ein reduziertes LGS

$$\tilde{A} \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} = \tilde{b} \in \mathbb{R}^{m-1}.$$

Wir betrachten ein Beispiel.

$$\begin{aligned} &\text{minimiere } z(x_1, x_2, x_3) = x_1 + 3x_2 + 4x_3 \\ &\text{unter d.N.} \\ &x_1 + 2x_2 + x_3 = 5 \\ &2x_1 + 3x_2 + x_3 = 6 \\ &x_2 \geq 0, x_3 \geq 0. \end{aligned}$$

Die freie Variable x_1 wird von der ersten Gleichung eliminiert:

$$x_1 = 5 - 2x_2 - x_3.$$

Durch Einsetzen dieses Ausdrucks erhält man (subtrahiere 5 in der Zielfunktion)

$$\begin{aligned} &\text{minimiere } z(x_1, x_2, x_3) = x_2 + 3x_3 \\ &\text{unter d.N.} \\ &x_2 + x_3 = 4 \\ &x_2 \geq 0, x_3 \geq 0. \end{aligned}$$

Lösung: $x_2 = 4, x_3 = 0 \Rightarrow x_1 = -3$.

2. Methode:

Setze $x_1 = u_1 - v_1, u_1 := \max\{0, x_1\} \geq 0, v_1 := \max\{0, -x_1\} \geq 0$. Dies ergibt ein LP in den $n + 1$ Variablen $u_1, v_1, x_2, \dots, x_n$.

Kapitel 4

Grundlagen der Polyedertheorie

Wir beschäftigen uns in diesem Kapitel mit einigen grundlegenden Konzepten der Polyedertheorie. Insbesondere werden anschaulich klare Begriffe wie Seitenflächen und Ecken von Polyedern definiert und charakterisiert.

Wie wir bereits gesehen haben, ist die Menge

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

ein Polyeder in Standardform. Da die Beschreibung des Polyeders durch die Matrizen A, b gegeben ist, schreiben wir auch $P^=(A, b) := K$. Ebenso ist die Menge

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

ein Polyeder und hierfür schreiben wir $P(A, b) := K$.

| **Definition 4.1.** Ein Polyeder das beschränkt ist, heisst Polytop.

4.1 Seitenflächen von Polyedern

Wir werden im Folgenden Polyeder der Form $P(A, b)$ betrachten, wobei $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.

| **Definition 4.2.** Es seien $K \subseteq \mathbb{R}^n, a \in \mathbb{R}^n, \alpha \in \mathbb{R}$.

1. Die Ungleichung $a^T x \leq \alpha$ heisst gültig bezüglich K , falls gilt

$$K \subseteq \{x \in \mathbb{R}^n \mid a^T x \leq \alpha\}.$$

2. Die Hyperebene $H = \{x \in \mathbb{R}^n \mid a^T x = \alpha\}, a \neq 0$, heisst Stützhyperebene von K , falls $a^T x \leq \alpha$ gültig bezüglich K und $K \cap H \neq \emptyset$ ist.

Beachte, dass $a = 0$ im ersten Punkt von Definition 4.2 zugelassen ist.

Definition 4.3. Sei $K \subset \mathbb{R}^n$. Eine Menge $F \subseteq K$ heisst Seitenfläche von K , wenn es eine bezüglich K gültige Ungleichung $d^T x \leq \delta$ gibt mit

$$F = K \cap \{x \in \mathbb{R}^n \mid d^T x = \delta\}.$$

Eine Seitenfläche heisst echt, wenn $F \neq K$ gilt. F heisst nichttrivial, falls $\emptyset \neq F \neq K$ gilt. Ist $d^T x \leq \delta$ gültig für K , dann heisst $K \cap \{x \in \mathbb{R}^n \mid d^T x = \delta\}$ die von $d^T x \leq \delta$ induzierte Seitenfläche.

Beachte wiederum, dass $d = 0$ in Definition 4.3 zugelassen ist.

Beispiel 4.4. Wir betrachten folgendes Beispiel.

$$\begin{aligned} 2x_1 + x_2 &\leq 8 \\ 4x_1 + 5x_2 &\leq 20 \\ 2x_1 + 5x_2 &\leq 15 \\ x_1, x_2 &\geq 0 \end{aligned}$$

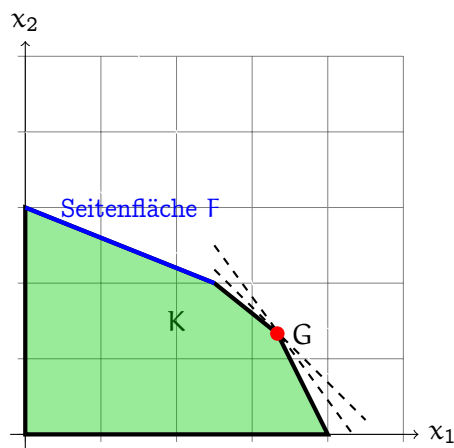


Abbildung 4.1: Beispiel einer Seitenfläche.

Die Seitenfläche F ist durch die Ungleichung

$$2x_1 + 5x_2 \leq 15$$

induziert, denn es gilt

$$F = K \cap \{x \in \mathbb{R}_+^2 \mid 2x_1 + 5x_2 = 15\}.$$

Ebenso ist die Menge $G = \{(\frac{10}{3}, \frac{4}{3})\}$ eine Seitenfläche, die man erhält, wenn beide folgenden

Ungleichungen gleichzeitig mit Gleichheit gelten (vgl. Proposition 4.8):

$$2x_1 + x_2 \leq 8$$

$$4x_1 + 5x_2 \leq 20.$$

Man beachte dass $G = \{(\frac{10}{3}, \frac{4}{3})\}$ als Seitenfläche von K auch durch die beiden gültigen Ungleichungen (siehe Abbildung 4.1)

$$4x_1 + 3x_2 \leq 52/3$$

$$x_1 + x_2 \leq 14/3$$

induziert wird. Eine Seitenfläche kann somit durch unterschiedliche gültige Ungleichungen induziert werden.

Proposition 4.5. Sei $K = P(A, b) \subset \mathbb{R}^n$ ein Polyeder. Dann gilt:

1. K ist eine Seitenfläche von sich selbst.
2. $\emptyset$ ist eine Seitenfläche von K .
3. Ist $F = \{x | d^T x = \delta\} \cap K$ eine nichttriviale Seitenfläche von K , dann gilt $d \neq 0$.

Beweis. (1): $K = K \cap \{x \in \mathbb{R}^n | 0^T x = 0\}$.

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

(3): Für $d = 0$ liefert nur der Fall $\delta \geq 0$ eine gültige Ungleichung der Form $d^T x \leq \delta$ und somit tritt einer der beiden ersten Fälle ein. $\square$

Satz 4.6. Sei $K = P(A, b)$ ein nichtleeres Polyeder und $c^T \in \mathbb{R}^n$. Betrachte das lineare Optimierungsproblem

$$\min\{cx | x \in K\}.$$

Sei F^* die Lösungsmenge und im Fall $F^* \neq \emptyset$ bezeichne $z^* = \min\{cx | x \in K\}$ den Optimalwert. Dann gilt:

1. Ist $F^* \neq \emptyset$, dann ist $F^* = \{x \in K | cx = z^*\}$ eine nichtleere Seitenfläche von K und $\{x \in \mathbb{R}^n | cx = z^*\}$ ist eine Stützhyperebene von K , falls $c \neq 0$.
2. Die Menge der Optimallösungen des linearen Problems $\min\{cx | x \in K\}$ ist eine Seitenfläche von K .

Beweis. Zu (1):

Sei $F^* \neq \emptyset$. Nach Definition von z^* ist $cx \geq z^*$ für alle $x \in K$, somit ist die Ungleichung $-cx \leq -z^*$ gültig für K . Desweiteren gilt $F^* = \{x \in \mathbb{R}^n | cx = z^*\} \cap K$, was impliziert dass F^* eine nichtleere Seitenfläche von K ist. Zudem folgt unmittelbar, dass die Menge $\{x \in \mathbb{R}^n | cx = z^*\}$ eine Stützhyperebene ist, falls $c \neq 0$.

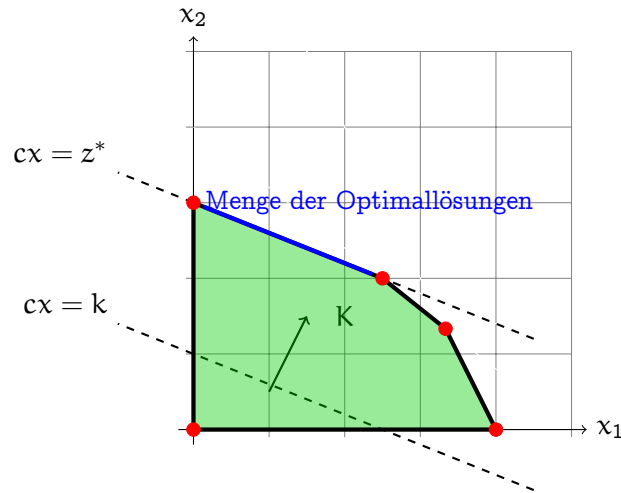


Abbildung 4.2: Graphische Veranschaulichung des Satzes (4.6).

Zu (2): Für $F^* = \emptyset$ ist die Aussage klar, ansonsten ist (2) nur eine Umformulierung von (1). $\square$

Eine Illustration des obigen Satzes findet sich in Abbildung 4.2.

Zu einer Matrix $A \in \mathbb{R}^{m \times n}$ mit Zeilenindexmenge $M := \{1, \dots, m\}$ und Spaltenindexmenge $N = \{1, \dots, n\}$ bezeichne mit $a_i, i \in M$ die Zeilen und $a^j, j \in N$ die Spalten von A . Für $I \subseteq M$ bezeichne A_I die aus den Zeilen $a_i, i \in I$ gebildete Teilmatrix.

Definition 4.7. Sei $K = P(A, b) \subset \mathbb{R}^n$ und M die Zeilenindexmenge von A . Für $F \subseteq K$ sei

$$eq(F) := \{i \in M \mid a_i x = b_i \forall x \in F\},$$

d.h. $eq(F)$ ist die Menge der für alle $x \in F$ aktiven Ungleichungen. Für $I \subseteq M$ bezeichne

$$fa(I) := \{x \in K \mid A_I x = b_I\}$$

die von I induzierte Seitenfläche (face).

Wir überprüfen nun, dass $fa(I)$ tatsächlich eine Seitenfläche von K ist.

Proposition 4.8. Die in Definition 4.7 definierte Menge $F := fa(I)$ ist eine Seitenfläche von K .

Beweis. Falls $I = \emptyset$, ist auch $F = K$ eine (triviale) Seitenfläche von K . Sei nun $|I| \geq 1$. Definiere

$$a^T := \sum_{i \in I} a_i \text{ und } \gamma := \sum_{i \in I} b_i.$$

Es gilt, dass $a^T x \leq \gamma$ gültige Ungleichung ist und weiterhin gilt für $x \in K \setminus F$, dass

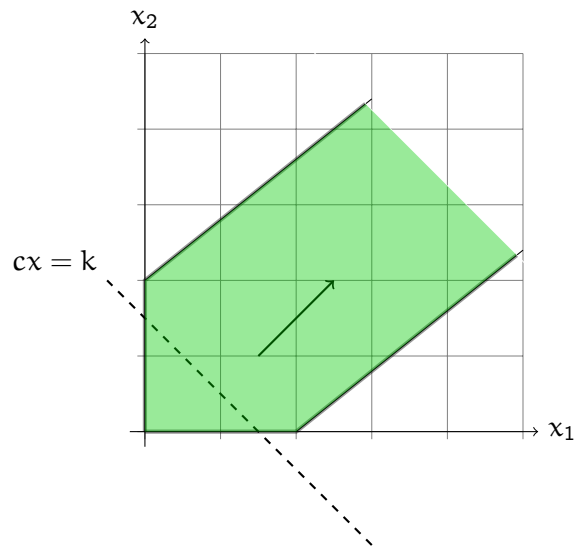


Abbildung 4.3: In diesem Beispiel ist die Menge der Optimallösungen leer.

mindestens eine der Ungleichungen nicht strikt ist und somit

$$a^T x = \gamma \quad \text{für } x \in F, \quad a^T x < \gamma \quad \text{für } x \in K \setminus F.$$

Es folgt $F = \{x \in K \mid a^T x = \gamma\} = K \cap \{x \in \mathbb{R}^n \mid a^T x = \gamma\}$. □

Wir betrachten nochmals Beispiel 4.4. Hier ist $M = \{1, 2, 3, 4, 5\}$, $fa(\{1, 2\}) = G$ und $eq(G) = \{1, 2\}$.

4.2 Ecken und Extrempunkte

Wir geben nun eine Definition von Ecken und Extrempunkten an.

Definition 4.9. Sei $K \subset \mathbb{R}^n$.

1. Ein Punkt $x \in K$ heißt Extrempunkt von K , wenn es nicht zwei verschiedene Punkte $y, z \in K$ gibt mit

$$x = \alpha y + (1 - \alpha)z \quad \text{für ein } \alpha \in (0, 1).$$

2. Ein Punkt $x \in K$ heißt Ecke von K , falls $\{x\}$ eine (nulldimensionale) Seitenfläche von K ist.

Man beachte dass Definition 4.9 für beliebige Mengen $K \subset \mathbb{R}^n$ definiert ist.

Satz 4.10. Sei $K = P(A, b) \subset \mathbb{R}^n$ ein Polyeder und $x \in K$. Dann sind die folgenden Aussagen äquivalent:



Abbildung 4.4: Die schwarzen Punkte des Polytopes sind die Ecken. Für einen abgeschlossenen Kreisscheibe sind genau die Randpunkte Ecken.

1. x ist eine Ecke von K .
2. $\{x\}$ ist eine nulldimensionale Seitenfläche von K .
3. x ist ein Extrempunkt von K .
4. $\text{rang}(A_{\text{eq}(\{x\})}) = n$.
5. Es existiert ein $c^T \in \mathbb{R}^n \setminus \{0\}$, so dass x die eindeutig bestimmte Optimallösung des linearen Problems $\min\{cy : y \in K\}$ ist.

Beweis. Die Äquivalenz der Aussagen (1) und (2) ist gerade die Definition 4.9 einer Ecke. Der Beweis besteht aus den Teilen (2) $\Rightarrow$ (5), (5) $\Rightarrow$ (3), (3) $\Rightarrow$ (4) und (4) $\Rightarrow$ (2).

(2) $\Rightarrow$ (5):

Nach Definition ist $\{x\}$ eine Seitenfläche, also existiert eine bezüglich K gültige Ungleichung $d^T y \leq \gamma$, so dass $\{y \in K \mid d^T y = \gamma\} = \{x\}$.

Folglich ist x die eindeutig bestimmte Optimallösung von $\min\{cy : y \in K\}$ für $c := -d^T$.

Ist $K \neq \{x\}$, so ist $c \neq 0$ wegen Proposition 4.5(3), andernfalls kann $c \neq 0$ beliebig gewählt werden.

(5) $\Rightarrow$ (3):

Sei x die eindeutige Optimallösung von $\min\{cy \mid y \in K\}$ mit Wert γ . Gilt $x = \lambda w + (1 - \lambda)z$ für $w, z \in K, w \neq z, 0 < \lambda < 1$, dann folgt

$$\begin{aligned} \gamma &= cx = c(\lambda w + (1 - \lambda)z) \\ &= \lambda cw + (1 - \lambda)cz \\ &> \lambda\gamma + (1 - \lambda)\gamma = \gamma, \end{aligned}$$

ein Widerspruch.

(3) $\Rightarrow$ (4):

Angenommen (4) gilt nicht. Dann existiert $d \neq 0$ mit $A_{\text{eq}(\{x\})}d = 0$. Für kleines $\epsilon > 0$ gilt dann

$$A(x \pm \epsilon d) \leq b,$$

denn die bindenden Nebenbedingungen bleiben aktiv, die inaktiven Nebenbedingungen bleiben für $\epsilon > 0$ klein genug inaktiv. Mit $y = x - \epsilon d$, und $z = x + \epsilon d$ gilt dann $y, z \in K$ und $x = 1/2y + 1/2z$. Also gilt auch (3) nicht.

(4) $\Rightarrow$ (2):

Wir wissen, dass $\text{fa}(\text{eq}(\{x\}))$ eine Seitenfläche ist und wegen $\text{rang}(A_{\text{eq}(\{x\})}) = n$ gilt

$$\text{fa}(\text{eq}(\{x\})) = \{y \in K \mid A_{\text{eq}(\{x\})}y = b_{\text{eq}(\{x\})}\} = \{x\}.$$

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

Wir übertragen nun die obige Charakterisierung von Ecken auf Polyeder in Standardform.

Satz 4.11. Sei $K = P^-(A, b)$, d.h.

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

$x \in K$ ist genau dann ein Extrempunkt (bzw. äquivalent Ecke), wenn die zu positiven Komponenten x_k gehörenden Spaltenvektoren a^k von A linear unabhängig sind.

Beweis. $\Rightarrow$: $x \in K$ sei eine Ecke. O.E. sei

$$x = (x_1, \dots, x_r, 0, \dots, 0)^T, x_i > 0, i = 1, \dots, r.$$

Ist $r = 0$, so ist die zugehörige Menge der Spaltenvektoren leer, also per Definition linear unabhängig. Für $r > 0$ gilt

$$\sum_{i=1}^r a^i x_i = b.$$

Annahme: $a^1, \dots, a^r$ seien linear abhängig. Dann gibt es Zahlen $d_1, \dots, d_r, d_i \neq 0$ für mindestens ein i , mit

$$\sum_{i=1}^r a^i d_i = 0.$$

Wenn $x_i > 0$ ist für ein hinreichend kleines $\epsilon > 0$ auch

$$x_i \pm \epsilon d_i > 0, \text{ für } i = 1, \dots, r.$$

Wir setzen

$$d := (d_1, \dots, d_r, 0, \dots, 0)^T, y := x + \epsilon d, z := x - \epsilon d.$$

Dann sind $y, z \geq 0$ und wegen

$$\sum_{i=1}^r a^i (x_i \pm \epsilon d_i) = \sum_{i=1}^r a^i x_i \pm \epsilon \underbrace{\sum_{i=1}^r a^i d_i}_{=0} = b,$$

gilt $y, z \in K$. Nun ist $x \neq y, z$, aber $x = \frac{y+z}{2}$ im Widerspruch dazu, dass x Ecke (Extrempunkt) von K ist. Daher sind $a^1, \dots, a^r$ linear unabhängig.

$\Leftarrow$: O.E. seien die ersten r Komponenten von x positiv, also nach Vor. $a^1, \dots, a^r$ linear unabhängig.

1. Fall: $r = 0 \Rightarrow x = 0$. Wäre $x = 0$ kein Extrempunkt (Ecke), so gäbe es $y, z \in K, y \neq z$

und $0 < \alpha < 1$ mit

$$0 = x = \alpha y + (1 - \alpha)z.$$

Wegen $y, z \geq 0$ und $\alpha \neq 0$ folgt hieraus $y = 0, z = 0$, Widerspruch.

2. Fall: $r > 0$: Nach Def. ist $\sum_{i=1}^r a^i x_i = b$.

Annahme: x ist kein Extrempunkt (Ecke) von K . Es gibt $y, z \in K$, $y \neq z$, und $0 < \alpha < 1$ mit

$$x = \alpha y + (1 - \alpha)z.$$

Wie im 1. Fall folgt

$$y_{r+1} = \dots = y_n = 0, z_{r+1} = \dots = z_n = 0.$$

Ferner ist

$$Ay = Az = b \text{ also } A(y - z) = 0 \Rightarrow \sum_{i=1}^r a^i (y_i - z_i) = 0$$

Weil $a^1, \dots, a^r$ linear unabhängig sind, folgt

$$y_i = z_i \text{ für } i = 1, \dots, r \Rightarrow y = z \Rightarrow x = y = z, \text{ Widerspruch.}$$

□

Definition 4.12. Ein Polyeder heißt spitz, wenn es eine Ecke besitzt.

Neben den Ecken eines Polyeders, sind auch weitere niedrigdimensionale Seitenflächen von Polyedern relevant. Wir definieren die Begriffe Kante und Linie in einem Polyeder.

Definition 4.13. Ein Polyeder $K \subset \mathbb{R}^n$ enthält eine Linie, falls $x \in K$ und eine Richtung $d \in \mathbb{R}^n \setminus \{0\}$ existiert, so dass gilt

$$x + \lambda d \in K \text{ für alle } \lambda \in \mathbb{R}.$$

Eine Kante von K ist eine Seitenfläche der Dimension 1. Bei einem spitzen Polyeder ist eine Kante entweder eine Strecke zwischen zwei Ecken oder ein von einer Ecke ausgehender Strahl.

Satz 4.14. Sei $K = P(A, b) \subset \mathbb{R}^n$ ein nichtleeres Polyeder. Die folgenden Bedingungen sind äquivalent.

1. K ist spitz.
2. $\text{rang}(A) = n$.
3. Jede nichtleere Seitenfläche von K ist spitz.

Beweis. (1) $\Rightarrow$ (2): Ist x eine Ecke von K , so gilt mit Satz 4.10 dass

$$n = \text{rang}(A_{\text{eq}(\{x\})}) \leq \text{rang}(A) \leq n.$$

Also $\text{rang}(A) = n$.

(2) $\Rightarrow$ (1) : Wir wählen $x \in K$ so, dass die Menge $I = \text{eq}(\{x\})$ inklusionsmaximal ist. Sei

$$F = \{x \in K \mid A_I x = b_I\}.$$

Falls $\text{rang}(A_I) = n$, so gilt mit Satz 4.10 bereits dass x eine Ecke ist, daher nehmen wir an dass $\text{rang}(A_I) < n$ gilt. Somit enthält der Kern von A_I einen Vektor $d \neq 0$ so dass gilt

$$x \pm \epsilon d \in K, \text{ für } \epsilon \text{ klein genug.}$$

Die Linie $\{x + \lambda d \mid \lambda \in \mathbb{R}\}$ trifft nun mindestens eine der Hyperebenen $H_j = \{x \in \mathbb{R}^n \mid a_j x = b_j\}$ für ein $j \notin I$. (Angenommen nicht, d.h. die gesamte Linie liegt vollständig in K . Dann gilt

$$a_i(x + \lambda d) \leq b_i, \text{ für alle Zeilenindizes } i, \text{ und alle } \lambda \in \mathbb{R}.$$

Diese Aussage impliziert $Ad = 0$ und mit $\text{rang}(A) = n$ folgt $d = 0$, Widerspruch.) Somit existiert $\delta \in \mathbb{R}$ sodass $x + \delta d \in K$ und $\text{eq}(\{x + \delta d\}) \supset I$, im Widerspruch zur Maximalität von I .

(3) $\Rightarrow$ (1): Ist trivial, da K eine Seitenfläche von sich selbst ist.

(2) $\Rightarrow$ (3): Sei $F = \{x \in K \mid dx = \delta\}$ eine nichtleere, durch eine gültige Ungleichung $dx \leq \delta$ induzierte Seitenfläche. Man erhält eine Ungleichungsdarstellung von F , indem man zu $Ax \leq b$ die Ungleichungen $dx \leq \delta$ und $-dx \leq -\delta$ hinzufügt. Die erweiterte Matrix hat weiterhin Rang n . Aus der bereits bewiesenen Äquivalenz folgt, dass F spitz ist. $\square$

Wir erhalten eine wichtige Folgerung für Polyeder in Standardform.

Korollar 4.15. Sei $K = P^=(A, b)$, dann gilt

$$K \neq \emptyset \Leftrightarrow K \text{ ist spitz.}$$

Beweis. Wir erhalten eine Darstellung von K durch

$$K = P^=(A, b) = P(D, d) \text{ mit } D = \begin{pmatrix} A \\ -A \\ -I \end{pmatrix}, d = \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}.$$

Offenbar hat D den Rang n . Aus Satz 4.14 folgt dann die Behauptung. $\square$

Für Polytope erhalten wir ein ähnliches Resultat.

Korollar 4.16. Sei $K = P(A, b)$ ein Polytop. Dann gilt

$$K \neq \emptyset \Leftrightarrow K \text{ ist spitz.}$$

Beweis. Da K beschränkt ist, gibt es einen Vektor u mit $K \subseteq \{x | x \leq u\}$. So erhalten wir eine Darstellung von K durch

$$K = P(A, b) = P(D, d) \text{ mit } D = \begin{pmatrix} A \\ I \end{pmatrix}, d = \begin{pmatrix} b \\ u \end{pmatrix}.$$

D hat den Rang n . Aus Satz 4.14 folgt dann die Behauptung. $\square$

Korollar 4.17. Sei $K = P(A, b)$ ein spitzes Polyeder und das LP

$$\min cx \text{ u.d.N. } x \in K$$

habe eine endliche Optimallösung. Dann hat das LP auch eine Optimallösung die eine Ecke ist.

Beweis. $F = \{x \in K | cx = \min\{cy | y \in K\}\}$ ist eine nichtleere Seitenfläche von K und besitzt nach Satz 4.14 eine Ecke. Eine Ecke von F ist auch Ecke von K : Bei einer echten konvexen Zerlegung in Punkte aus K müssten wegen der Optimalität beide Punkte ebenfalls in F liegen. $\square$

Korollar 4.18. Wenn K ein nichtleeres Polytop ist, dann hat jedes LP

$$\min cx \text{ u.d.N. } x \in K$$

auch eine optimale Ecklösung.

Wir fassen nun die bisherigen Resultate im folgenden wichtigen Hauptsatz der linearen Optimierung zusammen.

Satz 4.19. Ist das LP $\min\{cx | Ax = b, x \geq 0\}$ zulässig und seine Zielfunktion nach unten beschränkt, so besitzt es eine optimale Ecke.

Beweis. Sei x zulässig. Sind die zu $x_i > 0$ gehörenden Spalten linear abhängig, wähle $d \neq 0$ mit $Ad = 0$ und $d_i = 0$ für $x_i = 0$. Die Richtung kann so orientiert werden, dass mindestens ein $d_i > 0$ und $cd \geq 0$ gilt. Hat d beide Vorzeichen, wähle das Vorzeichen entsprechend cd . Andernfalls wähle $d \geq 0$; dann würde $cd < 0$ entlang $x + td$, $t \geq 0$, der Beschränktheit widersprechen. Setze

$$\mu = \min_{d_i > 0} \frac{x_i}{d_i}, \quad x' = x - \mu d.$$

Dann ist x' zulässig, $cx' \leq cx$, und x' hat weniger positive Komponenten. Nach endlich vielen Wiederholungen erhält man nach Satz 4.11 eine Ecke mit nicht größerem Zielfunktionswert. Es gibt nur endlich viele Ecken: Jeder linear unabhängige Spaltensatz bestimmt höchstens eine Lösung auf seinem Träger. Die beste dieser Ecken ist daher optimal. $\square$

Der Hauptsatz 4.19 kann im Prinzip zur Lösung eines LP's in Standardform benutzt werden indem man die endlich vielen Ecken von K "ausprobiert".

Beispiel 4.20.

$$\begin{aligned} \min & x_1 + 2x_2 + 3x_3 \\ \text{unter d.N.} \\ & 2x_1 + x_2 + 5x_3 = 5 \\ & x_1 + 2x_2 + x_3 = 4 \\ & x_i \geq 0, i = 1, 2, 3 \end{aligned}$$

Lösung: Für eine Ecke ist mindestens eine der Komponenten gleich 0.

1. Möglichkeit: $x_1 = 0$. Löse

$$\begin{bmatrix} 1 & 5 \\ 2 & 1 \end{bmatrix} \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \Rightarrow x_2 = 5/3, x_3 = 2/3$$

Ecke: $(0, 5/3, 2/3)^T$, Zielfunktion $16/3$.

2. Möglichkeit: $x_2 = 0$. Löse

$$\begin{bmatrix} 2 & 5 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \Rightarrow x_1 = 5, x_3 = -1$$

Nicht zulässig.

3. Möglichkeit: $x_3 = 0$. Löse

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \Rightarrow x_1 = 2, x_2 = 1$$

Ecke: $(2, 1, 0)^T$, Zielfunktion 4.

Da die zulässige Menge K beschränkt ist, nimmt die Zielfunktion im Punkt $(2, 1, 0)^T$ ihr Minimum an.

Im Gegensatz zum "Ausprobieren" stellt der Simplex-Algorithmus (siehe Kapitel 5) einen "effizienten" Algorithmus zur Berechnung einer optimalen Ecke dar.

4.3 Basislösungen

Im Fall $m = n$ und $\text{rank}(A) = n$ gibt es genau einen Kandidaten $x = A^{-1}b$; man prüft nur noch $x \geq 0$. Der eigentliche Simplex-Fall ist $m < n$.

Wir betrachten ein LP in Standardform:

$$\text{minimiere } z(x) = cx$$

unter d.N.

$$Ax = b$$

$$x \geq 0$$

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

Wir nehmen an, dass $\text{rang}(A) = m \leq n$ gilt. (Später werden wir sehen dass diese Voraussetzung keine Beschränkung der Allgemeinheit darstellt). Wir betrachten das lineare Gleichungssystem $Ax = b$, dessen Lösungsmenge ein $(n - m)$ -dimensionaler affiner Unterraum ist. Die Spalten von A seien $a^j, j = 1, \dots, n$.

Definition 4.21. Ein Indexvektor $B = (i_1, \dots, i_m)$ mit m verschiedenen Indizes $i_j \in \{1, \dots, n\}$ heißt Basis, wenn die zugehörigen Spaltenvektoren linear unabhängig sind. Ein zu B komplementärer Indexvektor $N = (j_1, \dots, j_{n-m})$, $j_k \in \{1, \dots, n\}$ heißt Nichtbasis. Es gilt also die Zerlegung $B \oplus N = \{1, \dots, n\}$. Mit A_B und A_N werden die Untermatrizen bezeichnet, die durch die Spaltenvektoren in B bzw. N definiert sind:

$A_B : m \times m$ Matrix mit Spaltenvektoren $a^i, i \in B$

$A_N : m \times (n - m)$ Matrix mit Spaltenvektoren $a^j, j \in N$

Für eine solche Zerlegung der Indexmenge $\{1, \dots, n\}$ schreibt sich die Menge K äquivalent als

$$(A_B, A_N) \begin{pmatrix} x_B \\ x_N \end{pmatrix} = b, x_B, x_N \geq 0.$$

Wir bezeichnen für eine Basis B , die Matrix A_B als Basismatrix und komplementär A_N als Nichtbasismatrix. Die Variablen $x_i, i \in B$ heißen Basisvariablen und die Variablen $x_j, j \in N$ heißen Nichtbasisvariablen. Der Vektor $x = (x_B, x_N)$ mit

$$x_B := A_B^{-1}b \text{ und } x_N := 0$$

wird als Basislösung zur Basis B bezeichnet. Eine Basislösung heißt zulässig, falls $x_B \geq 0$. Eine zulässige Basislösung heißt nicht-entartet, falls $x_B > 0$; falls aber mehr als $n - m$ Komponenten von x gleich Null sind, so sprechen wir von einer degenerierten oder entarteten Basislösung.

Wir erhalten unmittelbar folgende Charakterisierung von Basislösungen aus Satz 4.11:

Satz 4.22. Sei $K = P^=(A, b)$. Dann sind die folgenden Aussagen äquivalent:

1. $x \in K$ ist ein Extrempunkt.
2. $x \in K$ ist eine Ecke.
3. $x \in K$ ist eine zulässige Basislösung zu einer Basis B .

Beweis. (1) $\Leftrightarrow$ (2) folgt aus Satz 4.10.

(1) $\Rightarrow$ (3) : Sei $x \in K$ ein Extrempunkt. Nach Satz 4.11 sind alle, die zu positiven Komponenten gehörenden Spaltenvektoren linear unabhängig. Somit sind die Spaltenvektoren $a^j, j \in J$ mit $J := \{i \in \{1, \dots, n\} : x_i > 0\}$ linear unabhängig und können zu einer Basis B ergänzt werden. Nach Definition von J gilt dann $x_N = 0$, wobei N die zu B gehörende Nichtbasis ist. Damit folgt

$$b = Ax = A_B x_B,$$

und es gilt, da A_B invertierbar ist,

$$x_B = A_B^{-1}b.$$

Somit ist x eine zulässige Basis-Lösung zur Basis B .

(3) $\Rightarrow$ (1) : Sei x eine zulässige Basislösung zu einer Basis B , das heisst es gibt eine Darstellung $x = (x_B, x_N)$ mit $x_N = 0$ und $x_B := A_B^{-1}b \geq 0$. Die Menge der Indizes der positiven Komponenten von x ist eine Teilmenge von B . Weil alle Spaltenvektoren $a^i, i \in B$ nach Voraussetzung linear unabhängig sind, folgt mit Satz 4.11 die Behauptung. $\square$

Der obige Satz besagt dass Ecken eines Polyeders $P^=(A, b)$ genau zulässigen Basislösungen entsprechen. Wir betonen ausdrücklich, dass die Korrespondenz zwischen Ecken und Basen im allgemeinen keine Bijektion ist. Verschiedene Ecken liefern verschiedene zugehörige Basen, aber die Aussage gilt nicht umgekehrt: Zu einer entarteten Ecke - also einer Ecke mit weniger als m positiven Komponenten - können verschiedene Basen gehören, weil die Erweiterungen zu einer Basis nicht eindeutig sein müssen.

$$\text{Ecke} \xleftrightarrow{\text{eindeutig}} \text{zulässige Basislösung} \xleftrightarrow{\text{nicht-eindeutig}} \text{zulässige Basis}.$$

Wir erhalten eine Folgerung bzgl. der maximalen Anzahl von Ecken bzw. zulässigen Basislösungen eines Polyeders K mit $\text{rang}(A) = m \leq n$.

Korollar 4.23. Es gibt höchstens $\binom{n}{m}$ verschiedene Ecken von K .

Beweis. Jede Ecke führt zu mindestens einer zulässigen Basis-Lösung mit zugehöriger Basismatrix. Jede Basismatrix besteht aus m linear unabhängigen Spalten von A und wir

können nun auf höchstens $\binom{n}{m}$ verschiedene Arten m linear unabhängige Spalten aus A auswählen. $\square$

Wir beenden dieses Kapitel mit der Aussage, dass die Annahme $\text{rang}(A) = m$ ohne Beschränkung der Allgemeinheit ist.

Satz 4.24. Sei $K = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$ ein nichtleeres Polyeder in Standardform mit Matrix $A \in \mathbb{R}^{m \times n}$. Sei $\text{rang}(A) = k < m$ und seien die Zeilenvektoren $a_{i_1}, \dots, a_{i_k}$ linear unabhängig. Betrachte das Polyeder

$$P := \{x \in \mathbb{R}^n | a_{i_1}x = b_{i_1}, \dots, a_{i_k}x = b_{i_k}, x \geq 0\}.$$

Dann gilt $K = P$.

Beweis. Nach Umbenennung der Indizes können wir annehmen dass gilt $i_1 = 1, \dots, i_k = k$. Trivialerweise gilt $K \subseteq P$, da jeder Punkt aus K natürlich alle Gleichungen aus P erfüllt. Wir zeigen nun $P \subseteq K$. Mit $\text{rang}(A) = k$ gilt dass der Zeilenraum von A Dimension k hat und die Vektoren $a_1, \dots, a_k$ eine Basis des Zeilenraums bilden. Somit kann jede Zeile a_i von A in der Form $a_i = \sum_{j=1}^k \lambda_{ij} a_j$ dargestellt werden, wobei λ_{ij} geeignete Skalare sind. Betrachte ein $x \in K$. Es gilt

$$b_i = a_i x = \sum_{j=1}^k \lambda_{ij} a_j x = \sum_{j=1}^k \lambda_{ij} b_j, \quad i = 1, \dots, m. \quad (4.1)$$

Sei nun $y \in P$. Wir zeigen dass $y \in K$. Für alle $i = 1, \dots, m$ folgt mit (4.1)

$$a_i y = \sum_{j=1}^k \lambda_{ij} a_j y = \sum_{j=1}^k \lambda_{ij} b_j = b_i,$$

somit folgt $y \in K$. $\square$

4.4 Degeneriertheit von Basislösungen

Bei einer zulässigen Basislösung sind alle m Gleichungen erfüllt. Mindestens $n - m$ Nichtnegativitätsbedingungen sind aktiv, da alle Nichtbasisvariablen null sind. Falls jedoch mehr als $n - m$ Variablen gleich null sind, so spricht man von einer entarteten Basislösung, siehe Definition 4.21.

Wir geben einige Beispiele an.

Beispiel 4.25. 1. Redundante Variablen:

$$\begin{aligned} x_1 + x_2 &= 2 \\ x_3 &= 0 \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

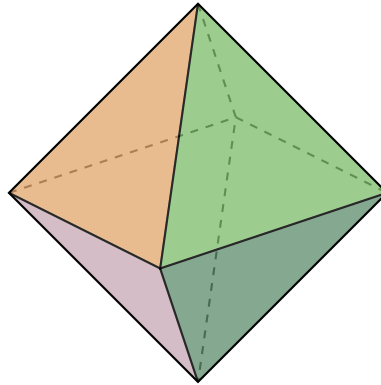


Abbildung 4.5: Die Ecken des Oktahedrons sind degeneriert.

2. Redundante Ungleichungen:

$$2x_1 + x_2 + y_1 = 3$$

$$x_1 + 2x_2 + y_2 = 3$$

$$x_1 + x_2 + y_3 = 2$$

$$x_1, x_2, y_1, y_2, y_3 \geq 0.$$

Im Punkt $x = (1, 1, 0, 0, 0)^T$ ist hier die Ungleichung $y_3 \geq 0$ redundant.

3. geometrische Gründe (siehe Oktahedron in Abbildung 4.5).

Bemerkung 4.26. Bei festem A mit vollem Zeilenrang macht eine Störung von b gemäß einer Verteilung mit Dichte alle Basislösungen mit Wahrscheinlichkeit eins nichtentartet. Eine solche Störung kann Zulässigkeit und Optimallösung des ursprünglichen Problems verändern.

Kapitel 5

Das Simplex Verfahren

Die am häufigsten benutzte Methode zur Lösung des Standardproblems

$$\begin{aligned} &\text{minimiere } z(x) = cx \\ &\text{unter d.N.} \\ &Ax = b \\ &x \geq 0. \end{aligned} \tag{5.1}$$

ist das Simplex-Verfahren. Generell wird $\text{rang}(A) = m < n$ vorausgesetzt. Nach dem Hauptsatz (4.19) ist die optimale Lösung (falls sie existiert) in einer Ecke der zulässigen Menge

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

zu suchen.

Das Simplex-Verfahren besteht aus den folgenden elementaren Schritten:

Ablauf auf Basisebene:

1. Bestimme eine zulässige Basis (Phase I), oder stelle Unzulässigkeit fest.
2. Sind alle reduzierten Kosten $r_j \leq 0$, ist die aktuelle Basislösung optimal. Andernfalls wähle $r_s > 0$.
3. Liefert der Quotiententest keine Ausgangsvariable, ist das LP unbeschränkt. Andernfalls führe einen zulässigen Basistausch aus und wiederhole Schritt 2.

Bei einem nichtentarteten Schritt sinkt der Zielfunktionswert und man gelangt zu einer benachbarten Ecke. Bei einem entarteten Schritt kann die Ecke unverändert bleiben; deshalb verfolgen wir Basen und nicht nur Ecken.

5.1 Parametrisierung des Lösungsraumes

Bei beliebiger Basis B geht das LGS $Ax = b$ über in

$$Ax = A_B x_B + A_N x_N = b.$$

Damit erhält man eine Parametrisierung des $n - m$ dimensionalen Lösungsraumes von $Ax = b$ durch

$$x_B = A_B^{-1}b - A_B^{-1}A_N x_N, \quad x_N \in \mathbb{R}^{n-m}, \quad (5.2)$$

wobei x_B die abhängige Variable und x_N die unabhängige Variable bezeichnet. Zerlege c entsprechend in $c_B \in \mathbb{R}^m$ und $c_N \in \mathbb{R}^{n-m}$. Durch Einsetzen des Ausdrucks (5.2) ergibt sich die Zielfunktion

$$\begin{aligned} z(x) &= cx = c_B x_B + c_N x_N \\ &= c_B A_B^{-1}b - (c_B A_B^{-1}A_N - c_N)x_N \\ &=: z_0 - r_N x_N, \end{aligned} \quad (5.3)$$

wobei

$$\begin{aligned} z_0 &:= c_B A_B^{-1}b \\ r_N &:= c_B A_B^{-1}A_N - c_N = (r_j)_{j \in N} \in \mathbb{R}^{n-m} \text{ (Vektor der reduzierten Kosten)} \end{aligned}$$

Nach Satz 4.22 sind die Basis-Lösungen gerade die Ecken von K . Die Darstellung (5.3) der Zielfunktion ergibt das folgende fundamentale Optimalitätskriterium:

Satz 5.1. Sei B eine Basis mit den Eigenschaften

1. die zugehörige Basis-Lösung x ist zulässig, d.h. $x_B \geq 0$,
2. $r_N = c_B A_B^{-1}A_N - c_N \leq 0$.

Dann ist x optimal für das LP (5.1) und der optimale Wert ist $z_0 = c_B A_B^{-1}b$.

Beweis. Für jeden zulässigen Punkt $\tilde{x}$ ist $\tilde{x}_N \geq 0$ und daher folgt aus (5.3) wegen $r_N \leq 0$

$$z(\tilde{x}) = c\tilde{x} = z_0 - r_N \tilde{x}_N \geq z_0 = z(x).$$

□

Für den Spezialfall eines LP mit Ungleichungsrestriktionen

$$\min cx \text{ u.d.N. } Ax \leq b, x \geq 0, \text{ wobei } b \geq 0,$$

erstellen wir ein äquivalentes Problem

$$\min cx + 0y \text{ u.d.N. } Ax + y = \tilde{A} \begin{pmatrix} x \\ y \end{pmatrix} = b, x \geq 0, y \geq 0,$$

wobei $\tilde{A} = (A, I)$. Wähle Basis $B = \{n+1, \dots, n+m\}$, $N = \{1, \dots, n\}$ und wir erhalten eine Basis-Lösung $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix} \in \mathbb{R}^{n+m}$. Die reduzierten Kosten ergeben sich mit $\tilde{A}_B = I$, $\tilde{c}_B = 0$, $\tilde{c}_N = c$ durch

$$r_N = \tilde{c}_B \tilde{A}_B^{-1} \tilde{A}_N - \tilde{c}_N = -c \in \mathbb{R}^n$$

$$z_0 = 0.$$

5.2 Basistausch

Das Simplex-Verfahren basiert nun auf dem hinreichenden Optimalitätskriterium von Satz 5.1. Man sucht eine Basis B mit

1. $x_B \geq 0$
2. $r_N \leq 0$.

Man startet dabei mit einer Basis B , für die 1. erfüllt ist. Wenn das Kriterium 2. verletzt ist, so geht man zu einer benachbarten Basis B' über (Basistausch), in der sich der Zielfunktionswert nicht vergrößert.

Definition 5.2. Zwei Basen B, B' heißen benachbart, wenn $|B \cap B'| = m-1$, also genau eine Basisvariable ausgetauscht wird. Die zugehörigen zulässigen Basislösungen können bei Entartung dieselbe Ecke beschreiben.

Zur Basis B benötigt man die allgemeine Darstellung (5.2)

$$x_B = A_B^{-1}b - A_B^{-1}A_N x_N, \quad x_N \in \mathbb{R}^{n-m}.$$

Die Multiplikation mit A_B^{-1} wird durch elementare Zeilenoperationen ausgeführt (Gauß-Jordan-Elimination). Das folgende Beispiel zeigt einen Basistausch.

Beispiel 5.3.

$$\begin{aligned} \text{minimiere } z(x_1, x_2, x_3) &= -x_1 - 2x_2 - 3x_3 \\ \text{unter d.N.} \\ 2x_1 + x_2 + 5x_3 &= 5 \\ x_1 + 2x_2 + x_3 &= 4 \end{aligned}$$

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0.$$

Wir wählen die Basis $B = (1, 2)$ und führen die Rechnungen in Tableau-Form an der Matrix $b|A$ durch:

b	x_1	x_2	x_3	
5	2	1	5	·1/2
4	1	2	1	

b	x_1	x_2	x_3	
5/2	1	1/2	5/2	-1. Zeile
4	1	2	1	

b	x_1	x_2	x_3	
5/2	1	1/2	5/2	·2/3
3/2	0	3/2	-3/2	

b	x_1	x_2	x_3	
5/2	1	1/2	5/2	-1/2 · 2. Zeile
1	0	1	-1	

b	x_1	x_2	x_3
2	1	0	3
1	0	1	-1

Wir erhalten nun folgende Darstellung:

$$x_B = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 3 \\ -1 \end{pmatrix} x_3 = A_B^{-1}b - A_B^{-1}A_N x_N.$$

Die Basislösung

$$x_1 = 2, x_2 = 1, x_3 = 0$$

ist zulässig aber nicht optimal, da für die reduzierten Kosten gilt

$$r_N = r_3 = c_B A_B^{-1} A_N - c_N = (-1, -2) \begin{pmatrix} 3 \\ -1 \end{pmatrix} - (-3) = 2 > 0.$$

Für $0 \leq x_3 \leq 2/3$ ist $x_B \geq 0$ zulässig: hier läuft man entlang einer Kante von K . Für $x_3 = 2/3$ erhält man die neue Basislösung $x = (0, 5/3, 2/3)^T$. Der Übergang $x_3 = 0 \rightarrow x_3 = 2/3$ entspricht dem Basistausch

$$B = (1, 2) \rightarrow B' = (3, 2), N = (3) \rightarrow N' = (1).$$

Die Nichtbasisvariable x_3 wird dabei mit der Basisvariable x_1 getauscht: $x_3 = 2/3 - 1/3x_1$. In der neuen Basis B' ergibt sich die Darstellung

$$x_{B'} = \begin{pmatrix} x_3 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2/3 \\ 5/3 \end{pmatrix} - \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix} x_1$$

mit

$$r_{N'} = r'_1 = (-3, -2) \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix} - (-1) = -2/3 < 0.$$

Also ist die gefundene Lösung optimal.

Wir werden nun die Idee des Basistausches formalisieren. Wir nehmen an, dass

$$A_B = I, B = (1, \dots, m) \text{ und } N = (m+1, \dots, n)$$

gelten sowie $x_B = b \geq 0$, d.h. die gegenwärtige Basislösung x sei zulässig. Ist $r_N \leq 0$, so ist x optimal nach Satz 5.1. Sei nun $r_s > 0$ für ein $s \in N$. In die allgemeinen Ausdrücke

$$x_B = A_B^{-1}b - A_B^{-1}A_N x_N, z(x) = z_0 - r_N x_N,$$

setze man ein

$$A_B = I, x_j = 0 \text{ für } j \in N \setminus \{s\}$$

und wir erhalten

$$\begin{aligned} x_B &= b - a^s x_s \\ z(x) &= z_0 - r_s x_s. \end{aligned} \tag{5.4}$$

1. Fall: $a^s \leq 0$.

In diesem Fall ist

$$x_B = b - a^s x_s \geq 0 \text{ für } x_s \rightarrow \infty.$$

Es folgt $z(x) = z_0 - r_s x_s \rightarrow -\infty$ für $x_s \rightarrow \infty$ und somit besitzt das Problem keine endliche Lösung; das Polyeder K ist unbeschränkt.

2. Fall: $a_{is} > 0$ für mindestens ein $i \in \{1, \dots, m\}$. Definiere Index $p \in B$ durch

$$\frac{b_p}{a_{ps}} = \min \left\{ \frac{b_i}{a_{is}} \mid a_{is} > 0, i = 1, \dots, m \right\}.$$

56 | Kapitel 5. Das Simplex Verfahren

Um $x_B \geq 0$ zu garantieren, kann x_s höchstens den Wert

$$x_s = \frac{b_p}{a_{ps}} \text{ annehmen} \Rightarrow x_p = 0.$$

Dies entspricht einem Tausch der Nichtbasisvariablen x_s gegen die Basisvariable x_p (beachte $B = (1, \dots, m)$). Der Wert $a_{ps} > 0$ heißt Pivotelement. Es gilt die Beziehung

$$z(x) = z_0 - r_s \frac{b_p}{a_{ps}}. \quad (5.5)$$

Wir können folgende Fälle unterscheiden:

- (a) $b_p = 0$: Basislösung ist entartet, $z(x) = z_0$
- (b) $b_p > 0$: Wir erhalten eine strikte Verbesserung $z(x) < z_0$

Wir folgern an dieser Stelle ein notwendiges Optimalitätskriterium:

Korollar 5.4. Die Basislösung x mit $x_B \geq 0$ und $x_N = 0$ sei optimal. Gilt $x_B > 0$, dann folgt $r_N \leq 0$.

Der Index $s \in N$ (Pivot Spalte) kann folgendermaßen festgelegt werden:

1. (Regel von Dantzig): Wähle $s \in N$ kleinstmöglich mit der Eigenschaft:

$$r_s = \max_{j \in N} r_j. \quad (5.6)$$

Wähle den kleinsten Index $p \in B$ mit

$$\frac{b_p}{a_{ps}} = \min \left\{ \frac{b_i}{a_{is}} \mid a_{is} > 0, i = 1, \dots, m \right\}. \quad (5.7)$$

2. (Regel von Bland): Wähle den kleinsten Index $s \in N$ mit $r_s > 0$. Wähle den kleinsten Index $p \in B$, unter allen Indizes mit $\frac{b_p}{a_{ps}} = \min \left\{ \frac{b_i}{a_{is}} \mid a_{is} > 0, i = 1, \dots, m \right\}$.

Wir beschreiben nun den Basistausch mittels der Gauss-Jordan Elimination. Nach Voraussetzung

$$A_B = I, B = (1, \dots, m), N = (m+1, \dots, n),$$

gilt

$$x_i = b_i - \sum_{j \in N} a_{ij} x_j, \quad i \in B. \quad (5.8)$$

Die p -te Gleichung kann wegen $a_{ps} > 0$ nach x_s aufgelöst werden:

$$x_p = b_p - \sum_{j \in N} a_{pj} x_j = b_p - \sum_{j \in N, j \neq s} a_{pj} x_j - a_{ps} x_s$$

Auflösen nach x_s ergibt:

$$x_s = \frac{b_p}{a_{ps}} - \sum_{j \in N, j \neq s} \frac{a_{pj}}{a_{ps}} x_j - \frac{1}{a_{ps}} x_p. \quad (5.9)$$

Setze x_s in die übrigen Gleichungen (5.8) mit $i \neq p$ ein:

$$x_i = \underbrace{b_i - a_{is} \frac{b_p}{a_{ps}}}_{=b'_i} - \sum_{j \in N, j \neq s} \left(a_{ij} - a_{is} \frac{a_{pj}}{a_{ps}} \right) x_j + \frac{a_{is}}{a_{ps}} x_p. \quad (5.10)$$

Definiere neue Basis: Übergang

$$B = (1, \dots, m) \rightarrow B' = (1, \dots, p-1, s, p+1, \dots, m).$$

Neue Nichtbasis:

$$N = (m+1, \dots, n) \rightarrow N' = (m+1, \dots, s-1, p, s+1, \dots, n).$$

Die Ausdrücke (5.9) und (5.10) haben die Form

$$x_{B'} = \begin{pmatrix} x_1 \\ \vdots \\ x_{p-1} \\ x_s \\ x_{p+1} \\ \vdots \\ x_m \end{pmatrix} = b' - A'_{N'} x_{N'},$$

und sind somit von der Gestalt (5.2). Die Elemente der Matrix

b'	$A'_{N'}$
------	-----------

sind gegeben durch

Pivotelement: reziproker Wert:

$$a'_{ps} = \frac{1}{a_{ps}}$$

Übrige Zeile p: teile durch Pivotelement:

$$a'_{pj} = \frac{a_{pj}}{a_{ps}}, j \neq s$$

$$b'_p = \frac{b_p}{a_{ps}}$$

Übrige Spalte s: teile durch negatives Pivotelement:

$$a'_{is} = -\frac{a_{is}}{a_{ps}}, i \neq p$$

Sonstige Elemente: subtrahiere das a_{is} -fache der neuen Zeile p von i-ter Zeile:

$$a'_{ij} = a_{ij} - a_{is} \frac{a_{pj}}{a_{ps}} = a_{ij} - a_{is} a'_{pj}, i \neq p, j \neq s$$

$$b'_i = b_i - a_{is} \frac{b_p}{a_{ps}} = b_i - a_{is} b'_p, i \neq p$$

Wir betrachten ein Beispiel.

	b	x_1	x_2	x_3	x_4	x_5	x_6
x_1	5	1	0	0	①	1	-1
x_2	3	0	1	0	2	-3	1
x_3	-1	0	0	1	-1	2	-1

Beachte, dass die aktuelle Basislösung gar nicht zulässig ist. Tausche Nichtbasis-variable x_4 gegen Basisvariable x_1 :

	b'	x_1	x_2	x_3	x_4	x_5	x_6
x_4	5	1	0	0	1	1	-1
x_2	-7	-2	1	0	0	-5	3
x_3	4	1	0	1	0	3	-2

Das Beispiel zeigt nur die algebraische Pivotoperation. Schon die Ausgangslösung ist unzulässig; die Zulässigkeitsgarantie des primalen Simplex-Verfahrens ist daher nicht anwendbar. Nach dem Pivot ist $x_2 = -7$.

Da die Einheitsmatrix $A_B = I$ (bzw. allgemein Permutationsmatrix) keine zusätzliche Information beinhaltet (bei Kenntnis von B, N, A_N, b) werden wir im Folgenden nur das

verkürzte Tableau betrachten.

	b'	x_1	x_5	x_6
x_4	5	1	1	-1
x_2	-7	-2	-5	3
x_3	4	1	3	-2

Wir überprüfen: z.B.:

$$b'_2 = 3 - 2 \cdot 5 = -7$$

$$b'_3 = -1 - (-1) \cdot 5 = 4$$

$$x_{B'} = \begin{pmatrix} x_4 \\ x_2 \\ x_3 \end{pmatrix} = b' = \begin{pmatrix} 5 \\ -7 \\ 4 \end{pmatrix}$$

$$B' = (4, 2, 3)$$

5.3 Zielfunktion und Reduzierte Kosten

Durch Elimination des Ausdrucks (5.9) für x_s folgt

$$\begin{aligned} z(x) &= z_0 - \sum_{j \in N} r_j x_j \\ &= z_0 - \sum_{j \in N, j \neq s} r_j x_j - r_s \left(\frac{b_p}{a_{ps}} - \sum_{j \in N, j \neq s} \frac{a_{pj}}{a_{ps}} x_j - \frac{1}{a_{ps}} x_p \right) \\ &= z_0 - r_s \frac{b_p}{a_{ps}} - \sum_{j \in N, j \neq s} \left(r_j - r_s \frac{a_{pj}}{a_{ps}} \right) x_j + \frac{r_s}{a_{ps}} x_p \\ &=: z'_0 - \sum_{j \in N'} r'_j x_j \end{aligned}$$

mit

$$\begin{aligned} z'_0 &:= z_0 - r_s \frac{b_p}{a_{ps}} \\ r'_p &:= -\frac{r_s}{a_{ps}} \\ r'_j &:= r_j - r_s \frac{a_{pj}}{a_{ps}} = r_j - r_s a'_{pj}, \quad j \in N \setminus \{s\} = N' \setminus \{p\}. \end{aligned} \tag{5.11}$$

Die Zeile (z_0, r_N) wird also wie die Zeilen in

b	A_N
---	-------

transformiert. Damit kann die Pivotoperation des Basis-Tausches an dem **erweiterten Simplex-Tableau** durchgeführt werden:

$$\begin{array}{c}
 \begin{array}{cc}
 & x_j, j \in N \\
 z(x) & \begin{array}{|c|c|}
 \hline
 z_0 & r_N \\
 \hline
 \end{array} \\
 x_i, i \in B & \begin{array}{|c|c|}
 \hline
 b & A_N \\
 \hline
 \end{array}
 \end{array}
 \end{array} \quad (5.12)$$

Abbildung 5.1: Erweitertes Simplex-Tableau.

Wir geben ein Beispiel.

Beispiel 5.5.

$$\begin{aligned}
 &\text{minimiere } z(x_1, x_2, x_3) = -x_1 - 2x_2 - 3x_3 \\
 &\text{unter d.N.} \\
 &2x_1 + x_2 + 5x_3 = 5 \\
 &x_1 + 2x_2 + x_3 = 4 \\
 &x_1 \geq 0, x_2 \geq 0, x_3 \geq 0.
 \end{aligned}$$

Für $B = (1, 2)$, $N = (3)$ kann das LGS überführt werden in das LGS mit $A_B = I$:

$$\begin{array}{|c|c|}
 \hline
 b & A_N \\
 \hline
 \end{array} = \begin{array}{|c|c|}
 \hline
 2 & 3 \\
 1 & -1 \\
 \hline
 \end{array}, \quad x_B = b = \begin{pmatrix} 2 \\ 1 \end{pmatrix} > 0$$

$$c = (-1, -2, -3)$$

$$z_0 = c_B b = (-1, -2) \begin{pmatrix} 2 \\ 1 \end{pmatrix} = -4$$

$$r_3 = c_B A_B^{-1} A_N - c_3 = 2 > 0.$$

Wir erhalten das folgende Simplex-Tableau:

Tausche x_3 gegen x_1 : $r'_1 = -2/3 < 0$, also ist die Basislösung $x_1 = 0, x_2 = 5/3, x_3 = 2/3$ optimal.

Für den Algorithmus verwenden wir die Tableau-Matrix aus Abbildung 5.1.

	x_N	
$z(x), r_N$	-4	2
x_1	2	3
x_2	1	-1

	x_N	
$z(x), r_N$	-16/3	-2/3
x_3	2/3	1/3
x_2	5/3	1/3

$$T = (t_{ij}) = \begin{array}{|c|c|} \hline z_0 & r_N \\ \hline b & A_N \\ \hline \end{array} \quad \begin{array}{l} 0 \leq i \leq m \\ 0 \leq j \leq n - m \end{array} \quad (5.13)$$

Nun können wir den Simplex-Algorithmus formal hinschreiben (hier mit der Auswahlregel von Dantzig).

1. **Start:** Sei B eine Basis mit $A_B = I$. Die zugehörige Basis-Lösung sei zulässig, d.h. $x_B = b \geq 0$. Berechne mit N die Tableau-Matrix $T = (t_{ij})$ in (5.13). Bezeichne mit $B(i)$ den zur i -ten Zeile von T gehörenden Basisindex für $1 \leq i \leq m$. Analog dazu bezeichne mit $N(j)$ den zur j -ten Spalte von T gehörenden Nichtbasisindex für $1 \leq j \leq n - m$.
2. Ist ein $t_{0s} > 0$ ($1 \leq s \leq n - m$), so gehe zu 3. Andernfalls ist die Basis-Lösung optimal. Setze

$$\begin{aligned} x_{B(i)} &:= t_{i0}, 1 \leq i \leq m \\ x_{N(j)} &:= 0, 1 \leq j \leq n - m \\ z &:= t_{00} \end{aligned}$$

Stopp.

3. **Bestimmung der Austausch-Spalte:** Wähle Index $1 \leq s \leq n - m$ (mit $N(s)$ kleinstmöglich) mit

$$t_{0s} = \max_{1 \leq j \leq n-m} t_{0j}.$$

4. Sind alle $t_{is} \leq 0$, $1 \leq i \leq m$, so gibt es keine endliche Lösung. **Stopp.**
Gibt es ein $t_{is} > 0$, so gehe zu Schritt 5.
5. **Bestimmung der Austausch-Zeile:** Wähle Index $1 \leq p \leq m$ (mit $B(p)$ kleinstmöglich) mit

$$\frac{t_{p0}}{t_{ps}} = \min \left\{ \frac{t_{i0}}{t_{is}} \mid t_{is} > 0, i = 1, \dots, m \right\}.$$

Gehe zu Schritt 6.

6. Vertausche das s -te Element von N mit dem p -ten Element von B :

$$(B(p), N(s)) \leftarrow (N(s), B(p)) \quad \text{gleichzeitig.}$$

Führe eine Pivotoperation mit Pivotelement $t_{ps} > 0$ durch, also setze:

- Pivotelement: $t'_{ps} := 1/t_{ps}$
- übrige Pivotzeile: $t'_{pj} := \frac{t_{pj}}{t_{ps}}, j = 0, 1, \dots, n - m, j \neq s$
- übrige Pivotspalte: $t'_{is} := -\frac{t_{is}}{t_{ps}}, i = 0, 1, \dots, m, i \neq p$
- sonstige Elemente: $t'_{ij} := t_{ij} - t_{is} \frac{t_{pj}}{t_{ps}}, i \neq p, j \neq s$
- Setze $t_{ij} := t'_{ij}$ und gehe zu Schritt 2.

Abbildung 5.2: Formaler Ablauf des Simplex-Verfahrens.

5.4 Anwendung auf LPs mit Ungleichungen

Das LP

$$\min\{cx \mid Ax \leq b, x \geq 0\}, x \in \mathbb{R}^n, b \in \mathbb{R}^m, b \geq 0$$

ist äquivalent zu

$$\min\{cx + 0 \cdot y \mid Ax + y = b, x \geq 0, y \geq 0\}, \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n+m}, b \in \mathbb{R}^m$$

Start: $B = (n+1, \dots, n+m), N = (1, \dots, n)$

Basis-Lösung: $x = 0, y = b \geq 0, z_0 = 0, r_N = -c$. Das zugehörige $(m+1) \times (n+1)$ dimensionale Simplex-Tableau lautet

z	0	$-c$
y	b	A

Als Beispiel lösen wir das Problem aus Abschnitt 3.1.1. Die folgenden Tableaus behandeln das Produktionsproblem als Minimierung des negativen Gewinns, also $\min(-16x_1 - 32x_2)$. Die erste Zeile enthält daher negative Gewinnwerte. Für den dargestellten Weg wählen wir zunächst die verbessernde Variable x_1 ; die Regel von Dantzig würde x_2 wählen.

$$\text{maximiere } z(x_1, x_2) = 16x_1 + 32x_2$$

unter d.N.

$$20x_1 + 10x_2 \leq 8000$$

$$4x_1 + 5x_2 \leq 2000$$

$$6x_1 + 15x_2 \leq 4500$$

$$x_1 \geq 0$$

$$x_2 \geq 0$$

Der geometrische Übergang der Ecken $A \rightarrow B \rightarrow C \rightarrow D$ wird rechnerisch in Tableau-Form nachvollzogen.

		x_1	x_2
z	0	16	32
y_1	8000	20	10
y_2	2000	4	5
y_3	4500	6	15

64 | Kapitel 5. Das Simplex Verfahren

Wir wählen das Element 20 als Pivotelement und tauschen x_1 mit y_1 . Dies ist der Übergang $A \rightarrow B$.

		y_1	x_2
z	-6400	$-4/5$	24
x_1	400	$1/20$	$1/2$
y_2	400	$-1/5$	(3)
y_3	2100	$-3/10$	12

Nun tauschen wir x_2 mit y_2 : Dies entspricht dem Übergang $B \rightarrow C$.

		y_1	y_2
z	-9600	$4/5$	-8
x_1	$1000/3$	$1/12$	$-1/6$
x_2	$400/3$	$-1/15$	$1/3$
y_3	500	($1/2$)	-4

Tausche y_1 mit y_3 : Dies entspricht dem Übergang $C \rightarrow D$.

		y_3	y_2
z	-10400	$-8/5$	$-8/5$
x_1	250	$-1/6$	$1/2$
x_2	200	$2/15$	$-1/5$
y_1	1000	2	-8

Hier ist nun $r_N = (-8/5, -8/5) < 0$, also ist $x_1 = 250, x_2 = 200, z_{\min} = -10400$, d.h. $z_{\max} = 10400$, die optimale Lösung.

Wir zeigen nun dass für nichtentartete Probleme, der Simplexalgorithmus stets terminiert.

Satz 5.6. Falls alle bei der Durchführung des Simplex-Algorithmus berechneten Basis-Lösungen nichtentartet sind, so terminiert der Algorithmus.

Beweis. Bei jedem Pivotschritt j erhalten wir eine Darstellung des neuen Zielfunktions-

werts

$$z_0^j = z_0^{j-1} - r_s^{j-1} \frac{b_p^{j-1}}{a_{ps}^{j-1}}.$$

Da für jeden Schritt j nach Annahme $b_p^j > 0$ gilt, erhalten wir eine streng monotone Folge

$$z_0^1 > z_0^2 > \dots$$

Somit wird keine Basis-Lösung mehrfach besucht und da es nur endlich viele Basis-Lösungen gibt, terminiert der Algorithmus. □

5.5 Kreiseln des Simplex Verfahrens und Lexikographische Auswahlregel

Bei einer Pivotoperation mit $b_p = 0$ ändert sich die Zielfunktion nicht, d.h.,

$$\begin{aligned} z(x) &= z_0 - r_s \frac{b_p}{a_{ps}} = t_{00} - t_{0s} \frac{t_{p0}}{t_{ps}} \\ &= z_0 \end{aligned}$$

Beispiele zeigen, dass bei entarteten Basis-Lösungen der Simplex-Algorithmus (mit der “falschen” Pivot-Regel) nicht endlich ist, sondern ins Kreiseln geraten kann (siehe Übung). Zur Vermeidung des Kreisels werden wir nun den lexikographischen Simplex-Algorithmus kennenlernen, der auf der lexikographischen Auswahl der Zeilen beruht.

Definition 5.7. Sei $k \in \mathbb{N}$. Dann ist auf $\mathbb{R}^k$ durch folgende Vorschrift

$$x \succeq y :\Leftrightarrow x = y \text{ oder } x_i > y_i \text{ für } i = \min\{s \mid x_s \neq y_s\}$$

eine totale Ordnung erklärt, die mit der Addition auf $\mathbb{R}^k$ verträglich ist. Man nennt diese die lexikographische Ordnung.

Ein Vektor $x \in \mathbb{R}^k$ heißt **lexikographisch positiv**, wenn $x \succ 0$, also $x \succeq 0$ und $x \neq 0$, gilt. Wir betrachten nun das sogenannte volle Tableau in Standardform, welches bei der ersten Iteration wie folgt aussieht:

$$T = (t_{ij})_{\substack{i=0,\dots,m \\ j=0,\dots,n}} = \begin{array}{|c|c|c|} \hline z_0 & 0 & r_N \\ \hline b & A_B = I & A_N \\ \hline \end{array}$$

Die Zeilen des Tableaus $t_i, i = 1, \dots, m$ sind dann also Vektoren im $\mathbb{R}^{n+1}$. Für eine Basis B bezeichnen wir wie üblich mit $B(i)$ den i -ten Eintrag der aktuellen Basis B (siehe vorangegangene Beispiele). Zu Beginn ist also $B = (1, \dots, m)$ und $N = (m+1, \dots, n)$. Die Spalten des vollständigen Tableaus behalten während aller Pivotoperationen ihre anfängliche Reihenfolge. Die Basisindizes werden aktualisiert, aber die Spalten werden nicht erneut nach Basis und Nichtbasis sortiert.

Definition 5.8 (Lexikographische Auswahlregel (LEX)). Es liege die Basis B vor und das zugehörige volle Tableau sei T .

- Wähle einen beliebigen Spaltenindex $s \in N$ mit $r_s = t_{0s} > 0$.
- Betrachte zur Zeilenauswahl die Menge $I = \{i \in \{1, \dots, m\} | t_{is} > 0\}$. Ist $I = \emptyset$, ist das LP unbeschränkt; der Algorithmus stoppt. Andernfalls Nun sei $p \in I$ so gewählt, dass gilt

$$\frac{t_p}{t_{ps}} = \text{lexmin} \left\{ \frac{t_i}{t_{is}} \mid i \in I \right\}.$$

Das so gewählte lexikographische Minimum erfüllt

$$\frac{t_i}{t_{is}} \succ \frac{t_p}{t_{ps}} \text{ für alle } i \in I \setminus \{p\}.$$

Dann ist $B(p)$ die Ausgangsvariable, welche die aktuelle Basis B verlässt, und s wird zur Eingangsvariable, welche in die neue Basis B' aufgenommen wird, d.h., $B'(p) = s$.

Man beachte, dass nach Wahl der Pivotzeile p insbesondere gelten muss, dass

$$\frac{t_{p0}}{t_{ps}} = \min \left\{ \frac{t_{i0}}{t_{is}} \mid i \in I \right\}.$$

Somit führt der Pivotschritt wiederum zu einer zulässigen Basislösung.

Satz 5.9. Beim Start des Simplex liege ein Tableau vor, dessen Zeilen (ausser möglicherweise der reduzierten Kosten) sämtlich lexikographisch positiv sind. Dann terminiert der Simplex-Algorithmus unter Verwendung von LEX.

Beweis. Wir sehen zunächst recht leicht, dass die Voraussetzung des Satzes leicht erfüllbar ist, indem wir durch Gauss-Jordan Transformation das Tableau in die Gestalt $(b|I|A_N)$ bringen (erste Zeile weggelassen). Wegen der Zulässigkeit $x_B \geq 0$ muss die erste Spalte nur Einträge grösser gleich 0 haben (erste Zeile wieder weggelassen). Da die erste Teilmatrix nun die Einheitsmatrix ist, sind alle Zeilenvektoren lexikographisch positiv.

Wir zeigen nun induktiv dass bei iterativer Ausführung von LEX, diese Eigenschaft erhalten bleibt. Wir betrachten dazu eine Basis B mit der gewünschten Eigenschaft. Es seien nun s und p gemäß LEX gewählt. Die neue zu B benachbarte Basis sei mit B' bezeichnet. Für die p -te Zeile t'_p des neuen Tableaus T' gilt nun

$$t'_p = \frac{t_p}{t_{ps}} \succ 0,$$

da nach Voraussetzung $t_p \succ 0$ und zudem $t_{ps} > 0$ gilt. Ist nun $i \in \{1, \dots, m\}$, $i \neq p$, so gilt

für die i -te Zeile t'_i des Tableaus T' :

$$t'_i = t_i - \frac{t_{is}}{t_{ps}} t_p.$$

Ist $t_{is} > 0$, so ist gemäß Regel LEX, die zur Wahl von p geführt hat,

$$t'_i = t_{is} \left(\frac{t_i}{t_{is}} - \frac{t_p}{t_{ps}} \right) \succ 0.$$

Ist andererseits $t_{is} \leq 0$, so folgt

$$t'_i = t_i + \frac{|t_{is}|}{t_{ps}} t_p \geq t_i \succ 0.$$

Daraus folgt insgesamt, dass die Zeilen lexikographisch positiv bleiben.

Wir betrachten schließlich die ersten Zeilen t_0 und t'_0 der Tableaus, welche die reduzierten Kosten enthalten. Es gilt

$$t'_0 = t_0 - \frac{t_{0s}}{t_{ps}} t_p.$$

Wegen $t_{0s} = r_s > 0$ gilt also

$$t_0 = t'_0 + \frac{t_{0s}}{t_{ps}} t_p \succ t'_0.$$

Dies zeigt, dass bei jedem Pivotschritt die 0-Zeile des Tableaus strikt lexikographisch fällt und somit sich kein Tableau während der Ausführung des Algorithmus wiederholen kann. $\square$

Es gibt nur endlich viele Basen und damit endlich viele zugehörige Tableaus. Da die erste Zeile bei jedem LEX-Schritt strikt lexikographisch fällt, kann keine Basis wiederholt werden; dies beweist die endliche Terminierung.

5.6 Berechnung einer zulässigen Basislösung - Die Zweiphasen-Methode

In den bisherigen Betrachtungen sind wir immer von einer bereits gegebenen zulässigen Basislösung ausgegangen und haben durch die Anwendung des Simplex-Algorithmus eine optimale Ecke gesucht (falls sie existiert). Jedoch ist es überhaupt nicht klar, ob es für ein gegebenes LP überhaupt eine zulässige Ecke gibt und falls doch, wie man diese findet. Man kann natürlich alle der $\binom{n}{m}$ möglichen Kombinationen von Basisvariablen ausprobieren, diese Vorgehensweise ist aber nicht effizient.

Wir betrachten wieder ein LP in Standardform:

$$\begin{aligned} \text{minimiere } z(x) &= cx \\ \text{unter d.N.} \\ Ax &= b \\ x &\geq 0. \end{aligned} \tag{5.14}$$

68 | Kapitel 5. Das Simplex Verfahren

Zusätzlich setzen wir ohne Einschränkung voraus, dass $b \geq 0$ ist. Falls es nämlich Zeilen gibt, in denen der Eintrag von b negativ ist, multiplizieren wir diese mit -1 und ändern damit nichts am Polyeder.

Interessanterweise gibt es ein relativ leicht zu lösendes Hilfsproblem, welches uns bei Existenz einer zulässigen Basislösung eine solche liefert. Dieses sieht wie folgt aus:

$$\begin{aligned} \text{minimiere } z(x, y) &= \sum_{i=n+1}^{n+m} y_i \\ \text{unter d.N.} & \\ Ax + y &= b \\ x, y &\geq 0. \end{aligned} \quad (5.15)$$

Man führt also in jeder der m Zeilen eine zusätzliche künstliche Variable y_i ein, wobei $i \in \{n+1, \dots, n+m\}$. Nun muss genau eine der beiden folgenden Möglichkeiten eintreten:

1. Das Hilfsproblem besitzt eine Optimallösung mit Zielfunktionswert 0. Dann ist die zugehörige Basislösung zulässig für das System $Ax = b, x \geq 0$.
2. Das Hilfsproblem besitzt eine Optimallösung mit Zielfunktionswert > 0 . Dann besitzt das System $Ax = b, x \geq 0$ keine zulässige Lösung.

Die Folgerungen sind offensichtlich. Ist nämlich der Zielfunktionswert gleich 0, so sind alle $y_i = 0$, da $y \geq 0$ ist. Dann ist aber auch $Ax = b$ mit $x \geq 0$. Ist jedoch der optimale Zielfunktionswert > 0 , so kann $Ax = b, x \geq 0$ keine Lösung besitzen, da dafür ja $y = 0$ gelten müsste.

Das Hilfsproblem lässt sich nun mit dem Simplex-Algorithmus lösen, wobei man trivialerweise als Basis die künstlichen Variablen $y_i, i \in \{n+1, \dots, n+m\}$ nutzen kann (da die rechte Seite nach Voraussetzung ≥ 0 ist). Es ist also

$$B = (n+1, \dots, n+m), N = (1, \dots, n), c_B = (1, 1, \dots, 1) \in \mathbb{R}^m \text{ und } c_N = 0 \in \mathbb{R}^n.$$

Daraus erhalten wir

$$\begin{aligned} z_0 = c_B b &= \sum_{i=1}^m b_i \geq 0, \\ r_N &= c_N A_N, \end{aligned} \quad (5.16)$$

wobei sich $r_j = \sum_{i=1}^m a_{ij}, j = 1, \dots, n$ aus der Summe der Einträge in den Spalten von A_N berechnet. Das Starttableau lautet dann

$$\begin{array}{cc|cc} & & & x \\ z(x, y), r_N & z_0 & r_N \\ y & b & A_N \end{array} \quad (5.17)$$

Die Zweiphasen-Methode zur Lösung des LP (5.14) besteht also aus

Phase I: Löse das Hilfsproblem (5.15), etwa mit LEX. Ist sein Optimalwert positiv, ist das Ausgangsproblem unzulässig und das Verfahren stoppt. Ist er null, sind alle künstlichen Variablen null; einige können trotzdem noch zur Basis gehören. Tausche jede solche Variable mit einer ursprünglichen Nichtbasisvariablen aus, deren Koeffizient in der betreffenden Tableauzeile nicht null ist. Wegen der rechten Seite null bleibt die Basislösung dabei unverändert und zulässig. Gibt es keinen solchen Koeffizienten, entferne die redundante Gleichung. Danach liegt eine zulässige Basis des Ausgangsproblems vor.

Phase II: Berechne Starttableau zur Lösung von LP (5.14) durch

- Streichung der zu $y_{n+1}, \dots, y_{n+m}$ gehörigen Spalten,
- Berechnung der neuen Werte z_0 und r_N .

Beispiel 5.10. Wir wenden die Zweiphasen-Methode auf folgendes Problem an:

$$\begin{array}{llll} \min & 4x_1 & + & x_2 & + & x_3 \\ \text{u.d.N.} & 2x_1 & + & x_2 & + & 2x_3 & = & 4 \\ & 3x_1 & + & 3x_2 & + & x_3 & = & 3 \\ & x_1 & , & x_2 & , & x_3 & \geq & 0 \end{array}$$

Das Hilfsproblem sieht dann folgendermaßen aus:

$$\begin{array}{llllll} \min & & & & y_4 & + & y_5 \\ \text{u.d.N.} & 2x_1 & + & x_2 & + & 2x_3 & + & y_4 & & = & 4 \\ & 3x_1 & + & 3x_2 & + & x_3 & & & + & y_5 & = & 3 \\ & x_1 & , & x_2 & , & x_3 & , & y_4 & , & y_5 & \geq & 0 \end{array}$$

Wir starten mit der Basis $B = (4, 5)$ und zugehöriger Nichtbasis $N = (1, 2, 3)$. Zusätzlich berechnen wir

$$z_0 = b_1 + b_2 = 4 + 3 = 7 \quad \text{und} \quad r_N = \left(\sum_{i=1}^2 a_{ij} \right)_{j=1,2,3} = (5, 4, 3).$$

70 | Kapitel 5. Das Simplex Verfahren

Damit erhalten wir das Starttableau

		x_1	x_2	x_3
$z(x, y), r_N$	7	5	4	3
y_4	4	2	1	2
y_5	3	3	3	1

Mit Dantzig's Auswahlregel erhalten wir noch die beiden folgenden Tableaus:

		y_5	x_2	x_3
$z(x, y), r_N$	2	-5/3	-1	4/3
y_4	2	-2/3	-1	4/3
x_1	1	1/3	1	1/3

		y_5	x_2	y_4
$z(x, y), r_N$	0	-1	0	-1
x_3	3/2	-1/2	-3/4	3/4
x_1	1/2	1/2	5/4	-1/4

Hier erkennen wir, dass $r_N \leq 0$ und $z(x, y) = 0$ ist, wobei y_4 und y_5 Nichtbasis-Variablen sind. Damit ist eine zulässige Lösung von $Ax = b, x \geq 0$ gegeben durch die Basis $B = (3, 1)$ mit den Belegungen

$$x_3 = 3/2, \quad x_1 = 1/2, \quad x_2 = 0.$$

An dieser Stelle wollen wir festhalten, dass durch das Lösen des Hilfsproblems die zugehörige Basisspalten im transformierten Tableau schon als Einheitsmatrix vorliegt und wir deshalb das Starttableau für Phase II relativ leicht aufstellen können.

Wollen wir nun Phase II ausführen, so streichen wir die zu y_4 und y_5 gehörenden Spalten aus dem Tableau. Es ist $c = (4, 1, 1)$. Bezüglich $B = (3, 1)$ und $N = (2)$ ist $c_B = (1, 4)$ und $b = \begin{pmatrix} 3/2 \\ 1/2 \end{pmatrix}$ (siehe Tableau). Berechnen wir nun noch die neuen Werte für die Zielfunktion und reduzierte Kosten, so erhalten wir

$$z_0 = c_B b = (1, 4) \begin{pmatrix} 3/2 \\ 1/2 \end{pmatrix} = 3/2 + 4/2 = 7/2 \text{ und}$$

$$r_N = r_2 = c_B a^2 - c_2 = (1, 4) \begin{pmatrix} -3/4 \\ 5/4 \end{pmatrix} - 1 = 13/4,$$

wobei a^2 aus dem Tableau entnommen werden muss.

Das Starttableau für Phase II hat also folgende Form:

	x_2	
$z(x), r_N$	$7/2$	$13/4$
x_3	$3/2$	$-3/4$
x_1	$1/2$	$5/4$

Nach einem Pivotschritt erhalten wir

	x_1	
$z(x), r_N$	$11/5$	$-13/5$
x_3	$9/5$	$3/5$
x_2	$2/5$	$4/5$

Da nun die reduzierten Kosten $r_1 = -13/5 < 0$ sind, ist die optimale Lösung des Ausgangsproblems

$$x_1 = 0, \quad x_2 = 2/5, \quad x_3 = 9/5 \quad \text{mit Zielfunktionswert } 11/5.$$

Mit der Zweiphasen-Methode lassen sich nicht nur zulässige Basislösungen für lineare Programme in Standardform finden, sondern auch für Probleme, in denen zusätzlich Ungleichungen enthalten sind. Gegeben sei also das Problem

$$\begin{aligned} &\text{minimiere } z(x) = cx \\ &\text{unter d.N.} \\ &a_i x = b_i, \quad i = 1, \dots, k \\ &a_i x \leq b_i, \quad i = k+1, \dots, m \\ &x \geq 0. \end{aligned} \tag{5.18}$$

Hier benötigen wir also künstliche Variablen $y_{n+1}, \dots, y_{n+k}$ für die ersten k Gleichungen sowie Schlupfvariablen $x_{n+k+1}, \dots, x_{n+m}$ um die Ungleichungen in Gleichungsform zu bringen. Da wir wie oben von $b \geq 0$ ausgehen, müssen wir in Phase I folgendes Hilfsproblem lösen (da die Schlupfvariablen ≥ 0 sein dürfen, kommen diese in der Zielfunktion nicht vor):

$$\begin{aligned}
& \text{minimiere } \sum_{i=n+1}^{n+k} y_i \\
& \text{unter d.N.} \\
& a_i x + y_{n+i} = b_i, \quad i = 1, \dots, k \\
& a_i x + x_{n+i} = b_i \quad i = k+1, \dots, m \\
& x, y \geq 0.
\end{aligned} \tag{5.19}$$

Das Starttableau ergibt sich wie in (5.17), wobei wir

$$z_0 = \sum_{i=1}^k b_i \text{ und } r_j = \sum_{i=1}^k a_{ij}, j = 1, \dots, n$$

wählen müssen. Bei Hilfoptimalwert null entfernen wir die künstlichen Basisvariablen gegebenenfalls durch die oben beschriebenen zusätzlichen Pivotschritte. Anschließend lässt sich dann wieder Phase II anwenden.

Bei diesen Betrachtungen haben wir wieder $b \geq 0$ vorausgesetzt. Dies können wir jedoch nicht ohne Einschränkung machen, da wir die Ungleichungen $a_i x \leq b_i$ nicht einfach so mit -1 multiplizieren können, sodass die rechte Seite nichtnegativ ist. Denn dann ist für die entsprechenden Zeilen $a_i x \geq b_i$ und die eingeführten Schlupfvariablen müssten von der linken Seite abgezogen werden, um eine Gleichung zu bekommen. In diesem Fall bilden die Schlupfvariablen keine zulässige Basis. Wir betrachten deshalb nun ein Problem mit „gemischten Ungleichungen“:

$$\begin{aligned}
& \text{minimiere } z(x) = cx \\
& \text{unter d.N.} \\
& a_i x \geq b_i, \quad i = 1, \dots, k \\
& a_i x \leq b_i, \quad i = k+1, \dots, m \\
& x \geq 0.
\end{aligned} \tag{5.20}$$

Negative rechte Seiten werden durch Multiplikation der betreffenden Zeile mit -1 beseitigt; dabei kehrt sich die Ungleichungsrichtung um. Anschließend werden die Zeilen neu gruppiert und k entsprechend angepasst, so dass $b \geq 0$ gilt.

Hier können wir nun wieder davon ausgehen, dass $b \geq 0$ ist. Durch Einführung von Schlupfvariablen $x_{n+1}, \dots, x_{n+m}$ erhalten wir ein Problem in Standardform:

$$\begin{aligned}
& \text{minimiere } z(x) = cx \\
& \text{unter d.N.} \\
& a_i x - x_{n+i} = b_i, \quad i = 1, \dots, k \\
& a_i x + x_{n+i} = b_i, \quad i = k+1, \dots, m \\
& x_i \geq 0, \quad i = 1, \dots, n+m.
\end{aligned} \tag{5.21}$$

5.6. Berechnung einer zulässigen Basislösung - Die Zweiphasen-Methode | 73

Um nun eine zulässige Basislösung zu erhalten, müssen wir wieder künstliche Variablen $y_{n+m+1}, \dots, y_{n+m+k}$ für die ersten k Gleichungen einführen (da hier die Schlupfvariablen negativ auftreten).

Das zugehörige Hilfsproblem lautet dann

$$\begin{aligned}
 & \text{minimiere} \quad \sum_{i=n+m+1}^{n+m+k} y_i \\
 & \text{unter d.N.} \\
 & a_i x - x_{n+i} + y_{n+m+i} = b_i, \quad i = 1, \dots, k \\
 & a_i x + x_{n+i} = b_i, \quad i = k+1, \dots, m \\
 & x_i \geq 0, \quad i = 1, \dots, n+m \\
 & y_i \geq 0, \quad i = n+m+1, \dots, n+m+k.
 \end{aligned} \tag{5.22}$$

Zur Bestimmung eines Starttableaus für dieses Hilfsproblem sortieren wir die $n+m+k$ Variablen in der Reihenfolge

$$\underbrace{x_1, \dots, x_n, x_{n+1}, \dots, x_{n+k}}_N, \underbrace{y_{n+m+1}, \dots, y_{n+m+k}, x_{n+k+1}, \dots, x_{n+m}}_B.$$

Damit ist dann

$$\begin{aligned}
 c_B &= (\underbrace{1, 1, \dots, 1}_k, \underbrace{0, 0, \dots, 0}_{m-k}) \in \mathbb{R}^m, \\
 z_0 &= c_B b = \sum_{i=1}^k b_i, \\
 r_j &= c_B a^j - c_j = \begin{cases} \sum_{i=1}^k a_{ij}, & j = 1, \dots, n \\ -1, & j = n+1, \dots, n+k \end{cases}
 \end{aligned}$$

Abschließend können wir damit das Starttableau für Phase I aufstellen zu

	$x_1 \cdots x_n$		$x_{n+1} \cdots x_{n+k}$
$z(x), r_N$	$\sum_{i=1}^k b_i$	$\sum_{i=1}^k a_{ij}$	$-1 \cdots -1$
y_{n+m+1}	b_1	a_1	$-I_k$
$\vdots$	$\vdots$	$\vdots$	
y_{n+m+k}	b_k	a_k	
x_{n+k+1}	b_{k+1}	a_{k+1}	0
$\vdots$	$\vdots$	$\vdots$	
x_{n+m}	b_m	a_m	

Bemerkung 5.11. Eine künstliche Basisvariable mit Wert null zeigt Entartung des Hilfsproblems an. Auch nach dem Entfernen aller künstlichen Variablen kann die Basislösung entartet sein. Bei vollem Zeilenrang von A lassen sich alle künstlichen Basisvariablen durch die beschriebenen Pivotschritte entfernen; sonst werden zusätzlich redundante Gleichungen gestrichen.

Kapitel 6

Dualität für Lineare Probleme

Wir betrachten wiederum ein LP in Standardform

$$\begin{aligned} \text{minimiere } z(x) &= cx \\ Ax &= b \\ x &\geq 0, \end{aligned} \tag{P}$$

für $A \in \mathbb{R}^{m \times n}$, $c^T \in \mathbb{R}^n$, $b \in \mathbb{R}^m$. Wir bezeichnen das Polyeder der zulässigen Punkte mit K , also

$$K = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}.$$

Sei $x \in K$ und betrachte einen beliebigen Vektor $\lambda^T \in \mathbb{R}^m$. Wir erhalten folgende äquivalente Darstellungen

$$b = Ax \Leftrightarrow \lambda b = \lambda Ax \quad \forall \lambda^T \in \mathbb{R}^m.$$

Wir erhalten nun eine untere Schranke an den Zielfunktionswert cx durch folgende Ungleichung:

$$\lambda b \leq cx \quad \forall \lambda^T \in \mathbb{R}^m \text{ mit } \lambda A \leq c.$$

Unser Ziel ist nun, die “beste” untere Schranke an cx zu bestimmen. Das führt auf folgendes duale Optimierungsproblem:

$$\begin{aligned} \text{maximiere } z^*(\lambda) &= \lambda b \\ \lambda A &\leq c. \end{aligned} \tag{D}$$

Wir bezeichnen mit K^* das Polyeder für obiges Problem.

$$K^* := \{\lambda^T \in \mathbb{R}^m | \lambda A \leq c\}.$$

Man beachte dass hier λ^T keinerlei Vorzeichenbeschränkung unterliegt.

6.1 Duale Programme

Allgemein kann jedem LP ein duales LP zugeordnet werden und zwischen diesen beiden bestehen wichtige Beziehungen, die wir später behandeln werden. Die Dualitätsbeziehungen können am besten für LP's mit Ungleichungen in symmetrischer Form ausgedrückt werden.

Symmetrische Form:

Primalproblem (P-sym):

minimiere cx

$$Ax \geq b \quad (\text{P-sym})$$

$$x \geq 0,$$

Dualproblem (D-sym):

maximiere λb

$$\lambda A \leq c \quad (\text{D-sym})$$

$$\lambda \geq 0.$$

Dabei ist A eine $m \times n$ Matrix, $x \in \mathbb{R}^n$ Spaltenvector, und $\lambda \in \mathbb{R}^m$ Zeilenvector. Wie vorher werden die jeweiligen zulässigen Mengen mit K und K^* bezeichnet.

Bemerkung 6.1. Das Dualproblem des Dualproblems ist wieder das Primalproblem.
Beweis: Übung.

Wir geben ein Beispiel an.

Primalproblem (P):

minimiere $3x_1 + 4x_2 + 5x_3$

$$x_1 + 2x_2 + 3x_3 \geq 5$$

$$2x_1 + 2x_2 + x_3 \geq 6$$

$$x_i \geq 0, i = 1, 2, 3$$

Dualproblem (P):

maximiere $5\lambda_1 + 6\lambda_2$

$$\lambda_1 + 2\lambda_2 \leq 3$$

$$2\lambda_1 + 2\lambda_2 \leq 4$$

$$3\lambda_1 + \lambda_2 \leq 5$$

$$\lambda_i \geq 0, i = 1, 2.$$

Das LP in Standardform kann auf die symmetrische Form zurückgeführt werden. Die Gleichung

$$Ax = b$$

entspricht

$$Ax \geq b$$

$$-Ax \geq -b.$$

Das Dualproblem für die Variable $w \in \mathbb{R}^{2m}$ ist dann nach (P-sym)

$$\begin{aligned} &\text{maximiere } w \begin{pmatrix} b \\ -b \end{pmatrix} \\ &w \begin{pmatrix} A \\ -A \end{pmatrix} \leq c \\ &w \geq 0. \end{aligned}$$

Setzt man nun $w := (u, v)$, $u, v \in \mathbb{R}^m$, so erhält man

$$\begin{aligned} &\text{maximiere } ub - vb \\ &uA - vA \leq c \\ &u \geq 0, v \geq 0. \end{aligned}$$

Mit $\lambda := u - v$ wird daraus das bereits definierte Dualproblem (D). Zu beachten ist dass die duale Variable $\lambda = u - v$ keinen Vorzeichenbeschränkungen unterliegt.

Wir geben tabellarisch einige Rechenregeln bezüglich der Dualisierung.

Tabelle 6.1: Dualisierung.

Primales LP	Duales LP
(P ₁) $\max cx, Ax \leq b, x \geq 0$	(D ₁) $\min \lambda b, \lambda A \geq c, \lambda \geq 0$
(P ₂) $\min cx, Ax \geq b, x \geq 0$	(D ₂) $\max \lambda b, \lambda A \leq c, \lambda \geq 0$
(P ₃) $\max cx, Ax = b, x \geq 0$	(D ₃) $\min \lambda b, \lambda A \geq c$
(P ₄) $\min cx, Ax = b, x \geq 0$	(D ₄) $\max \lambda b, \lambda A \leq c$
(P ₅) $\max cx, Ax \leq b$	(D ₅) $\min \lambda b, \lambda A = c, \lambda \geq 0$
(P ₆) $\min cx, Ax \geq b$	(D ₆) $\max \lambda b, \lambda A = c, \lambda \geq 0$

Wir geben noch eine allgemeine Form der Dualität an.

Lemma 6.2. Es seien dimensionsverträgliche Matrizen A, B, C, D und Vektoren a, b, c, d gegeben. Zum primalen LP

$$\begin{aligned} &\text{maximiere } cx + dy \\ &Ax + By \leq a \\ &Cx + Dy = b \\ &x \geq 0, \end{aligned} \tag{LP}$$

gehört das duale LP

$$\begin{aligned}
& \text{minimiere } ua + vb \\
& \quad uA + vC \geq c \\
& \quad uB + vD = d \\
& \quad u \geq 0.
\end{aligned} \tag{DP}$$

Die Variablen y und die Multiplikatoren v sind ohne Vorzeichenbeschränkung.

Beweis. Übung. □

Bevor wir nun zum zentralen starken Dualitätssatz kommen, werden wir zunächst ein Beispiel und eine Interpretation von zueinander dualen Programmen geben.

6.2 Duale Programme für Transportprobleme

Wir wiederholen das in Abschnitt 3.1.3 eingeführte Transportproblem.

Primalproblem:

In m Fabriken werden jeweils die Warenmengen a_i , $1 \leq i \leq m$ erzeugt. Diese Waren sollen an n Abnehmer, die jeweils die Warenmengen b_j , $1 \leq j \leq n$ benötigen, so versandt werden, dass die Transportkosten minimiert werden. Bezeichnen wir mit c_{ij} die Transportkosten pro Wareneinheit von Fabrik i zum Abnehmer j , so lautet das zugehörige LP, in dem mit x_{ij} die tatsächlich von i nach j versandte Menge bezeichnet wird:

$$\begin{aligned}
& \text{minimiere } z(x_1, \dots, x_n) = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \\
& \quad \sum_{j=1}^n x_{ij} = a_i, 1 \leq i \leq m, \\
& \quad \sum_{i=1}^m x_{ij} = b_j, 1 \leq j \leq n, \\
& \quad x_{ij} \geq 0, i = 1, \dots, m, j = 1, \dots, n.
\end{aligned} \tag{P}$$

Dualproblem:

Ein Unternehmer meint, den Versand kostengünstiger durchführen zu können. Er bietet dem Erzeuger an, seine Waren am Erzeugungsort zu kaufen und am Bestimmungsort wieder zu verkaufen. Der Preis der Produkte variiert von Ort zu Ort und wird vom Unternehmer im Voraus festgesetzt. Er muss jedoch die Preise so gestalten, dass sein Angebot attraktiv für den Erzeuger ist. Somit muss der Unternehmer Preise $\lambda_1, \dots, \lambda_m$ für die m Produktionsorte und $\mu_1, \dots, \mu_n$ für die n Bestimmungsorte wählen. Um gegenüber herkömmlichen Transportmöglichkeiten konkurrenzfähig zu sein, müssen seine Preise

$$\mu_j - \lambda_i \leq c_{ij}, 1 \leq i \leq m, 1 \leq j \leq n$$

erfüllen. Die Gewinn-Maximierung des Unternehmers führt auf das Problem

$$\begin{aligned} \text{maximiere } z^*(\lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_n) &= \sum_{j=1}^n \mu_j b_j - \sum_{i=1}^m \lambda_i a_i \\ \mu_j - \lambda_i &\leq c_{ij}, \quad 1 \leq i \leq m, 1 \leq j \leq n. \end{aligned} \quad (D)$$

6.3 Der Dualitätssatz

Wir bereiten nun den starken Dualitätssatz der linearen Optimierung vor. Dazu betrachten wir das Eingangs eingeführte Paar von zueinander dualen linearen Programmen

$$\begin{aligned} \text{minimiere } z(x) &= cx \\ Ax &= b \\ x &\geq 0, \end{aligned} \quad (P)$$

$$\begin{aligned} \text{maximiere } z^*(\lambda) &= \lambda b \\ \lambda A &\leq c. \end{aligned} \quad (D)$$

Wir bezeichnen wiederum die zulässigen Mengen der beiden LPs mit K und K^* .

Wir erhalten unmittelbar, den folgenden Zusammenhang zwischen Lösungen von (P) und (D).

| Satz 6.3 (Schwacher Dualitätssatz). Für $x \in K, \lambda^T \in K^*$ gilt $\lambda b \leq cx$.

Beweis. Sei $x \in K$ und $\lambda^T \in K^*$. Wir erhalten

$$b = Ax \Rightarrow \lambda b = \lambda Ax \underset{\lambda^T \in K^*}{\leq} cx.$$

□

Es ergibt sich folgendes wichtiges Korollar.

| Korollar 6.4. Seien $x^* \in K$ und $(\lambda^*)^T \in K^*$ mit $\lambda^* b = cx^*$. Dann sind jeweils x^* und $(\lambda^*)^T$ optimal für die jeweiligen Probleme (P) und (D).

Beweis. Für beliebiges $x \in K$ gilt

$$cx \geq \lambda^* b = cx^*.$$

Desweiteren gilt für beliebiges $\lambda \in K^*$

$$\lambda b \leq cx^* = \lambda^* b.$$

□

Satz 6.5 (Dualitätssatz der linearen Optimierung). Besitzt eines der Probleme (P) oder (D) eine endliche Optimallösung, so besitzt auch das jeweils andere Problem eine endliche Optimallösung und der Optimalwert der beiden Zielfunktionen ist gleich:

$$\min\{cx \mid x \in K\} = \max\{\lambda b \mid \lambda^T \in K^*\}.$$

Beweis. Ohne Einschränkung sei $x^* \in K$ optimal mit $cx^* = z_0$ endlich. Definiere folgende Menge

$$C := \{(r, w) = t(cx - z_0, b - Ax)^T \mid x \geq 0, t > 0\} \subset \mathbb{R}^{m+1}.$$

C ist ein konvexer Kegel mit Spitze im Nullpunkt.

Wir erhalten folgende alternative Darstellung von C :

$$C = \{(r, w) = (cx - tz_0, tb - Ax)^T \mid x \geq 0, t > 0\} \subset \mathbb{R}^{m+1}.$$

(Definiere $\tilde{x} = x/t$). Da $Ax^* = b$ und $cx^* = z_0$ gilt, ändert die Ersetzung $(x, t) \mapsto (x + x^*, t + 1)$ den dargestellten Punkt nicht. Daher kann man auch $t = 0$ zulassen, ohne C zu vergrößern. Somit ist C der von den endlich vielen Vektoren $(c_j, -a^j)^T$ und $(-z_0, b)^T$ erzeugte konvexe Kegel und ist abgeschlossen. Zur Begründung der letzten Aussage: Jede nichtnegative Linearkombination kann durch sukzessive Elimination linearer Abhängigkeiten auf einen linear unabhängigen Träger reduziert werden. Der Kegel ist daher eine endliche Vereinigung von Kegeln mit linear unabhängigen Erzeugern. Jeder dieser Kegel ist abgeschlossen, da die Koeffizienten durch eine stetige Linksinverse bestimmt werden und bei Grenzübergängen nichtnegativ bleiben.

Wir behaupten nun, dass der Punkt $(-1, 0)$ nicht in C liegt, d.h., $(-1, 0) \notin C$. Dies folgt unmittelbar, da $t > 0$ und $z_0 \leq cx$ für alle $x \in K$. Somit können wir den Trennungssatz für konvexe und abgeschlossene Kegel anwenden (vgl. Satz 2.9). Es existiert ein Vektor $(\lambda_0, \lambda)^T \in \mathbb{R}^{m+1} \setminus \{0\}$ mit $\lambda_0 \in \mathbb{R}, \lambda^T \in \mathbb{R}^m$ so dass gilt

$$(\lambda_0, \lambda) \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -\lambda_0 < 0 \leq (\lambda_0, \lambda) \begin{pmatrix} cx - tz_0 \\ tb - Ax \end{pmatrix}$$

für alle $t > 0, x \geq 0$. Es gilt $\lambda_0 > 0$ und ohne Einschränkung können wir $\lambda_0 = 1$ setzen (zu erreichen durch Multiplikation mit $1/\lambda_0$). Indem wir das Skalarprodukt ausrechnen erhalten wir für alle $t > 0, x \geq 0$

$$\begin{aligned} 0 &\leq (cx - tz_0) + \lambda(tb - Ax) \\ \Leftrightarrow tz_0 - cx &\leq \lambda(tb - Ax). \end{aligned} \tag{6.1}$$

Indem wir nun den Limes $t \downarrow 0$ betrachten, erhalten wir

$$cx \geq \lambda Ax \text{ für alle } x \geq 0$$

und somit auch

$$c \geq \lambda A$$

was $\lambda \in K^*$ impliziert. Setzen wir nun in (6.1) die Werte $t = 1, x = 0$ ein, so erhalten wir

$$z_0 \leq \lambda b.$$

Mit dem schwachen Dualitätssatz 6.3 folgt $z_0 = \lambda b$ und nach Korollar 6.4 sind dann auch $x^* \in K$ und $\lambda \in K^*$ optimal. $\square$

Bemerkung 6.6. Der Beweis benutzt nicht die Rangvoraussetzung $\text{Rang}(A) = m$ des Simplexverfahrens. Bei Existenz einer nichtentarteten optimalen Basis-Lösung ist der Beweis deutlich einfacher.

Mit den Sätzen (6.3) und (6.5) können in einfacher Weise die folgenden Existenz- und Optimalitätskriterien gewonnen werden.

Satz 6.7 (Existenzsatz). Die folgenden Aussagen sind äquivalent.

1. (P) und (D) haben beide zulässige Lösungen.
2. (P) hat eine endliche Optimallösung.
3. (P) und (D) haben endliche Optimallösungen, deren Zielfunktionswerte übereinstimmen.
4. (D) hat eine endliche Optimallösung.
5. (P) hat zulässige Lösungen und ist nach unten beschränkt oder (D) hat zulässige Lösungen und ist nach oben beschränkt.

Beweis. (1) $\Rightarrow$ (2) : Mit Satz (6.3) gilt, dass die Zielfunktion von (P) nach unten beschränkt ist. Somit folgt (da (P) nach Voraussetzung eine zulässige Lösung besitzt) mit dem Hauptsatz 4.19 die Existenz einer optimalen Lösung.

(2) $\Rightarrow$ (3) : Dies entspricht ja genau der Aussage von Satz (6.5).

(3) $\Rightarrow$ (4) $\Rightarrow$ (5) ist unmittelbar. Aus (5) folgt mit dem Hauptsatz 4.19, gegebenenfalls nach Umformung in Standardform, die Existenz einer Optimallösung für eines der beiden Probleme. Der starke Dualitätssatz liefert dann (1). $\square$

Satz 6.8 (Komplementarität). Seien $x \in K$ und $\lambda \in K^*$ zulässig für (P) und (D). Dann sind die folgenden Aussagen äquivalent.

1. x und λ sind optimal.
2. $(\lambda A - c)x = 0$

3. Für alle $i = 1, \dots, n$ gilt

$$x_i > 0 \Rightarrow \lambda a^i = c_i$$

$$\lambda a^i < c_i \Rightarrow x_i = 0.$$

Beweis: Übung (trivial).

6.4 Das Duale Simplex-Verfahren

Der Dualitätssatz 6.5 kann unter den Rangvoraussetzungen des Simplex-Verfahrens einfacher bewiesen werden. Gleichzeitig gelangt man so zum dualen Simplex-Verfahren. Gegeben seien die primal-dualen Programme

$$\min\{cx \mid Ax = b, x \geq 0\} \quad (P)$$

$$\max\{\lambda b \mid \lambda A \leq c\}. \quad (D)$$

Wir setzen $\text{rang}(A) = m$ voraus. Die Basis-Lösung zu einer Basis B von (P) ist nach dem vorigen Kapitel gegeben durch

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

Der Zeilenvektor

$$\lambda := c_B A_B^{-1} \in \mathbb{R}^m \quad (6.2)$$

heißt Simplex- oder Lagrange-Multiplikator. Damit sind die reduzierten Kosten gegeben durch

$$r_N = c_B A_B^{-1} A_N - c_N = \lambda A_N - c_N.$$

Wir erhalten folgende Folgerungen aus dem starken Dualitätssatz.

Satz 6.9. 1. $\lambda = c_B A_B^{-1}$ ist zulässig für $(D) \Leftrightarrow r_N \leq 0$.

2. Die Basis-Lösung x sei optimal und nicht entartet, d.h.,

$$x_B = A_B^{-1}b > 0, x_N = 0.$$

Dann ist $\lambda = c_B A_B^{-1}$ optimal für (D) und es gilt

$$\lambda b = cx.$$

Beweis. Zu (1): Mit $\lambda = c_B A_B^{-1}$ gilt

$$\lambda A \leq c \Leftrightarrow (\lambda A_B, \lambda A_N) \leq (c_B, c_N)$$

$$\Leftrightarrow \lambda A_N \leq c_N$$

$$\Leftrightarrow r_N = \lambda A_N - c_N \leq 0.$$

Zu (2): Ist x optimal und nicht entartet, so gilt notwendigerweise $r_N \leq 0$ (siehe Korollar 5.4). Also ist $\lambda = c_B A_B^{-1}$ zulässig für (D) (nach (1)) und es gilt

$$\lambda b = c_B A_B^{-1} b = c_B x_B = c x.$$

Wegen Korollar (6.4) ist daher $\lambda = c_B A_B^{-1}$ optimal für D. $\square$

Die Verknüpfung von primaler Optimalität und dualer Zulässigkeit wird in folgender Definition aufgegriffen.

Definition 6.10. Eine Basis-Lösung x mit

$$x_B = A_B^{-1} b, x_N = 0$$

heisst dual zulässig für (P), genau dann, wenn

$$\lambda = c_B A_B^{-1}$$

zulässig für (D) ist, d.h., wenn $r_N \leq 0$ gilt.

Wir skizzieren nun die Idee des dualen Simplex-Verfahrens. Es liege eine Basis B und eine Basislösung x vor, die

- nicht primal zulässig ist, d.h. es gilt nicht $x_B = A_B^{-1} b \geq 0$
- dual zulässig ist, d.h. es gilt $r_N \leq 0$.

Man geht vom “alten” primalen Simplextableau aus und löst (P) vom dualen Standpunkt:

1. erhalte die Optimalitätsbedingung $r_N \leq 0$
2. erreiche die Zulässigkeit von x , d.h. $x_B = A_B^{-1} b \geq 0$
3. Der Zielfunktionswert der aktuellen Basislösung nimmt nicht ab; bei einem entarteten Schritt kann er unverändert bleiben.

Mit der beim Simplex-Verfahren getroffenen Annahme $A_B = I$ ist das Simplex Tableau für (P) nach Formel (5.13) durch

$$\begin{array}{c}
 x_j, j \in N \\
 \begin{array}{c|cc}
 z(x) & z_0 & r_N \\
 \hline
 x_i, i \in B & b & A_N
 \end{array}
 \end{array} \tag{6.3}$$

Abbildung 6.1: Erweitertes Simplex-Tableau.

mit der zugehörigen Tableau-Matrix:

Das Tableau (6.4) ist

$$T = (t_{ij}) = \begin{array}{c|c} z_0 & r_N \\ \hline b & A_N \end{array} \quad \begin{array}{l} 0 \leq i \leq m \\ 0 \leq j \leq n - m \end{array} \quad (6.4)$$

1. primal zulässig, wenn $b \geq 0$,
2. dual zulässig, wenn $r_N \leq 0$,
3. optimal, wenn $b \geq 0, r_N \leq 0$.

Nun können wir den dualen Simplex-Algorithmus formal hinschreiben (siehe Abb. 6.2).
Wir geben eine Anwendung. Das Primalproblem sei in der Form

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

gegeben. $Ax \geq b$ ist äquivalent zu

$$Ax - y = b, y \geq 0.$$

Wegen $c \geq 0$ ist die Schlupfvariablenbasis dual zulässig. Man kann daher für (P) direkt mit dem dualen Simplex-Verfahren starten. Dazu betrachtet man das LGS

$$-Ax + y = -b, y \geq 0, x \geq 0.$$

Wir erhalten das Start-Tableau

$$B = (n + 1, \dots, n + m), N = (1, \dots, n) \\ r_N = -c \leq 0.$$

Somit ist $x = 0, y = -b$ dual zulässig und wir erhalten das folgende Tableau:
Wir lösen nun damit das Beispiel aus dem letzten Abschnitt:

$$\begin{aligned} &\text{minimiere } 3x_1 + 4x_2 + 5x_3 \\ &x_1 + 2x_2 + 3x_3 \geq 5 \\ &2x_1 + 2x_2 + x_3 \geq 6 \\ &x_i \geq 0, i = 1, 2, 3 \end{aligned}$$

Das Start-Tableau hat die Form:

		x_1	x_2	x_3
z	0	-3	-4	-5
y_1	-5	-1	-2	-3
y_2	-6	-2	-1	-1

1. **Start:** Sei x dual zulässig zur Basis B . Berechne die Matrix $T = (t_{ij})$, wie in (6.4). Bezeichne mit $B(i)$ den zur i -ten Zeile von T gehörenden Basisindex für $1 \leq i \leq m$. Analog dazu bezeichne mit $N(j)$ den zur j -ten Spalte von T gehörenden Nichtbasisindex für $1 \leq j \leq n - m$.

2. Ist $t_{i0} \geq 0$ ($1 \leq i \leq m$), so ist die Lösung optimal. Setze

$$\begin{aligned} x_{B(i)} &:= t_{i0}, 1 \leq i \leq m \\ x_{N(j)} &:= 0, 1 \leq j \leq n - m \\ z &:= t_{00} \end{aligned}$$

Stopp. Andernfalls gehe zu 3.

3. **Bestimmung der Austausch-Zeile:** Wähle Index $p \in [m]$ (mit $B(p)$ kleinstmöglich) mit

$$t_{p0} = \min_{1 \leq i \leq m} t_{i0} \quad (t_{i0} = b_i).$$

4. Sind alle $t_{pj} \geq 0$, $1 \leq j \leq n - m$, so ist das primale Problem (P) unzulässig und das duale Problem (D) unbeschränkt. **Stopp.**
Gibt es ein $t_{pj} < 0$, so gehe zu Schritt 5.

5. **Bestimmung der Austausch-Spalte:** Wähle Index $s \in [n - m]$ (mit $N(s)$ kleinstmöglich) mit

$$\frac{t_{0s}}{t_{ps}} = \min \left\{ \frac{t_{0j}}{t_{pj}} \mid t_{pj} < 0, j = 1, \dots, n - m \right\} \quad (t_{0j} = r_{N(j)} \leq 0).$$

Gehe zu Schritt 6.

6. Vertausche die Indizes gleichzeitig: $(B(p), N(s)) \leftarrow (N(s), B(p))$. Führe an der Matrix T die Pivotoperation mit Pivotelement $t_{ps} < 0$ durch wie in Abbildung 5.2 Punkt (6):

- Pivotelement: $t'_{ps} := 1/t_{ps}$
- übrige Pivotspalte: $t'_{is} := -\frac{t_{is}}{t_{ps}}, i = 0, 1, \dots, m, i \neq p$
- übrige Pivotzeile: $t'_{pj} := \frac{t_{pj}}{t_{ps}}, j = 0, 1, \dots, n - m, j \neq s$
- sonstige Elemente: $t'_{ij} := t_{ij} - t_{is} \frac{t_{pj}}{t_{ps}}, i \neq p, j \neq s$
- Setze $t_{ij} := t'_{ij}$ und gehe zu Schritt 2.

Abbildung 6.2: Duales Simplex-Verfahren

$$\begin{array}{c} x \\ z \quad y \end{array} \begin{array}{|c|c|} \hline z_0 & r_N = -c \leq 0 \\ \hline -b & -A \\ \hline \end{array} \quad (6.5)$$

Wir erhalten

$$\min\{-5, -6\} = -6 \Rightarrow p = 2$$

$$\min\{3/2, 4/2, 5/1\} = 3/2 \Rightarrow s = 1.$$

Tausche y_2 und x_1 :

		y_2	x_2	x_3
z	9	$-3/2$	-1	$-7/2$
y_1	-2	$-1/2$	-1	$-5/2$
x_1	3	$-1/2$	1	$1/2$

Tausche y_1 und x_2 :

		y_2	y_1	x_3
z	11	-1	-1	-1
x_2	2	$1/2$	-1	$5/2$
x_1	1	-1	1	-2

Das Tableau ist optimal; die optimale Lösung ist

$$x_1 = 1, x_2 = 2, x_3 = 0.$$

Die optimale Lösung des dualen Problems ist

$$\lambda = (1, 1).$$

6.5 Sensitivität und Schattenpreise

Das LP

$$\min\{cx \mid Ax = b, x \geq 0\} \quad (6.6)$$

habe eine optimale Basis B mit nichtentarteter Basislösung $x_B = A_B^{-1}b > 0$, $x_N = 0$. Die zugehörige optimale duale Lösung ist $\lambda = c_B A_B^{-1}$. Bei einer hinreichend kleinen Änderung $b \mapsto b + \Delta b$ bleibt B zulässig. Die neue Basislösung lautet

$$x'_B = A_B^{-1}(b + \Delta b) = x_B + \Delta x_B, \quad x'_N = 0, \quad \Delta x_B = A_B^{-1}\Delta b.$$

Da die reduzierten Kosten $r_N = c_B A_B^{-1} A_N - c_N \leq 0$ nicht von b abhängen, bleibt diese Basislösung optimal, solange sie zulässig ist. Bezeichnet $z(b)$ den Optimalwert, so gilt

$$z(b + \Delta b) - z(b) = cx' - cx = c_B \Delta x_B = \lambda \Delta b.$$

Lokal um b erhalten wir somit die Schattenpreisformel

$$\frac{\partial z}{\partial b_j}(b) = \lambda_j, \quad j = 1, \dots, m. \quad (6.7)$$

λ_j ist der marginale Preis einer Änderung von b_j . Bei größeren Änderungen kann ein Basiswechsel erforderlich werden und damit ein anderer Schattenpreis gelten.

Kapitel 7

Postoptimierung

Wir betrachten in diesem Abschnitt verschiedene Möglichkeiten eine schon berechnete Optimallösung eines LPs bei gestörten Eingabedaten, so zu verändern, dass die veränderte Lösung optimal bzgl. der neuen Eingabedaten ist. Die Veränderungen sollen mit möglichst wenig Aufwand berechnet werden; es gibt ja immer die Möglichkeit alles “from scratch” neu zu berechnen.

Wir betrachten wiederum ein LP in Standardform

$$\begin{aligned} \text{minimiere } z(x) &= cx \\ Ax &= b \\ x &\geq 0, \end{aligned} \tag{LP}$$

sowie eine optimale Basis B .

7.1 Änderung der rechten Seite b

Wir betrachten eine Änderung der rechten Seite der Form

$$b' = b + d \text{ für ein } d \in \mathbb{R}^m.$$

Dann ergibt sich mit der optimalen Basis B für das ungestörte Problem eine neue Basis für das gestörte Problem durch

$$x'_B := A_B^{-1}(b + d) = x_B + A_B^{-1}d \text{ und } x'_N = 0.$$

Sollte nun $x'_B \geq 0$ gelten, ist die so geänderte Basislösung weiterhin optimal, da sich die reduzierten Kosten nicht geändert haben

$$r'_N = r_N = c_B A_B^{-1} A_N - c_N \leq 0.$$

Der neue Zielfunktionswert ist gegeben durch

$$z' = c_B A_B^{-1} b' = z + c_B A_B^{-1} d.$$

Sollte jedoch $x'_B \not\geq 0$ gelten, so ist die zugehörige Basislösung zwar nicht primal, aber **dual** zulässig. Damit können wir direkt das duale Simplexverfahren mit Basismatrix A_B

7.2 Änderung der Zielfunktion c

Wir ersetzen c im LP durch $c' = c + d$. Dann ist der neue Zielfunktionswert für die gegebene Basis B gegeben durch

$$z' = (c + d)x = (c_B + d_B)A_B^{-1}b = z + d_B A_B^{-1}b.$$

Ob nun die aktuelle Basislösung optimal bleibt, liegt an den geänderten reduzierten Kosten. Sie berechnen sich durch

$$\begin{aligned} r'_N &= (c_B + d_B)A_B^{-1}A_N - c_N - d_N \\ &= r_N + d_B A_B^{-1}A_N - d_N. \end{aligned}$$

Falls $r'_N \leq 0$, so bleibt die aktuelle Basislösung optimal auch wenn sich z ändert. Andernfalls gibt es positive reduzierte Kosten $r'_j > 0$. Dann können wir das primale Simplex-Verfahren anwenden mit B als aktueller zulässiger Basis.

7.3 Hinzufügen einer neuen Variablen

Wir wollen eine neue Variable x_{n+1} mit Zielfunktionskoeffizient c_{n+1} in das bestehende LP einführen. Somit ergibt sich auch eine neue Spalte a^{n+1} in der bestehenden Matrix A . Wir können x_{n+1} als zusätzliche Nichtbasisvariable behandeln und das bestehende Tableau T um eine Spalte erweitern. Tableauspalten werden im Folgenden mit $(a')^j$ bezeichnet.

$$\begin{aligned} (a')^{n+1} &= A_B^{-1}a^{n+1} \\ r'_{n+1} &= c_B A_B^{-1}a^{n+1} - c_{n+1} = \lambda a^{n+1} - c_{n+1}. \end{aligned}$$

Falls nun $r'_{n+1} \leq 0$, so liegt weiterhin Optimalität von B vor. Andernfalls können wir das primale Simplex-Verfahren benutzen und x_{n+1} in die Basis bringen.

Der Quotiententest kann dabei auch zeigen, dass das geänderte LP unbeschränkt ist; nicht jede verbessernde neue Variable führt zu einem endlichen Basistausch.

7.4 Änderung einer Spalte von A

Wir ersetzen die Spalte a^j von A durch $a^j + d$. Es ergeben sich zwei Fälle:

1. Fall: $j \in N$.

Da dann die Basismatrix A_B unverändert bleibt, können wir die neue j -te Spalte $(a')^j$ des Tableaus wie folgt berechnen:

$$(a')^j = A_B^{-1}(a^j + d). \quad (7.1)$$

Die Basis bleibt weiterhin primal zulässig, jedoch können sich die reduzierten Kosten

verändern:

$$r'_j = c_B A_B^{-1} (a^j + d) - c_j.$$

Entweder bleibt die aktuelle Basis optimal, oder man führt nun ausgehend von dem geänderten Tableau das primale Simplex-Verfahren mit zulässiger Basis B durch.

Fall 2: $j \in B$: In diesem Fall kann sowohl die primale als auch die duale Zulässigkeit der aktuellen Basis verlorengehen. Auch kann die geänderte Basis Matrix A'_B singulär werden.

Um einige dieser Probleme zu beheben, werden wir eine zusätzliche Spalte der Form

$$s := a^j + d$$

an A anfügen. Somit erhalten wir ein Problem mit einer zusätzlichen Variablen, die wir nun allerdings x'_j nennen. Der Zielfunktionskoeffizient c_j bleibt derselbe. Das Tableau wird um folgende Spalte erweitert

$$\begin{pmatrix} r'_j \\ s' \end{pmatrix} := \begin{pmatrix} c_B A_B^{-1} s - c_j \\ A_B^{-1} s \end{pmatrix}.$$

Falls der Eintrag s'_j in der mit x_j indizierten Zeile von s' nicht 0 ist, können wir diesen Wert als Pivotelement wählen und die alte Basisvariable x_j gegen die neue x'_j tauschen. Nach diesem Schritt lassen wir einfach die alte Variable x_j weg und benennen die neue x'_j in x_j um. Man beachte jedoch dass nach diesem Schritt die aktuelle Basis B weder primal noch dual zulässig sein muss. Falls einer der beiden Zulässigkeiten vorliegt, fahren wir mit dem entsprechenden Simplex-Verfahren fort. Falls keiner der beiden Fälle eintritt, müssen wir die erste Phase des primalen Simplex durchlaufen um eine Startbasis zu erhalten.

Im Fall $s'_j = 0$ wäre selbst nach dem Tausch von x_j und x'_j , B keine Basis mehr; somit kann man nicht viel mehr tun als das Simplex-Verfahren von Anfang an neu anzuwenden.

Wir geben ein Beispiel bezüglich des Abänderns einer Spalte an.

Beispiel 7.1 (Änderung der Spalte). Betrachte folgendes Problem:

Primalproblem (P):

$$\begin{aligned} \min \quad & x_1 + 2x_2 \\ & -x_1 - x_2 \leq -2 \\ & -2x_1 - x_2 \leq -3 \\ & x_i \geq 0, i = 1, 2 \end{aligned}$$

Wir benutzen Schlupfvariablen y_3 und y_4 und erhalten folgendes Tableau:

		x_1	x_2
z	0	-1	-2
y_3	-2	-1	-1
y_4	-3	-2	-1

Die aktuelle Basis ist dual zulässig und nach einigen Austauschschritten mittels des [dualen Simplex-Verfahrens](#) erhalten wir folgendes optimales Tableau:

		x_2	y_3
z	2	-1	-1
y_4	1	1	-2
x_1	2	1	-1

Mit optimaler Basis $B = (4, 1)$.

Wir ersetzen nun die Nichtbasisspalte a^2 durch

$$a^2 + d := \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

und erhalten mit $r_2 = -1$ und

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

die Gleichung

$$r'_2 = r_2 + \lambda d = -1 - \alpha.$$

Also gilt $r'_2 \leq 0$ genau dann wenn $\alpha \geq -1$. Somit ist, falls $\alpha \geq -1$ gilt, die berechnete Basis-Lösung mit $B = (4, 1)$ optimal für alle linearen Programme der Form

$$\begin{aligned} \min \quad & x_1 + 2x_2 \\ & -x_1 + (\alpha - 1)x_2 \leq -2 \\ & -2x_1 + (\beta - 1)x_2 \leq -3 \\ & x_i \geq 0, i = 1, 2. \end{aligned} \quad (\text{LP}(\alpha, \beta))$$

Nun betrachten wir den Fall dass die Basisspalte a^1 geändert wird. Wir erhalten

$$s := a^1 + d = \begin{pmatrix} -1 \\ -2 \end{pmatrix} + \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}$$

mit Tableau-Spalte

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

sowie

$$r'_1 = c_B s' - c_1 = (0, 1) \begin{pmatrix} -\alpha \\ 1 - \alpha \end{pmatrix} - 1 = -\alpha.$$

Somit müssen wir den Vektor $\begin{pmatrix} -\alpha \\ -\alpha \\ 1 - \alpha \end{pmatrix}$ als neue Spalte an das Tableau zur ursprünglich optimalen Basis B anfügen:

	x_2	y_3	x'_1
z	2	-1 -1	$-\alpha$
y_4	1	1 -2	$-\alpha$
x_1	2	1 -1	$(1 - \alpha)$

Die für diesen Fall diskutierte Methode ist nun anwendbar wenn $\alpha \neq 1$ gilt (da ja dann das erwähnte Pivotelement nicht gleich 0 ist). Der Pivotschritt mit dem Pivotelement $(1 - \alpha)$ ergibt folgendes Tableau:

	x_2	y_3	x_1
z	$\frac{2}{1-\alpha}$	$\frac{2\alpha-1}{1-\alpha}$	$\frac{-1}{1-\alpha}$
y_4	$\frac{1+\alpha}{1-\alpha}$	$\frac{1}{1-\alpha}$	$\frac{\alpha-2}{1-\alpha}$
x'_1	$\frac{2}{1-\alpha}$	$\frac{1}{1-\alpha}$	$\frac{1}{1-\alpha}$

Wie beschrieben, werden wir nun die neue x_1 Spalte streichen. Somit erhalten wir:

		x_2	y_3
z	$\frac{2}{1-\alpha}$	$\frac{2\alpha-1}{1-\alpha}$	$\frac{-1}{1-\alpha}$
y_4	$\frac{1+\alpha}{1-\alpha}$	$\frac{1}{1-\alpha}$	$\frac{\alpha-2}{1-\alpha}$
x'_1	$\frac{2}{1-\alpha}$	$\frac{1}{1-\alpha}$	$\frac{-1}{1-\alpha}$

Man sieht nun leicht, dass T' bezüglich B

- genau für $\alpha \leq 1/2$ dual zulässig
- und genau für $-1 \leq \alpha < 1$ primal zulässig

ist. Wenn einer der Fälle eintritt, kann man mit dem entsprechenden Simplex-Verfahren fortfahren. Auch folgt direkt, dass für

$$-1 \leq \alpha \leq 1/2$$

die aktuelle Basis B optimal bleibt. Somit hat das Programm

$$\begin{aligned} \min \quad & x_1 + 2x_2 \\ (\alpha - 1)x_1 - x_2 & \leq -2 \\ (\alpha - 2)x_1 - x_2 & \leq -3 \\ x_i & \geq 0, i = 1, 2 \end{aligned}$$

für $-1 \leq \alpha \leq 1/2$ die optimalen Lösungen

$$x_1^* = \frac{2}{1-\alpha}, x_2^* = 0.$$

Der Zielfunktionswert ist $z^* = \frac{2}{1-\alpha}$.

Kapitel 8

Parametrische Lineare Optimierung

Wir untersuchen nun wie sich Optimallösungen von LPs als Funktionen von Eingangsparametern beschreiben lassen.

8.1 Störung der Zielfunktion c

Dazu betrachten wir nun folgendes parametrisierte LP:

$$\begin{aligned} \min \quad & cx \\ \text{s.t.} \quad & Ax \leq b \\ & x \geq 0, \end{aligned} \tag{LP(c)}$$

In dieser Formulierung sei nun c variabel.

Wir erhalten folgenden Satz.

Satz 8.1. Sei A eine $m \times n$ Matrix und $b \in \mathbb{R}^m$, so dass

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

ein nichtleeres Polyeder ist. Definiere

$$BP(A, b) := \{c \in \mathbb{R}^n | cx \text{ ist auf } P^+(A, b) \text{ nach unten beschränkt}\}.$$

Für $c \in BP(A, b)$ sei

$$M(c) := \min\{cx | x \in P^+(A, b)\}.$$

Dann gilt

1. $BP(A, b)$ ist ein konvexer Kegel.
2. Es gibt endlich viele Vektoren $v_1, \dots, v_N \in \mathbb{R}^n$ mit

$$M(c) = \min\{cv_i | i = 1, \dots, N\}$$

für alle $c \in BP(A, b)$.

3. $M(c)$ ist **konkav** auf $BP(A, b)$.

Beweis. Zu (1): Offensichtlich ist $BP(A, b)$ ein Kegel. Seien nun $c, c' \in BP(A, b)$ und $\alpha \in (0, 1)$. Dann folgt für alle $x \in P^+(A, b)$

$$(\alpha c + (1 - \alpha)c')x = \alpha cx + (1 - \alpha)c'x \geq \alpha M(c) + (1 - \alpha)M(c')$$

also folgt $\alpha c + (1 - \alpha)c' \in BP(A, b)$.

Zu 2.: Da mit dem Hauptsatz der linearen Optimierung jede beschränkte lineare Zielfunktion ihr Minimum auf einer der endlich vielen Ecken von $P^+(A, b)$ annimmt, können wir die v_i als Ecken von $P^+(A, b)$ wählen.

Zu 3. Mit Aussage 2. sehen wir, dass wir $M(c)$ als Minimum von konkaven (sogar linearen) Funktion darstellen können, welches stets konkav ist (vgl. eine ähnlich Aussage über konvexe Funktionen in einem vorherigen Kapitel über Konvexität). □

Korollar 8.2. Die Aussagen von Satz 8.1 gelten auch für Polyeder der Form

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

Wir erhalten ein ähnliches Resultat für Maximierungsprobleme:

$$\max\{cx | x \in P^+(A, b)\}.$$

Satz 8.3. Sei A eine $m \times n$ Matrix und $b \in \mathbb{R}^m$, so dass

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

ein nichtleeres Polyeder ist. Definiere

$$BP^*(A, b) := \{c \in \mathbb{R}^n | cx \text{ ist auf } P^+(A, b) \text{ nach oben beschränkt}\}.$$

Für $c \in BP^*(A, b)$ sei

$$M^*(c) := \max\{cx | x \in P^+(A, b)\}.$$

Dann gilt

1. $BP^*(A, b)$ ist ein konvexer Kegel.
2. Es gibt endlich viele Vektoren $v_1, \dots, v_N \in \mathbb{R}^n$ mit

$$M^*(c) = \max\{cv_i | i = 1, \dots, N\}$$

für alle $c \in BP^*(A, b)$.

3. $M^*(c)$ ist konvex auf $BP^*(A, b)$.

8.2 Störung der rechten Seite b

Als nächstes betrachten wir lineare Programme der Form

$$\begin{aligned} \min & cx \\ & Ax \geq b \\ & x \geq 0, \end{aligned} \quad (\text{LP}(b))$$

In dieser Formulierung sei nun b variabel. (Wir benutzen hier die Formulierung $Ax \geq b$ aus Gründen der einfacheren Darstellung, wie sich später zeigen wird.) Es sei

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

Satz 8.4. Sei A eine $m \times n$ Matrix und $c \in \mathbb{R}^n$, sodass die lineare Zielfunktion cx auf mindestens einem der nichtleeren Polyeder der Form $P^-(A, b)$ nach unten beschränkt ist.

Definiere

$$AP(A, c) := \{b \in \mathbb{R}^m | P^-(A, b) \neq \emptyset, cx \text{ ist auf } P^-(A, b) \text{ nach unten beschränkt}\}.$$

Für $b \in AP(A, c)$ sei

$$L(b) := \min\{cx | x \in P^-(A, b)\}.$$

Dann gilt

1. $AP(A, c)$ ist ein konvexer Kegel.
2. Es gibt endlich viele Vektoren $u_1, \dots, u_N \in \mathbb{R}^m$ mit

$$L(b) = \max\{b^T u_i | i = 1, \dots, N\}$$

für alle $b \in AP(A, c)$.

3. $L(b)$ ist **konvex** auf $AP(A, c)$.

Beweis. Wir betrachten das folgende zu $\text{LP}(b)$ duale LP:

$$\begin{aligned} \max & \lambda b \\ & \lambda A \leq c \\ & \lambda \geq 0, \end{aligned} \quad (\text{DP}(b))$$

welches ein parametrisches LP bezüglich des Zielfunktionsvektors b , bei fester rechten Seite c ist. Der starke Dualitätssatz besagt dann

$$AP(A, c) = BP^*(A^T, c^T),$$

wobei

$$BP^*(A^T, c^T) = \{b \in \mathbb{R}^m \mid \lambda b \text{ ist auf } P^+(A^T, c^T) \text{ nach oben beschränkt}\}.$$

Desweiteren folgt

$$L(b) = M^*(b),$$

und somit folgt aus Satz 8.3 die Behauptung.

□

Die obigen Sätze sind zwar theoretisch recht interessant, liefern aus “praktischer Sicht” jedoch kein nützliches Verfahren um eine parametrisierte Lösung zu berechnen. Dies wollen wir nun an einigen Beispielen nachholen.

Wir betrachten folgendes lineare Programm:

$$\begin{aligned} \min \quad & -x_1 - 3x_2 \\ & x_1 + x_2 \leq 6 \\ & -x_1 + 2x_2 \leq 6 \\ & x_i \geq 0, i = 1, 2 \end{aligned}$$

Es sollte klar sein, wie man nun Schlupfvariablen einführt (y_3, y_4) und dann das Problem löst.

Wir erhalten die optimale Basis $B = (1, 2)$ mit zugehörigem Tableau:

		y_3	y_4
z	-14	$-5/3$	$-2/3$
x_1	2	$2/3$	$-1/3$
x_2	4	$1/3$	$1/3$

Die Zielfunktion soll nun in Richtung $d = (2, 1)$ gestört werden: Die Zielfunktionskoeffizienten und die Störungskomponenten der Schlupfvariablen sind null.

$$c' = c + td = (-1, -3) + t(2, 1) \text{ mit } t \geq 0.$$

Für die gegebene Basis $B = (1, 2)$ führt c' zu den reduzierten Kosten

$$\begin{aligned} r'_N &= c'_B A_B^{-1} A_N - c'_N \\ &= r_N + t(d_B A_B^{-1} A_N - d_N) \\ &:= r_N + td'. \end{aligned}$$

Im Beispiel gilt $r_N = (-5/3, -2/3)$ sowie

$$d' = (2, 1) \begin{pmatrix} 2/3 & -1/3 \\ 1/3 & 1/3 \end{pmatrix} - (0, 0) = (5/3, -1/3).$$

Somit folgt

$$r'_N = (-5/3 + 5/3t, -2/3 - 1/3t).$$

Die bisherige optimale Basis mit zugehöriger Optimallösung bleibt optimal, solange $r'_N \leq 0$ gilt. Das ist für $0 \leq t \leq 1$ der Fall. Also ist

		y_3	y_4
z	$-14 + 8t$	$-5/3 + 5/3t$	$-2/3 - 1/3t$
x_1	2	$2/3$	$-1/3$
x_2	4	$1/3$	$1/3$

ein optimales Tableau für $LP(t)$ für alle $t \in [0, 1]$. Der Eintrag $-14 + 8t$ ergibt sich durch

$$c'_B A_B^{-1} b = c_B A_B^{-1} b + t d_B A_B^{-1} b = -14 + t(2, 1) \begin{pmatrix} 2 \\ 4 \end{pmatrix} = -14 + 8t.$$

Weil nun für $t > 1$, $r'_3 > 0$ würde, werden wir für solche Werte von t die Variable y_3 in die Basis bringen und das eben berechnete Tableau mit $t = 1$ verwenden. (Wegen $r'_r = 0$ für diesen Fall, wird weiterhin Optimalität garantiert).

Wir ersetzen also das Tableau

		y_3	y_4
z	-6	0	-1
x_1	2	$2/3$	$-1/3$
x_2	4	$1/3$	$1/3$

durch das Tableau

		x_1	y_4
z	-6	0	-1
y_3	3	$3/2$	$-1/2$
x_2	3	$-1/2$	$1/2$

welches immer noch optimal ist.

Wir berechnen nun die reduzierten Kosten $r'_{N'}$ als Funktion von $t \geq 1$ bzgl. Basis $B' = (3, 2)$ und $N' = (1, 4)$ neu und erhalten

$$\begin{aligned} r'_{N'} &= c'_{B'} A_{B'}^{-1} A_{N'} - c'_{N'} \\ &= ((0, -3) + t(0, 1)) \begin{pmatrix} 3/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} - ((-1, 0) + t(2, 0)) \\ &= (3/2 - 1/2t, -3/2 + 1/2t) - (-1 + 2t, 0) \\ &= (5/2 - 5/2t, -3/2 + 1/2t). \end{aligned}$$

Nun gilt $r'_{N'} \leq 0$ für $0 \leq t \leq 3$. Mit

$$c'_{B'} x_{B'} = (c_{B'} + t d_{B'}) A_{B'}^{-1} b = ((0, -3) + t(0, 1)) \begin{pmatrix} 3 \\ 3 \end{pmatrix} = -9 + 3t$$

erhalten wir das Tableau

		x_1	y_4
z	$-9 + 3t$	$5/2 - 5/2t$	$-3/2 + 1/2t$
y_3	3	$3/2$	$-1/2$
x_2	3	$-1/2$	$1/2$

Dieses Tableau ist für $1 \leq t \leq 3$ optimal und für $t = 3$ erhalten wir $r'_4 = 0$, somit bringen wir dann für $t \geq 3$ die Variable y_4 in die Basis.

		x_1	x_2
z	0	-5	0
y_3	6	1	1
y_4	6	-1	2

Wenn man nun für $B'' = (3, 4)$ wieder die reduzierten Kosten ausrechnet, bekommt man

$$\begin{aligned} r'_{N''} &= c'_{B''} A_{B''}^{-1} A_{N''} - c'_{N''} \\ &= ((0, 0) + t(0, 0)) \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} - ((-1, -3) + t(2, 1)) \end{aligned}$$

$$\begin{aligned}
 &= (0, 0) - (-1 + 2t, -3 + t) \\
 &= (1 - 2t, 3 - t).
 \end{aligned}$$

Weil nun $r_{N''} \leq 0$ für alle $t \geq 3$, bleibt die aktuelle Basis B'' optimal für diesen Fall. Das Tableau ist

		x_1	x_2
z	0	$1 - 2t$	$3 - t$
y_3	6	1	1
y_4	6	-1	2

Wir haben nun für alle $t \geq 0$ die zugehörigen optimalen Lösungen und Zielfunktionswerte in Abhängigkeit von t berechnet.

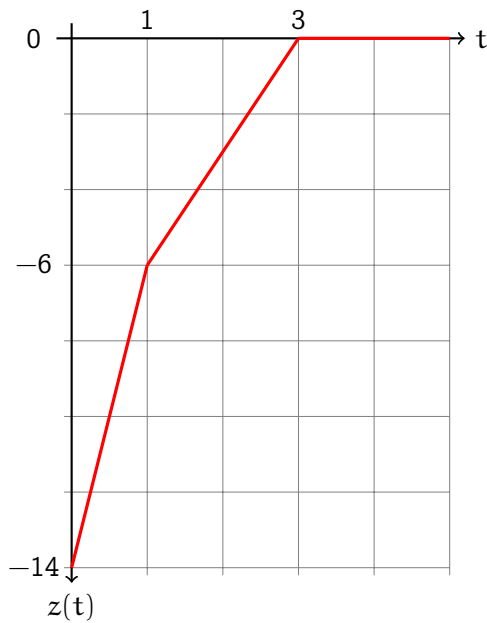


Abbildung 8.1: Optimalwertfunktion $z(t)$.

Eine ähnliche Vorgehensweise wird für Störungen der rechten Seite und/oder Störungen in der Koeffizientenmatrix durchgeführt (siehe Übungen).

Kapitel 9

Komplexität der Linearen Optimierung

In diesem Kapitel werden komplexitätstheoretische Grundlagen gelegt um eine Laufzeitanalyse eines polynomialen Algorithmus (der sogenannten [Ellipsoidmethode](#)) für LPs zu entwickeln. Zunächst werden wir jedoch zeigen, dass der Simplex-Algorithmus (mit einer [schlechten](#) Pivot-Regel) exponentielle Laufzeit haben kann.

9.1 Klee-Minty Würfel

Im Kapitel über das Simplex-Verfahren hatten wir gesehen, dass Dantzig's Pivot-Regel nicht notwendigerweise konvergiert sondern Kreiseln kann. Andererseits hatten wir für die lexikographische Pivot-Regel gezeigt, dass nach endlich vielen Schritten ein Optimum erreicht wird (falls es existiert). Nun stellt sich natürlicherweise die Frage, ob man eine [polynomielle](#) Laufzeitbeschränkung für besonders geeignete Pivotregeln angeben kann. Diese Frage ist bis heute ungelöst und wir werden später noch eingehend darauf zurückgehen. Wir konstruieren nun einen verbessernden Simplex-Pfad, der exponentiell viele Ecken besucht. Eine geeignete Wahl der Pivotschritte kann also exponentielle Laufzeit erzwingen. Damit ist noch keine Aussage für jede einzelne bekannte Pivot-Regel bewiesen. Dazu definieren wir die folgenden [Klee-Minty](#) Würfel.

Definition 9.1. Die d -dimensionalen Klee-Minty Würfel sind durch folgende Ungleichungen definiert:

$$\epsilon \leq x_1 \leq 1 \tag{9.1}$$

$$\epsilon x_{j-1} \leq x_j \leq 1 - \epsilon x_{j-1} \text{ für } j = 2, \dots, d, \tag{9.2}$$

wobei $0 < \epsilon < 1/2$ gelte. Die Wahl $\epsilon = 0$ würde auf den gewöhnlichen d -dimensionalen Würfel führen.

Wir geben ein Beispiel für $d = 3$ und $\epsilon = 1/4$ an.

$$1 \leq 4x_1 \leq 4 \tag{9.3}$$

$$x_1 \leq 4x_2 \leq 4 - x_1 \tag{9.4}$$

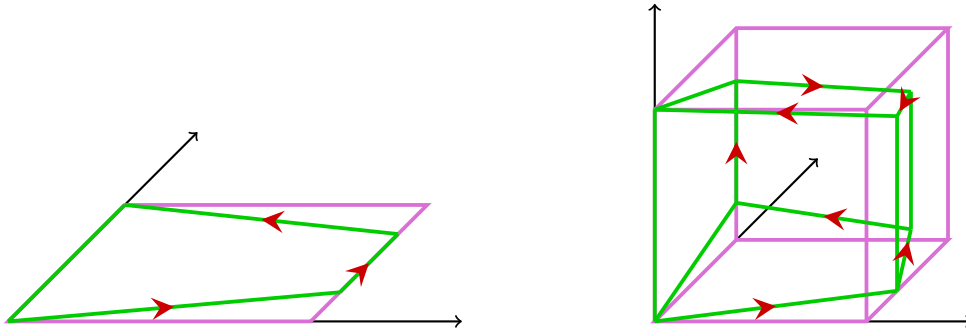


Abbildung 9.1: Der Klee-Minty Würfel mit entsprechendem Simplex-Pfad für Dimension $d = 2$ und $d = 3$.

$$x_2 \leq 4x_3 \leq 4 - x_2 \quad (9.5)$$

In Abbildung 9.1 haben wir die Ecken und Kanten des Klee-Minty Würfels angegeben. Der dort eingezeichnete Weg über die acht Ecken der rechten schematischen Darstellung entspricht der Ausführung des Simplex-Algorithmus (mit entsprechender Pivot-Regel) für das LP mit $d = 3$, bei dem die $-x_3$ -Koordinate minimiert wird.

Wir werden nun diese Beobachtung auf d -dimensionale Klee-Minty Würfel verallgemeinern.

$$\begin{aligned} \min -x_d \\ x_1 - r_1 &= \epsilon \\ x_1 + s_1 &= 1 \\ x_j - \epsilon x_{j-1} - r_j &= 0 \text{ für } j = 2, \dots, d \\ x_j + \epsilon x_{j-1} + s_j &= 1 \text{ für } j = 2, \dots, d \\ x, r, s &\geq 0, \end{aligned} \quad (\text{KM}(d))$$

wobei wieder $0 < \epsilon < 1/2$ gilt.

Wir erhalten folgendes Lemma, welches die Basen von $\text{KM}(d)$ charakterisiert.

Lemma 9.2. Die zulässigen Basen für $\text{KM}(d)$ entsprechen bijektiv den Teilmengen von $\{1, \dots, d\}$ und sind alle nichtentartet. Dabei gilt folgende Darstellung für die Indizes einer Basis in Mengenschreibweise:

$$B_S := \{x_1, \dots, x_d\} \cup \{r_j | j \in S\} \cup \{s_j | j \notin S\}. \quad (9.6)$$

Beweis. Es sei B eine zulässige Basis für $\text{KM}(d)$ mit zugehöriger Basislösung x . Aus den Ungleichungen $x_1 \geq \epsilon$ und $x_j \geq \epsilon x_{j-1}$ für $j = 2, \dots, d$ folgt dass

$$x_j \geq \epsilon^j > 0 \text{ für } j = 1, \dots, d,$$

was wiederum impliziert dass die Indizes $1, \dots, d$ sämtlich in der Basis enthalten sind.

Desweiteren muss gelten, dass für jedes j mindestens eine der beiden Variablen r_j bzw. s_j auch in B_S enthalten sein muss. Angenommen es gilt $r_j = s_j = 0$. Für den Fall $j = 1$ ergibt sich dann $x_1 = \epsilon = 1$, Widerspruch. Für $j > 1$ ergibt sich mit $x_j = \epsilon x_{j-1}$ und $x_j + \epsilon x_{j-1} = 1$ die Bedingung

$$2\epsilon x_{j-1} = 1.$$

Wegen $x_{j-1} \leq 1$ und $\epsilon < 1/2$ ergibt das ebenfalls einen Widerspruch. Also ist für jede Variable x_j mindestens eine der beiden Variablen r_j bzw. s_j positiv und gehört somit zur Basis. Damit haben wir mindestens $2d$ Basisvariablen mit positiven Einträgen gefunden. Wegen $m = 2d$ kann es nicht mehr als $2d$ Basisvariablen geben und somit ist genau eine der beiden Variablen r_j oder s_j positiv, so dass die Basis B_S in der Form (9.6) geschrieben werden kann. Desweiteren folgt dass die Basis nichtentartet ist.

Umgekehrt ist es leicht ausgehend von S eine zulässige Basislösung zu berechnen (Übung). $\square$

Wir erhalten mit Satz 4.22 folgende Folgerung:

Korollar 9.3. Der in $KM(d)$ beschriebene Klee-Minty Würfel hat genau 2^d Ecken.

Wir halten noch weitere Beobachtungen fest, die sich später als nützlich herausstellen werden.

Korollar 9.4. Sei B_S eine zulässige Basislösung für $KM(d)$ wie in (9.6) angegeben. Dann gelten folgende Aussagen:

- Es gilt $x_1 = 1$, falls $1 \in S$ und $x_1 = \epsilon$, sonst.
- Für $j \in S$ mit $j \geq 2$ gilt $x_j = 1 - \epsilon x_{j-1}$.
- Für $j \notin S$ mit $j \geq 2$ gilt $x_j = \epsilon x_{j-1}$.

Mit Korollar 9.4 kann man nun in einfacher Weise für eine vorgegebene Basis B_S die dazugehörige Basislösung (rekursiv) berechnen.

Als nächstes betrachten wir die für den Simplex-Pfad wichtige d -te Koordinate von x .

Lemma 9.5. Seien S und S' zwei Teilmengen von $\{1, \dots, d\}$ und x^S bzw. $x^{S'}$ die zugehörigen Ecken des KM -Würfels. Falls $d \in S \setminus S'$ gilt, folgt $x_d^S > x_d^{S'}$. Wenn $d \geq 2$ und $S' = S \setminus \{d\}$ gilt, folgt $x_d^{S'} = 1 - x_d^S$.

Beweis. Für $d = 1$ gilt $x_1^{\{1\}} = 1 > \epsilon = x_1^\emptyset$. Im Folgenden sei $d \geq 2$. Wir benutzen die Aussagen von Korollar (9.4). Mit $d \in S, x_{d-1}^S \leq 1$ und $\epsilon < 1/2$ erhalten wir $x_d^S = 1 - \epsilon x_{d-1}^S > 1/2$. Andererseits folgt aus $d \notin S'$ die Bedingung $x_d^{S'} = \epsilon x_{d-1}^{S'} < 1/2$, was dann $x_d^S > x_d^{S'}$ zeigt. Falls zusätzlich noch $S' = S \setminus \{d\}$ gilt, erhält man durch rekursive

Anwendung von Korollar (9.4) die Gleichungen $x_j^S = x_j^{S'}$ für $j = 1, \dots, d-1$. Damit ergibt sich sofort

$$x_d^{S'} = \epsilon x_{d-1}^{S'} = \epsilon x_{d-1}^S = 1 - (1 - \epsilon x_{d-1}^S) = 1 - x_d^S. \quad \square$$

Wir zeigen ein weiteres Lemma.

Lemma 9.6. Es seien die $D = 2^d$ Teilmengen von $\{1, \dots, d\}$ so nummeriert, dass

$$x_d^{S_1} \leq x_d^{S_2} \leq \dots \leq x_d^{S_D}. \quad (9.7)$$

Dann sind alle Ungleichungen strikt, und die Ecken x^{S_i} und $x^{S_{i+1}}$ sind für alle $i \in \{1, \dots, D-1\}$ adjazent.

Beweis. Wir beweisen das Lemma per Induktion nach $d \in \mathbb{N}$. Für $d = 1$ gilt mit $S_1 = \emptyset$ und $S_2 = \{1\}$

$$x_1^{S_1} = \epsilon \text{ und } x_1^{S_2} = 1,$$

die Behauptung ist also korrekt. Für den Induktionsschluss $d \rightarrow d+1$ verwenden wir, dass $S_1, \dots, S_D$ auch stets Teilmengen von $\{1, \dots, d+1\}$ sind und somit auch 2^d Ecken der insgesamt 2^{d+1} vielen Ecken des $(d+1)$ -dimensionalen KM-Würfels beschreiben. Beachte dass sich dabei die $d+1$ -erste Koordinate nach Korollar (9.4) durch

$$x_{d+1}^{S_i} = \epsilon x_d^{S_i}$$

ergibt. Mit der Induktionsvoraussetzung (9.7) ergibt sich

$$x_{d+1}^{S_1} < x_{d+1}^{S_2} < \dots < x_{d+1}^{S_D}. \quad (9.8)$$

Die restlichen D Teilmengen von $\{1, \dots, d+1\}$ haben die Form

$$\bar{S}_i := S_i \cup \{d+1\} \text{ für } i = 1, \dots, D.$$

Nach Lemma 9.5 gilt nun $(d+1 \in \bar{S}_i \setminus S_j)$

$$x_{d+1}^{\bar{S}_i} > x_{d+1}^{S_j} \text{ für } i, j = 1, \dots, D,$$

sowie $(S_i = \bar{S}_i \setminus \{d+1\})$

$$x_{d+1}^{\bar{S}_i} = 1 - x_{d+1}^{S_i} \text{ für } i = 1, \dots, D.$$

Somit erhalten wir mit (9.8) die Ungleichungen

$$x_{d+1}^{S_1} < x_{d+1}^{S_2} < \dots < x_{d+1}^{S_D} < x_{d+1}^{\bar{S}_D} < x_{d+1}^{\bar{S}_{D-1}} < \dots < x_{d+1}^{\bar{S}_1}.$$

Desweiteren sind die Ecken x^{S_i} und $x^{S_{i+1}}$ aufgrund der Induktionsannahme für alle $i \in \{1, \dots, D-1\}$ adjazent. Dies impliziert nun dass auch $x^{\bar{S}_i}$ und $x^{\bar{S}_{i+1}}$ immer adjazent sind: die Basis B_{S_i} geht aus der Basis $B_{S_{i+1}}$ durch Austausch einer Variablen hervor und das gilt

auch für die Basen $B_{\bar{s}_i}$ und $B_{\bar{s}_{i+1}}$. Schlussendlich sind auch die Ecken x^{s_D} und $x^{\bar{s}_D}$ adjazent, da $B_{\bar{s}_D}$ aus B_{s_D} durch Weglassen von s_{d+1} und Hinzufügen von r_{d+1} hervorgeht. $\square$

Kombination dieser Lemmata impliziert folgendes Resultat:

Satz 9.7. Für jedes $d > 1$ gibt es ein LP in Standardform mit $2d$ Gleichungen und $3d$ Variablen, ganzzahligen Koeffizienten mit Betrag ≤ 4 sowie eine Pivot-Regel, für das das zugehörige Simplex-Verfahren $2^d - 1$ Iterationen zur Bestimmung des Optimums benötigen kann.

Beweis. Wir setzen $\epsilon := 1/4$ und multiplizieren die Gleichungen in $KM(d)$ mit 4 womit alle Koeffizienten ganzzahlig werden. Mit der Zielfunktion $-x_d$, folgt die Behauptung direkt aus Lemma 9.6. $\square$

Wir schliessen diesen Abschnitt mit einigen Anmerkungen zur Existenz von Pivot-Regeln, die in polynomieller Zeit ein Optimum berechnen. Dazu definieren wir zunächst den Begriff des **Diameters von Polyedern**.

- Definition 9.8.**
1. Die Distanz $d(x, y)$ zwischen zwei Ecken x, y eines Polyeders P ist definiert als die minimale Kantenanzahl eines Weges (definiert als Folge von adjazenten Ecken) von x zu y .
 2. Der Diameter $D(P)$ eines Polyeders mit mindestens einer Ecke ist das Maximum $d(x, y)$ über alle Paare von Ecken (x, y) in P .
 3. $\Delta(n, m)$ ist definiert als das maximale $D(P)$ über alle volldimensionalen **Polytope** im $\mathbb{R}^n$ mit höchstens m Facetten, wobei $m \geq n + 1$.
 4. $\Delta_u(n, m)$ ist definiert als das maximale $D(P)$ über alle volldimensionalen **spitzen Polyeder** im $\mathbb{R}^n$ mit höchstens m Facetten. Spitz bedeutet hier, dass das Polyeder keine Linie enthält.

Offensichtlich ist der Diameter eines Polyeders (oder Polytops) eine **untere** Schranke an die Laufzeit des Simplex-Verfahren **unabhängig** von der Auswahl der Pivot-Regel.

Die folgende berühmte **Hirsch-Vermutung** besagt Folgendes:

Vermutung 9.9 (Polynomiale Hirsch-Vermutung).

$$\Delta(n, m) \leq \text{poly}(n, m).$$

Bekannte **Schranken** sind:

Untere Schranke: Das unbeschränkte Klee-Walkup-Polyeder hat Dimension 4, acht Facetten und Diameter 5. Daher gilt $\Delta_u(4, 8) \geq 5 > 8 - 4$.

Obere Schranke:

$$\Delta(n, m) \leq \Delta_u(n, m) \leq m^{1+\log_2(n)} = (2n)^{\log_2(m)}$$

- Die **Hirsch-Vermutung** $\Delta(n, m) \leq m - n$ wurde 2010 von Paco Santos für $n = 43, m = 86$ widerlegt.

Bemerkung 9.10. Es gibt einige erfolgreiche Versuche das gute empirische Verhalten des Simplex-Algorithmus mathematisch zu erklären. Zum Beispiel hat Borgwardt in einigen Arbeiten gezeigt, dass unter bestimmten Voraussetzungen **die erwartete Laufzeit des Simplex-Algorithmus** in einem gewissen probabilistischen Modell polynomiell ist.

9.2 Polynomiale Algorithmen

Wir werden in den folgenden Abschnitten die erste bekannte Methode zur Lösung von linearen Programmen in **Polynomialzeit** kennenlernen. Sie geht auf den russischen Mathematiker Leonid Khachiyan (1979) zurück. Zunächst müssen wir jedoch formal den Begriff **Laufzeit eines Algorithmus** definieren.

Um Laufzeiten von Algorithmen zu beschreiben, benötigt man eine Definition des **Maschinenmodells**, **Kodierungsvorschriften** etc. Wir werden hier nur die für die Behandlung von LPs wichtigen Begriffe einführen.

Zunächst müssen wir beachten dass jede Zahl in einem Maschinenmodell nur durch endlich viele Bits dargestellt werden kann, daher nehmen wir im Folgenden an, dass die auftretenden LPs nur **rationale Koeffizienten** haben.

Haben wir eine Ungleichung der Form

$$ax \leq \alpha$$

mit $a \in \mathbb{Q}^n, \alpha \in \mathbb{Q}$ gegeben, so können wir das kleinste gemeinsame Vielfache p der Nenner der Komponenten von a und des Nenners von α bestimmen. Die Ungleichung

$$pax \leq p\alpha$$

hat dann **ganzzahlige** Koeffizienten und ist offenbar äquivalent zur vorigen Ungleichung.

Annahme 9.11. Alle Daten der auftretenden LPs sind **ganzzahlig**, d.h., für $\min\{cx | Ax \leq b, x \geq 0\}$ oder $\max\{cx | Ax \leq b, x \geq 0\}$ gilt $c^T \in \mathbb{Z}^n, A \in \mathbb{Z}^{m \times n}, b \in \mathbb{Z}^m$.

Der Computer (also auch ein vernünftiges Maschinenmodell) arbeitet mit **Binärcodes**. Die binäre Darstellung einer ganzen Zahl n benötigt $\lceil \log_2(|n| + 1) \rceil$ viele Stellen (bits) und eine Stelle für das Vorzeichen.

Definition 9.12. Für $n \in \mathbb{Z}$ heisst

$$\langle n \rangle := \lceil \log_2(|n| + 1) \rceil + 1$$

die **Kodierungslänge** von n . Für jede nichtganzzahlige rationale Zahl $r = p/q$ in teiler-

fremder 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 einen Vektor) ist

$$\langle A \rangle = \sum_{i=1}^m \sum_{j=1}^n \langle a_{ij} \rangle.$$

Wir beschreiben nun einige nützliche Eigenschaften der Kodierungslänge.

Lemma 9.13. 1. Für jede Zahl $r \in \mathbb{Q}$ gilt:

$$|r| \leq 2^{\langle r \rangle - 1} - 1.$$

2. Für je zwei Zahlen $r, s \in \mathbb{Q}$ gilt

$$\langle rs \rangle \leq \langle r \rangle + \langle s \rangle.$$

3. Für jeden Vektor $x \in \mathbb{Q}^n$ gilt

$$\|x\|_2 \leq \|x\|_1 \leq 2^{\langle x \rangle - n} - 1.$$

4. Für jede Matrix $A \in \mathbb{Q}^{n \times n}$ gilt:

$$|\det(A)| \leq 2^{\langle A \rangle - n^2} - 1.$$

Beweis. Aussagen (1),(2) werden in den Übungen behandelt.

Zu (3): Sei $x \in \mathbb{Q}^n$. Dann gilt mit (1)

$$1 + \|x\|_1 = 1 + \sum_{i=1}^n |x_i| \leq \prod_{i=1}^n (1 + |x_i|) \leq \prod_{i=1}^n 2^{\langle x_i \rangle - 1} = 2^{\langle x \rangle - n}.$$

Die Ungleichung

$$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2} \leq \sum_{i=1}^n |x_i| = \|x\|_1$$

ist trivial.

Zu (4): Es gilt folgende Ungleichungskette:

$$1 + |\det(A)| \leq 1 + \prod_{j=1}^n \|a^j\|_2 \leq \prod_{j=1}^n (1 + \|a^j\|_2) \leq \prod_{j=1}^n 2^{\langle a^j \rangle - n} = 2^{\langle A \rangle - n^2}.$$

Die erste Ungleichung folgt aus der [Hadamard Ungleichung](#). Die dritte gilt wegen (3). $\square$

Im Folgenden werden wir die Laufzeit eines Algorithmus in Form der Anzahl von **elementaren Rechenoperationen** angeben. Elementare Rechenoperationen sind **Addition, Subtraktion, Multiplikation, Division** und **Vergleich** von ganzen oder rationalen Zahlen. Zudem müssen wir die Multiplikation großer Zahlen mitberücksichtigen.

Definition 9.14. Die **Laufzeit eines Algorithmus** A auf einer Instanz Π , bezeichnet mit $L_A(\Pi)$, ist die Anzahl der elementaren Bitoperationen während seiner Ausführung.

Arithmetische Operationen auf ganzen oder rationalen Zahlen lassen sich mit einer in der Kodierungslänge polynomialen Anzahl von Bitoperationen ausführen. Für polynomielle Bitlaufzeit genügt es daher, sowohl die Anzahl arithmetischer Operationen als auch die Kodierungslängen aller Zwischenergebnisse polynomial zu beschränken. Ihr einfaches Produkt ist im Allgemeinen nicht die tatsächliche Bitlaufzeit.

Nun können wir auch definieren was **polynomielle Laufzeit** bedeutet.

Definition 9.15. Sei A ein Algorithmus zur Lösung einer Klasse von Problemen mit Instanzen Π . Sei $\langle \Pi \rangle$ die Kodierungslänge der Daten zur Beschreibung von Π .

1. Die Funktion $f : \mathbb{N} \rightarrow \mathbb{N}$ definiert durch

$$f_A(n) := \max\{L_A(\Pi) \mid \langle \Pi \rangle \leq n\}$$

heißt **Laufzeitfunktion** von A .

2. Der Algorithmus A hat eine **polynomielle Laufzeit** (A ist ein polynomialer Algorithmus), wenn es ein Polynom $p : \mathbb{N} \rightarrow \mathbb{N}$ gibt, so dass

$$f_A(n) \leq p(n) \quad \forall n \in \mathbb{N}.$$

Bemerkung 9.16. Zur Lösung von LPs suchen wir also einen Algorithmus, dessen Laufzeit durch ein Polynom in $\langle A \rangle + \langle b \rangle + \langle c \rangle$ beschränkt ist.

9.3 Reduktion auf Zulässigkeitsprobleme – LP-Dualität

Die Ellipsoidmethode (die später eingeführt wird) ist im strengen Sinne kein Optimierungsverfahren sondern ein Algorithmus, der für ein gegebenes Polytop einen zulässigen Punkt findet oder feststellt, dass es leer ist. Wir müssen daher die Lösung eines LPs auf ein solches Zulässigkeitsproblem zurückführen.

Betrachten wir ein LP der Form

$$\begin{aligned} \min \quad & cx \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0, \end{aligned} \tag{LP}$$

so erhalten wir das duale LP

$$\begin{aligned} \max \lambda b \\ \lambda A \leq c. \end{aligned} \quad (\text{DP})$$

Aus dem starken Dualitätssatz wissen wir, dass die beiden Probleme genau dann Optimallösungen mit gleichem Zielfunktionswert haben, wenn beide Probleme **zulässig** sind. Somit folgt, dass jeder Punkt des Polyeders $P^\leq$ definiert als

$$\begin{pmatrix} c & -b^\top \\ -A & 0 \\ A & 0 \\ 0 & A^\top \\ -I & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda^\top \end{pmatrix} \leq \begin{pmatrix} 0 \\ -b \\ b \\ c^\top \\ 0 \end{pmatrix}. \quad (9.9)$$

ein Optimalpunkt von (LP) sein muss (betrachte dazu nur die Variablenbelegung von x). Der Nachteil dieser Reduktion ist, dass das entstehende Polyeder $P^\leq$ häufig nicht **volldimensional** ist, d.h. $\text{int}(P^\leq) = \emptyset$ auch wenn das primale Polyeder volldimensional war.

9.4 Reduktion mit Binärsuche

Wir benutzen Abschätzungen der “Größen” von Ecken um das Lösen eines LPs auf ein Zulässigkeitsproblem auf Polytopen zurückzuführen.

Diese Polytope müssen nicht volldimensional sein. Wir setzen im Folgenden $\text{rank}(A) = m$ voraus. Gaußelimination erkennt widersprüchliche Gleichungen und entfernt redundante Gleichungen in Polynomialzeit. Bleiben keine Gleichungen übrig, lässt sich das Problem über $x \geq 0$ direkt lösen.

Satz 9.17. Für jede Ecke $v = (v_1, \dots, v_n)$ eines Polyeders $P = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$ mit $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{Z}^m$ und $\text{rang}(A) = m$ gilt:

$$v_i \leq 2^{\langle A \rangle + \langle b \rangle - m^2} \text{ für alle } i = 1, \dots, n.$$

Beweis. Ist v eine Ecke von P , so gibt es eine reguläre $m \times m$ Teilmatrix A_B von A , so dass v_B die eindeutige Lösung des LGS

$$A_B x_B = b$$

ist. Mit der Cramerschen Regel gilt dann

$$v_i = \frac{\det A_B^i}{\det A_B} = \frac{|\det A_B^i|}{|\det A_B|} \text{ für } i \in B \text{ und } v_i = 0 \text{ für } i \in N,$$

wobei A_B^i aus A_B dadurch entsteht dass die i -te Spalte von A_B durch den Vektor b ersetzt wird. Beachte ausserdem dass $v_i \geq 0$ gilt, somit folgt

$$\frac{\det A_B^i}{\det A_B} = \frac{|\det A_B^i|}{|\det A_B|}.$$

Wir wiederholen kurz den Laplaceschen Determinanten-Entwicklungssatz (nach der i -ten Zeile einer $n \times n$ Matrix F):

$$\det F = \sum_{j=1}^n (-1)^{i+j} f_{ij} \det F_{ij}, \quad (9.10)$$

wobei F_{ij} die $(n-1) \times (n-1)$ Untermatrix von F ist, in der die i -te Zeile und j -te Spalte gestrichen wird.

Aus (9.10) folgt

$$|\det A_B| \geq 1.$$

Ausserdem ist A_B eine Teilmatrix von A , und somit folgt aus Lemma 9.13(4):

$$|\det A_B^i| \leq 2^{\langle A \rangle + \langle b \rangle - m^2}.$$

□

Aus diesem Satz erhalten wir folgende Folgerung.

Satz 9.18. Das primale (LP) hat eine endliche Optimallösung genau dann wenn folgende zwei LPs

$$\begin{aligned} \min cx \\ Ax = b \\ x_i \leq 2^{\langle A \rangle + \langle b \rangle - m^2} \text{ für } i = 1, \dots, n \\ x \geq 0 \end{aligned} \quad (\text{LP-Polytop})$$

$$\begin{aligned} \min cx \\ Ax = b \\ x_i \leq 2^{\langle A \rangle + \langle b \rangle - m^2} + 1 \text{ für } i = 1, \dots, n \\ x \geq 0. \end{aligned} \quad (\text{LP-Polytop}+1)$$

endliche Optimallösungen mit gleichem Zielfunktionswert besitzen. Das Problem (LP) ist genau dann unbeschränkt wenn die beiden Zielfunktionswerte der Probleme (LP-Polytop) und (LP-Polytop+1) nicht übereinstimmen. Das Problem (LP) ist genau dann unzulässig wenn eines der Probleme (LP-Polytop) bzw. (LP-Polytop+1) unzulässig ist.

Nun haben wir das Lösen eines allgemeinen LPs in Standardform auf das Lösen eines LPs über einem Polytop zurückgeführt.

Wir müssen im Folgenden also “nur” LPs über Polytope lösen.

Satz 9.19. Betrachte (LP) mit $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{Z}^m$, $c \in \mathbb{Z}^n$, so dass $P^=(A, b)$ ein nicht-leeres Polytop ist. Dann nimmt die Zielfunktion cx nur Optimalwerte in der endlichen Menge

$$\mathcal{S} := \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z}, |p| \leq n 2^{\langle A \rangle + \langle c \rangle + \langle b \rangle - m^2}, 1 \leq q \leq 2^{\langle A \rangle - m^2} \right\} \quad (9.11)$$

an.

Beweis. Mit dem Hauptsatz der linearen Optimierung müssen wir nur Ecken v des Polytops $P^=(A, b)$ betrachten:

$$cv = \sum_{i=1}^n c_i v_i.$$

Wie im vorigen Beweis erhalten wir

$$\sum_{i=1}^n c_i v_i = \sum_{i \in B} c_i \frac{|\det A_B^i|}{|\det A_B|} =: \frac{p}{q} \in \mathbb{Q}.$$

Wir setzen

$$q := |\det A_B|$$

und

$$p := \sum_{i \in B} c_i |\det A_B^i|$$

und erhalten

$$|p| \leq n 2^{\langle A \rangle + \langle c \rangle + \langle b \rangle - m^2}.$$

□

Dieser Satz erlaubt es nun eine **binäre Suche** über den endlichen Suchraum $\mathcal{S}$ durchzuführen.

Binäre Suche

Input: Ein LP der Form $\min\{cx \mid Ax = b, x \geq 0\}$ mit $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{Z}^m$ und $c \in \mathbb{Z}^n$ und beschränktem Zulässigkeitsbereich.

Output: Eine Optimallösung x^* , oder die Aussage dass $P^-(A, b) = \emptyset$.

1. Überprüfe ob

$$P^-(A, b) = \emptyset.$$

Falls ja, dann STOP, LP ist unzulässig.

2. Wähle ein Element $s \in S$ mit der Eigenschaft dass für $S' = \{t \in S \mid t \leq s\}$ und $S'' = \{t \in S \mid t > s\}$ gilt:

$$|S'| \leq |S''| \leq |S'| + 1.$$

3. Überprüfe ob

$$P_s := P^-(A, b) \cap \{x \in \mathbb{R}^n \mid cx \leq s\} = \emptyset.$$

4. Ist $P_s = \emptyset$, so setze $S \leftarrow S''$, andernfalls $S \leftarrow S'$.

5. Ist $|S| = 1$, so gilt für $s \in S$:

$$s = \min\{cx \mid x \in P^-(A, b)\},$$

Rufe das Zulässigkeitsverfahren für diesen endgültigen Wert s nochmals auf, um einen Punkt in P_s zu erhalten. Dieser Punkt ist optimal. Anderenfalls, gehe zu (2).

Satz 9.20. Binäre Suche ist ein korrektes Verfahren und terminiert nach höchstens

$$\log_2(n) + 2\langle A \rangle + \langle c \rangle + \langle b \rangle - 2m^2 + 3$$

vielen Ausführungen von Schritt (3).

Beweis. Da in jeder Iteration die Kardinalität von S (nahezu) halbiert wird, sind höchstens

$$N' = \lceil \log_2(|S|) \rceil$$

Aufspaltungen von S in zwei (nahezu) gleich grosse Teilmengen nötig, um eine einelementige Menge zu erhalten. Mit Satz 9.19 folgt

$$|S| \leq \left(2n \cdot 2^{\langle A \rangle + \langle c \rangle + \langle b \rangle - m^2} + 1\right) \cdot 2^{\langle A \rangle - m^2} \leq 2^2 \cdot n \cdot 2^{2\langle A \rangle + \langle c \rangle + \langle b \rangle - 2m^2}$$

und somit auch, dass Schritt (3) höchstens

$$\begin{aligned} N' &\leq \log_2(|S|) + 1 \leq \log_2\left(2^2 \cdot n \cdot 2^{2\langle A \rangle + \langle c \rangle + \langle b \rangle - 2m^2}\right) + 1 \\ &= \log_2 n + \log_2 2^2 + \log_2\left(2^{2\langle A \rangle + \langle c \rangle + \langle b \rangle - 2m^2}\right) + 1 \end{aligned}$$

$$= \log_2 n + 2 + 2\langle A \rangle + \langle c \rangle + \langle b \rangle - 2m^2 + 1$$

viele Schritte benötigt. Diese Anzahl ist polynomiell in $\langle A \rangle + \langle c \rangle + \langle b \rangle$ beschränkt. $\square$

Damit ist die Anzahl der Zulässigkeitsaufrufe polynomiell beschränkt. Das explizite Auflisten von S oder das Bestimmen seines Medians ist dadurch noch nicht in Polynomialzeit gezeigt. Für eine polynomielle Reduktion benutzen wir stattdessen das primal-duale System (9.9).

Satz 9.21. Es gibt einen polynomialen Algorithmus zur Lösung von (LP), wenn es einen polynomialen Algorithmus gibt, der entscheidet ob ein Polyeder $P^{\leq}(A, b)$ mit ganzzahligen Daten leer ist, und falls $P^{\leq}(A, b) \neq \emptyset$, einen Punkt in $P^{\leq}(A, b)$ berechnet.

Beweis. Teste zunächst primale Zulässigkeit. Ist das Problem unzulässig, gib dies aus. Andernfalls teste duale Zulässigkeit. Ist das duale Problem unzulässig, ist das primale Problem unbeschränkt. Sind beide zulässig, ist nach starker Dualität auch (9.9) zulässig; jeder Punkt dieses Systems liefert ein primales Optimum. Alle drei Systeme haben polynomielle Kodierungslänge. $\square$

Für $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{Z}^m$ löst die **Ellipsoidmethode** zunächst folgendes Problem unter der Voraussetzung, dass P beschränkt ist:

1. Finde $x \in P := P^{\leq}(A, b) = \{x \in \mathbb{R}^n | Ax \leq b\}$ oder
2. stelle fest, dass $\overset{o}{P} = \emptyset$.

9.5 Geometrische Grundlagen von Ellipsoiden

Definition 9.22. Eine Menge $E \subset \mathbb{R}^n$ heißt **Ellipsoid**, falls es einen Vektor $a \in \mathbb{R}^n$ und eine symmetrische positiv definite $n \times n$ Matrix A gibt, so dass gilt

$$E = E(A, a) := \{x \in \mathbb{R}^n | (x - a)^T A^{-1} (x - a) \leq 1\}. \quad (9.12)$$

Man beachte, dass für eine positiv definite Matrix A auch stets A^{-1} positiv definit ist. Mit der **Ellipsoid-Norm** auf $\mathbb{R}^n$ definiert durch eine positiv definite $(n \times n)$ -Matrix A :

$$\|x\|_A := \sqrt{x^T A^{-1} x}$$

ergibt sich eine äquivalente Darstellung von $E(A, a)$ durch

$$E(A, a) = \{x \in \mathbb{R}^n | \|x - a\|_A \leq 1\}. \quad (9.13)$$

Somit ist E die Einheitskugel mit Zentrum a im $\mathbb{R}^n$ bezüglich der Norm $\|\cdot\|_A$.

Desweiteren gilt, dass $E(A, a)$ das Bild der Einheitskugel

$$B_1(0) := \{u \in \mathbb{R}^n \mid \|u\|_2 \leq 1\}$$

unter einer **affinen Transformation** ist.

Satz 9.23. Es gilt

$$E(A, a) = f(B_1(0)) := A^{1/2}B_1(0) + a,$$

wobei $A^{1/2}$ die eindeutig bestimmte symmetrische positiv definite Wurzel der symmetrischen positiv definiten Matrix A ist, für die gilt

$$A^{1/2}A^{1/2} = A.$$

Beweis. Wir erhalten

$$\begin{aligned} A^{1/2}B_1(0) + a &= \{A^{1/2}y \in \mathbb{R}^n : y \in B_1(0)\} + a \\ &= \{x \in \mathbb{R}^n : (A^{1/2})^{-1}x \in B_1(0)\} + a \\ &= \{x \in \mathbb{R}^n : ((A^{1/2})^{-1}x)^T((A^{1/2})^{-1}x) \leq 1\} + a \\ &= \{x \in \mathbb{R}^n : x^T(A^{1/2}A^{1/2})^{-1}x \leq 1\} + a \\ &= \{x \in \mathbb{R}^n : x^TA^{-1}x \leq 1\} + a \\ &= \{x + a \in \mathbb{R}^n : x^TA^{-1}x \leq 1\} \\ &= \{z \in \mathbb{R}^n : (z - a)^TA^{-1}(z - a) \leq 1\} \\ &= E(A, a). \end{aligned}$$

□

Ein Ellipsoid hat weitere schöne geometrische Eigenschaften, die man aus den algebraischen Eigenschaften von A ableiten kann:

1. Der **Durchmesser** von $E(A, a)$ (definiert als längste **Achse** von $E(A, a)$) ist gegeben durch $2 \cdot \sqrt{\lambda^{\max}}$, wobei $\lambda^{\max}$ der grösste Eigenwert von A ist.
2. Die **Breite** (definiert als kürzeste **Achse** von $E(A, a)$) ist gegeben durch $2 \cdot \sqrt{\lambda^{\min}}$, wobei $\lambda^{\min}$ der kleinste Eigenwert von A ist.
3. Die **Symmetrieachsen** entsprechen den Eigenvektoren von A .

Diese Beobachtungen implizieren dass $B_{\sqrt{\lambda^{\max}}}(a)$ die kleinste euklidische Kugel (bzgl. des Radius) im $\mathbb{R}^n$ ist, die $E(A, a)$ enthält. Desweiteren ist $B_{\sqrt{\lambda^{\min}}}(a)$ die grösste euklidische Kugel (bzgl. Radius), die in $E(A, a)$ enthalten ist.

Das Ellipsoid in Abb. 9.2 ist definiert durch

$$A = \begin{pmatrix} 16 & 0 \\ 0 & 4 \end{pmatrix}.$$

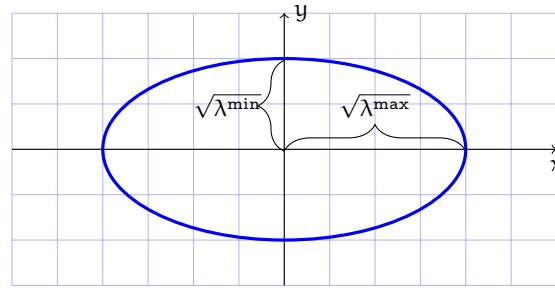


Abbildung 9.2: Ellipsoid $E(A, 0)$ mit größter enthaltener und kleinster umschließender euklidischer Kugel um 0.

Die Eigenwerte sind $\lambda_1 = 16, \lambda_2 = 4$ mit Eigenvektoren $v_1 = (1, 0)^T$ und $v_2 = (0, 1)^T$. Der Durchmesser von $E(A, 0)$ ergibt sich durch $2 \cdot \sqrt{\lambda_{\max}} = 8$ und die Breite durch $2 \cdot \sqrt{\lambda_{\min}} = 4$.

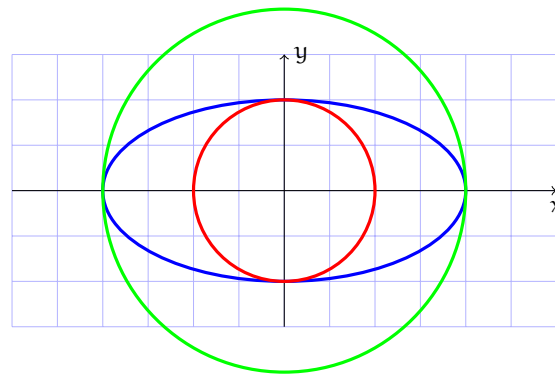


Abbildung 9.3: Ellipsoid $E(A, 0)$ mit “kleinster” und “grösster” Einheitskugel.

Im Folgenden wird das **Volumen** von Ellipsoiden eine große Rolle spielen.

Definition 9.24. Sei $K \subset \mathbb{R}^n$ Lebesgue-messbar. Das Volumen von K ist definiert als

$$\text{Vol}(K) := \int_{x \in K} dx.$$

Im Kontext von Ellipsoiden gilt folgende Eigenschaft.

Satz 9.25. Das Volumen eines Ellipsoids $E(A, a)$ ist gegeben durch

$$\text{Vol}(E(A, a)) = \sqrt{\det A} \cdot V_n,$$

wobei V_n das Volumen der Einheitskugel $B_1(0)$ im $\mathbb{R}^n$ ist.

Beweis. Wir benutzen

$$E(A, a) = f(B_1(0)) = A^{1/2}B_1(0) + a,$$

somit folgt mittels Variablentausch

$$\int_{y \in f(B_1(0))} dy = \int_{x \in B_1(0)} |\det A^{1/2}| dx = \sqrt{\det A} \cdot V_n.$$

□

Für V_n gilt folgende Darstellung:

$$V_n = \frac{\pi^{n/2}}{\Gamma(n/2 + 1)} \approx \frac{1}{\sqrt{\pi n}} \left(\frac{2e\pi}{n} \right)^{n/2}, \quad (9.14)$$

wobei

$$\Gamma(x) := \int_0^\infty e^{-t} t^{x-1} dt, \text{ für } x > 0$$

die Gamma-Funktion bezeichnet.

Wir werden häufiger Abschätzungen von V_n benötigen. Es gilt folgende grobe Abschätzung:

Lemma 9.26.

$$n^{-n} \leq V_n \leq 2^n.$$

Beweis. Es gilt dass $B_1(0)$ die Menge $B^1 := \{x \in \mathbb{R}^n | 0 \leq x_i \leq 1/n, i = 1, \dots, n\}$ enthält. Das Volumen von B^1 ist gegeben durch $\text{Vol}(B^1) = n^{-n}$. Desweiteren ist $B_1(0)$ vollständig in $B^2 := \{x \in \mathbb{R}^n | \|x\|_\infty \leq 1\}$ enthalten. Es gilt $\text{Vol}(B^2) = 2^n$. □

Wir werden nun das Volumenverhältnis von zwei Ellipsen vergleichen. Mit der Formel aus Satz 9.25 folgt für eine invertierbare affine Transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^n, x \mapsto Dx + d$

$$\text{Vol}(T(E(A, a))) = |\det D| \sqrt{\det A} V_n.$$

Somit gilt für das Verhältnis von zwei verschiedenen Ellipsoiden $E(A, a)$ und $E(B, b)$

$$\frac{\text{Vol}(E(A, a))}{\text{Vol}(E(B, b))} = \frac{\text{Vol}(T(E(A, a)))}{\text{Vol}(T(E(B, b)))}. \quad (9.15)$$

Im Weiteren wird die Optimierung einer linearen Funktion über ein Ellipsoid von Interesse sein. Wir werden hierzu die Tatsache benutzen dass eine lineare Funktion $cx, c^T \neq 0$ über einer Einheitskugel $B_1(a)$ mit Zentrum $a \in \mathbb{R}^n$ ihr Maximum im Punkt $a + \frac{c^T}{\|c^T\|_2}$ annimmt.

Satz 9.27. Sei $E(A, a) \subset \mathbb{R}^n$ ein Ellipsoid und sei $c^T \neq 0$. Setze $b := Ac^T / \sqrt{cAc^T}$, $z_{\max} := a + b$ und $z_{\min} := a - b$. Diese Punkte nehmen Maximum bzw. Minimum an. Es gilt

$$\begin{aligned} cz_{\max} &:= \max\{cx | x \in E(A, a)\} = ca + \sqrt{cAc^T} = ca + \left\| c^T \right\|_{A^{-1}}, \\ cz_{\min} &:= \min\{cx | x \in E(A, a)\} = ca - \sqrt{cAc^T} = ca - \left\| c^T \right\|_{A^{-1}}. \end{aligned} \quad (9.16)$$

Beweis. Setze $Q := A^{1/2}$. Somit gilt

$$Q^{-1}E(A, a) = B_1(0) + Q^{-1}a = B_1(Q^{-1}a)$$

und somit folgt

$$\begin{aligned} \max\{cx | x \in E(A, a)\} &= \max\{cQQ^{-1}x | Q^{-1}x \in Q^{-1}E(A, a)\} \\ &= \max\{cQy | y \in B_1(Q^{-1}a)\} \\ &= cQ \frac{1}{\|Qc^T\|_2} Qc^T + cQQ^{-1}a \\ &= c \frac{1}{\sqrt{cAc^T}} Ac^T + ca \\ &= ca + \sqrt{cAc^T}. \end{aligned}$$

Der maximierende Punkt ist $a + Ac^T/\sqrt{cAc^T} = z_{\max}$. Wendet man dieselbe Rechnung auf $-c$ an, erhält man das Minimum in $z_{\min}$.

□

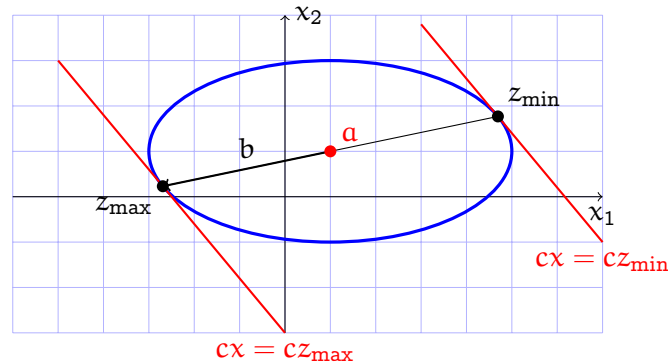


Abbildung 9.4: Maximum und Minimum einer linearen Funktion auf $E(A, a)$, mit $z_{\max} = a + b$ und $z_{\min} = a - b$.

Wir werden nun folgendes klassisches Resultat festhalten (ohne Beweis), dass jede konvexe, kompakte und volldimensionale Menge im $\mathbb{R}^n$ in einem eindeutig definierten Ellipsoid kleinsten Volumens enthalten ist, und, ein eindeutiges Ellipsoid größten Volumens enthält.

Wir wiederholen zunächst noch einmal die Begriffe [affine Unabhängigkeit](#) und [Dimension](#) einer Menge im $\mathbb{R}^n$.

Definition 9.28. Die Punkte $x_1, \dots, x_k$ im $\mathbb{R}^n$ heißen [affin unabhängig](#), falls die Punkte $x_2 - x_1, \dots, x_k - x_1$ linear unabhängig sind. Für $S \subseteq \mathbb{R}^n$ ist der [affine Rang](#) ($\text{arank}(S)$) definiert als die Kardinalität einer größten affin unabhängigen Teilmenge von S . Die Dimension von S ($\dim(S)$) ist definiert als

$$\dim(S) = \text{arank}(S) - 1.$$

| $S \subseteq \mathbb{R}^n$ heisst **volldimensional**, falls $\dim(S) = n$ gilt.

Satz 9.29 (Löwner, John). Für jede konvexe, kompakte und volldimensionale Menge $K \subset \mathbb{R}^n$ existiert ein eindeutiges Ellipsoid $E = E(A, a)$ mit kleinstem Volumen, welches K enthält.

Man bezeichnet das zu K gehörende Ellipsoid als **Löwner-John Ellipsoid**.

9.6 Die Ellipsoid-Methode

Wir werden nun eine geometrische Beschreibung der Ellipsoid-Methode angeben.

Gegeben sei ein Polytop $P := P^{\leq}(C, d)$. Wir wollen ein $x \in P$ finden oder zeigen dass $\overset{o}{P} = \emptyset$ ist.

Vorab prüfen wir alle Nullzeilen von C . Eine verletzte Ungleichung $0 \leq d_i$ zeigt sofort Unzulässigkeit; erfüllte Ungleichungen dieser Form werden gestrichen. Damit sind alle für Updates verwendeten Normalen ungleich null.

1. Konstruiere ein Ellipsoid $E_0 = E(A_0, a_0)$ mit $P \subseteq E_0$; $k \leftarrow 0$.
2. Ist das Volumen von E_k "zu klein", dann STOP: $\overset{o}{P} = \emptyset$.
3. Teste ob $a_k \in P$.
4. Gilt $a_k \in P$, dann STOP.
5. Gilt $a_k \notin P$, dann gibt es eine verletzte Ungleichung

$$ca_k > d_i,$$

wobei c die i -te Zeile der Matrix C ist. Bezeichne mit

$$E'_k := E_k \cap \{x \in \mathbb{R}^n | cx \leq ca_k\}$$

das **Halbellipsoid** von E_k , welches P enthält. Konstruiere das zu E'_k gehörige Löwner-John Ellipsoid und bezeichne es mit E_{k+1} .

6. Setze $k \leftarrow k + 1$ und gehe zu (2).

In dieser geometrischen Variante der Ellipsoid-Methode müssen wir nun noch genauer erklären wie wir das Anfangsellipsoid bzw. das zu dem Halbellipsoid gehörende Löwner-John Ellipsoid konstruieren.

Das Anfangsellipsoid wird eine Kugel mit Zentrum 0 sein.

Enthält das Polytop $P = P^{\leq}(C, d)$ explizite Ungleichungen der Form

$$\ell_i \leq x_i \leq u_i \text{ für } i = 1, \dots, n,$$

so kann

$$R := \max \left\{ 1, \sqrt{\sum_{i=1}^n (\max\{|u_i|, |\ell_i|\})^2} \right\}$$

als Radius der Kugel gewählt werden. Im Allgemeinen gilt folgendes:

Lemma 9.30. Sei $P = \{x \in \mathbb{R}^n | Cx \leq d\}$ ein Polytop und C, d ganzzahlig. Dann gilt $P \subseteq B_R(0)$ wobei $R := \sqrt{n} 2^{\langle C \rangle + \langle d \rangle - n^2}$.

Beweis. Wie im Beweis von Satz 9.17 kann man leicht zeigen dass für jede Ecke v von P gilt:

$$|v_i| \leq 2^{\langle C \rangle + \langle d \rangle - n^2} \text{ für alle } i = 1, \dots, n.$$

Daraus folgt für die Euklidische Norm von v :

$$\|v\|_2 = \sqrt{\sum_{i=1}^n v_i^2} \leq \sqrt{n} \max\{|v_i|\} \leq \sqrt{n} 2^{\langle C \rangle + \langle d \rangle - n^2}.$$

Also ist jede Ecke von P in $B_R(0)$ enthalten und weil P ein Polytop ist, gilt $P \subseteq B_R(0)$. $\square$

Nun werden wir spezifizieren, wie wir in Iteration k der Ellipsoid-Methode ausgehend von einem Ellipsoid E_k und einer verletzten Ungleichung

$$ca_k > d_i$$

das zu

$$E'_k := E_k \cap \{x \in \mathbb{R}^n | cx \leq ca_k\}$$

gehörige Löwner-John Ellipsoid konstruieren.

Wir setzen $n \geq 2$ voraus; für $n = 1$ genügt die Bestimmung eines zulässigen Intervalls.

Wir setzen folgende Werte fest:

$$b := \frac{1}{\sqrt{cA_k c^T}} A_k c^T \tag{9.17}$$

$$a_{k+1} := a_k - \frac{1}{n+1} b \tag{9.18}$$

$$A_{k+1} := \frac{n^2}{n^2 - 1} \left(A_k - \frac{2}{n+1} b b^T \right). \tag{9.19}$$

Das so definierte neue Ellipsoid $E_{k+1} = E_{k+1}(A_{k+1}, a_{k+1})$ ist genau das gesuchte Löwner-John Ellipsoid. Geometrisch kann die Konstruktion von E_{k+1} folgendermaßen interpretiert werden. Mit $c \neq 0$, kann das Maximum und Minimum von cx über E_k explizit angegeben

werden, siehe Satz 9.27. Nun sieht man leicht, dass der Mittelpunkt des neuen Ellipsoids E_{k+1} auf dem Geradenstück zwischen a_k und $z_{\min}$ liegt. Die Länge des Geradenstücks ist $\|b\|_2$ und a_{k+1} erreicht man von a_k aus, indem man einen Schritt der Länge $\frac{1}{n+1} \|b\|_2$ in Richtung $-b$ macht. Der Punkt $z_{\min}$ ist enthalten im Durchschnitt des Randes des Ellipsoids E_{k+1} mit dem Rand von E'_k .

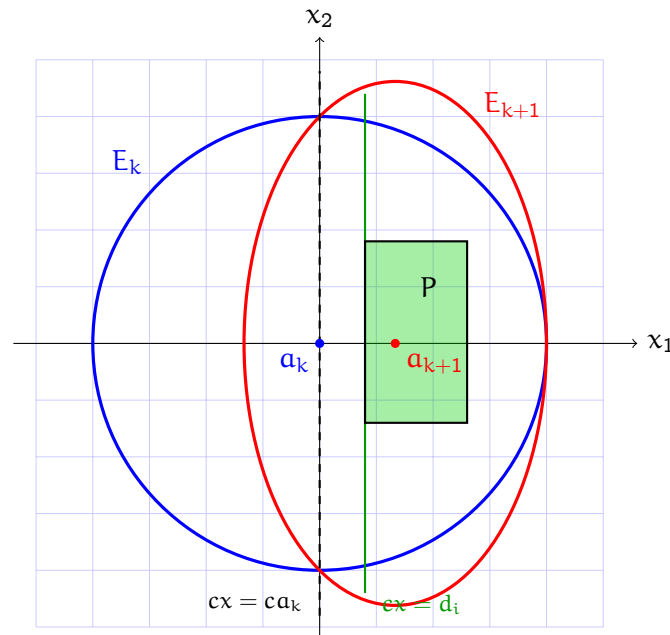


Abbildung 9.5: Ein Ellipsoidschritt für $n = 2$, $A_k = I$, $a_k = 0$ und $c = (-1, 0)$. Das neue Ellipsoid hat Mittelpunkt $(1/3, 0)$ und Halbachsen $2/3$ und $2/\sqrt{3}$.

Wir müssen zeigen, dass die Ellipsoid Methode nach polynomiell vielen Schritten terminiert. Dazu werden wir ein Volumenargument benutzen: In jeder Iteration konstruieren wir eine Ellipse mit strikt geringerem Volumen. Die **Geschwindigkeit** der Volumenabnahme werden wir im Folgenden untersuchen.

Lemma 9.31. Sei $E_k = E(A_k, a_k)$ ein Ellipsoid und $E_{k+1} = E(A_{k+1}, a_{k+1})$ das Löwner-John Ellipsoid, welches durch die Formeln (9.17), (9.18), (9.19) konstruiert wurde, wobei $c \neq 0$ gilt. Dann gilt

$$\frac{\text{Vol}(E_{k+1})}{\text{Vol}(E_k)} = \left(\left(\frac{n}{n+1} \right)^2 \left(\frac{n^2}{n^2-1} \right)^{n-1} \right)^{1/2} < e^{-1/(2(n+1))} < 1.$$

Beweis. Wir werden zunächst den Fall untersuchen dass das Ellipsoid E_k die Einheitskugel ist und dass der Vektor c gegeben ist durch

$$c = (-1, 0, \dots, 0).$$

Somit erhalten wir durch Benutzung von (9.17),(9.18),(9.19):

$$\begin{aligned} \mathbf{b} &= (-1, 0, \dots, 0)^\top \\ \mathbf{a}_{k+1} &= \mathbf{a}_k - \frac{1}{n+1} \mathbf{b} = \left(\frac{1}{n+1}, 0, \dots, 0 \right)^\top \\ \mathbf{A}_{k+1} &= \frac{n^2}{n^2-1} \left(\mathbf{I} - \frac{2}{n+1} (-1, 0, \dots, 0)^\top (-1, 0, \dots, 0) \right) \\ &= \text{diag} \left(\frac{n^2}{(n+1)^2}, \frac{n^2}{n^2-1}, \dots, \frac{n^2}{n^2-1} \right). \end{aligned}$$

Das neue Ellipsoid enthält die Halbkugel: Für $0 \leq x_1 \leq 1$ und $\sum_{i=2}^n x_i^2 \leq 1 - x_1^2$ gilt

$$\frac{(n+1)^2}{n^2} \left(x_1 - \frac{1}{n+1} \right)^2 + \frac{n^2-1}{n^2} \sum_{i=2}^n x_i^2 \leq 1 + \frac{2(n+1)}{n^2} (x_1^2 - x_1) \leq 1.$$

Für das Volumenverhältnis der beiden Ellipsoiden $E_k = E(\mathbf{I}, 0)$ und E_{k+1} erhalten wir $E_k := E(\mathbf{I}, 0)$ und E_{k+1} :

$$\begin{aligned} \frac{\text{Vol}(E_{k+1})}{\text{Vol}(E_k)} &= \frac{\sqrt{\det \mathbf{A}_{k+1}} V_n}{\sqrt{\det \mathbf{A}_k} V_n} = \sqrt{\det \mathbf{A}_{k+1}} = \left(\frac{n^{2n}}{(n+1)^2 (n^2-1)^{n-1}} \right)^{1/2} \\ &= \left(\frac{n}{n+1} \right) \left(\frac{n^2}{n^2-1} \right)^{\frac{n-1}{2}}. \end{aligned}$$

Wir formen nun um und schätzen ab:

$$\begin{aligned} \left(\frac{n}{n+1} \right) \left(\frac{n^2}{n^2-1} \right)^{\frac{n-1}{2}} &= \left(1 + \frac{-1}{n+1} \right) \left(1 + \frac{1}{n^2-1} \right)^{\frac{n-1}{2}} \\ &\leq \left(e^{\frac{-1}{n+1}} \right) \left(e^{\frac{1}{n^2-1}} \right)^{\frac{n-1}{2}} \\ &= \left(e^{\frac{-1}{n+1}} \right) \left(e^{\frac{1}{2(n+1)}} \right) \\ &= e^{\frac{-1}{2(n+1)}}, \end{aligned} \tag{9.20}$$

wobei (9.20) durch $1+x \leq e^x$ für alle $x \in \mathbb{R}$ folgt. (Man überprüfe, dass die Funktion $f(x) := 1+x - e^x$ ihr globales Maximum in $x=0$ annimmt.)

Nun betrachten wir den allgemeinen Fall mit beliebigem E_k und $c \neq 0$. Wir werden den allgemeinen Fall durch Angabe einer geeigneten affinen Transformation von E_k und E_{k+1} auf den bisher betrachteten Fall zurückführen.

Sei $F := B_1(0)$ und F' das Ellipsoid, welches wie im ersten Fall durch anwenden der Regeln (9.17),(9.18),(9.19) aus F entsteht, wobei hier noch $c = (-1, 0, \dots, 0)$ vorausgesetzt wird.

Mit Satz 9.23 folgt, dass E_k die Form

$$E_k(\mathbf{A}_k, \mathbf{a}_k) = \mathbf{A}_k^{1/2} B_1(0) + \mathbf{a}_k = \mathbf{A}_k^{1/2} F + \mathbf{a}_k$$

hat.

Aus der linearen Algebra wissen wir, dass eine orthogonale Matrix Q existiert, welche $A_k^{1/2}c^\top$ in ein Vielfaches des Vektors $(-1, 0, \dots, 0)$ rotiert, d.h.

$$(-1, 0, \dots, 0)^\top = \frac{1}{\|A_k^{1/2}c^\top\|_2} Q A_k^{1/2} c^\top.$$

Die gesuchte (bijektive) affine Transformation hat also die Form

$$T(x) = A_k^{1/2} Q^\top x + a_k,$$

mit Inversem

$$T^{-1}(x) = Q A_k^{-1/2} (x - a_k).$$

Man prüfe nun nach dass gilt

$$T(F) = E_k \text{ und } T(F') = E_{k+1}.$$

Nun benutzen wir dass das **relative** Volumenverhältnis von zwei Ellipsoiden nach einer affinen Transformation konstant bleibt (siehe (9.15)) und erhalten für die beiden Ellipsoide

$$\frac{\text{Vol}(E_{k+1})}{\text{Vol}(E_k)} = \frac{\text{Vol}(T(F'))}{\text{Vol}(T(F))} = \frac{\text{Vol}(F')}{\text{Vol}(F)} = \left(\frac{n}{n+1} \right) \left(\frac{n^2}{n^2-1} \right)^{\frac{n-1}{2}} \leq e^{\frac{-1}{2(n+1)}}.$$

□

Wir müssen nun auch zeigen, dass wir nicht “zu viele” Iterationen benötigen, um zu entscheiden ob das Polytop P leer ist oder nicht. Dazu werden wir im Folgenden eine untere Schranke an das Mindestvolumen eines volldimensionalen Polytops herleiten.

Lemma 9.32. Sei $P = \{x \in \mathbb{R}^n | Cx \leq d\}$ ein volldimensionales Polytop und die Matrix C sowie der Vektor d ganzzahlig. Dann gilt

$$\text{Vol}(P) \geq 2^{-(n+1)\langle C \rangle + n^3}.$$

Beweis. Ein volldimensionales Polytop in $\mathbb{R}^n$ enthält $n+1$ **Ecken**, die affin unabhängig sind.

Die konvexe Hülle dieser Punkte spannt ein Simplex $S := \text{co}(\{v_0, \dots, v_n\}) \subseteq P$ auf. Somit ist das Volumen von P mindestens so groß wie das von S . Für das Volumen von Simplexes gibt es eine geschlossene Formel

$$\text{Vol}(S) = \frac{1}{n!} \left| \det \begin{pmatrix} 1 & \dots & 1 \\ v_0 & \dots & v_n \end{pmatrix} \right|.$$

(Diese Formel ergibt sich wiederum durch den Transformationssatz, da S eine affine Transformation des Einheitssimplex $S_n := \{x \in \mathbb{R}_+^n | \sum_{i=1}^n x_i \leq 1\}$ ist mit Volumen $\text{Vol}(S_n) = \frac{1}{n!}$.)

Nun benutzen wir, dass für jede Ecke v_i des Polytops P die j -te Komponente die Form $\frac{\det B_{ij}}{\det B_i}$ hat, wobei B_i eine reguläre Teilmatrix von C ist und B_{ij} eine Teilmatrix von (C, d) ist. Wir erhalten (unter Benutzung von $\det(A \cdot B) = \det A \cdot \det B$ für alle $n \times n$ Matrizen A, B):

$$\begin{aligned} \left| \det \begin{pmatrix} 1 & \dots & 1 \\ v_0 & \dots & v_n \end{pmatrix} \right| &= \left| \det \begin{pmatrix} \det B_0 & \dots & \det B_n \\ \det B_{01} & \dots & \det B_{n1} \\ & \dots & \\ \det B_{0n} & \dots & \det B_{nn} \end{pmatrix} \operatorname{diag} \left(\frac{1}{\det B_0}, \dots, \frac{1}{\det B_n} \right) \right| \\ &= \left| \frac{1}{\det B_0 \cdots \det B_n} \right| \left| \det \begin{pmatrix} \det B_0 & \dots & \det B_n \\ u_0 & \dots & u_n \end{pmatrix} \right|, \end{aligned}$$

wobei $u_0, \dots, u_n$ Vektoren im $\mathbb{Z}^n$ sind.

Wir wissen dass der zweite Ausdruck eine positive ganze Zahl grösser gleich 1 ist. Wir müssen somit nur $\left| \frac{1}{\det B_0 \cdots \det B_n} \right|$ nach unten abschätzen. Aus Lemma 9.13 folgt

$$|\det B_j| \leq 2^{\langle B_j \rangle - n^2} - 1 \leq 2^{\langle C \rangle - n^2} \text{ für alle } j = 0, \dots, n,$$

somit gilt

$$\left| \det \begin{pmatrix} 1 & \dots & 1 \\ v_0 & \dots & v_n \end{pmatrix} \right| \geq \left| \frac{1}{\det B_0 \cdots \det B_n} \right| \geq \left(2^{\langle C \rangle - n^2} \right)^{-(n+1)}.$$

Mit der Ungleichung $n! \leq 2^{n^2}$ gilt schlussendlich

$$\operatorname{Vol}(P) \geq \operatorname{Vol}(S) \geq \frac{1}{n!} \left(2^{\langle C \rangle - n^2} \right)^{-(n+1)} \geq 2^{-(n+1)\langle C \rangle + n^3}.$$

□

Zuletzt schätzen wir die Anzahl der Iterationen der Ellipsoid-Methode ab. Nach Lemma 9.31 sinkt das Volumen bei jedem Update um einen Faktor von höchstens $e^{-1/(2(n+1))}$.

Lemma 9.33. Sei $P = \{x \in \mathbb{R}^n \mid Cx \leq d\}$ ein volldimensionales Polytop mit ganzzahligen Daten und $n \geq 2$. Setze $T := \langle C \rangle + \langle d \rangle$ und $E_0 = B_R(0)$ mit $R := \sqrt{n} 2^{T-n^2}$. Nach

$$N := (2n+2)((n+1)\langle C \rangle + nT)$$

Ellipsoid-Updates gilt

$$\operatorname{Vol}(E_N) < 2^{-N/(2n+2)+nT} \leq \operatorname{Vol}(P).$$

Beweis. Mit Lemma 9.30 und $n \geq 2$ erhalten wir

$$\begin{aligned}\text{Vol}(E_0) &\leq (2R)^n = 2^{n(T-n^2+1+\log_2(n)/2)} < 2^{nT}, \\ \text{Vol}(E_N) &< e^{-N/(2n+2)} \text{Vol}(E_0) < 2^{-N/(2n+2)+nT}.\end{aligned}$$

Die gewählte Schranke erfüllt

$$N \geq (2n+2)((n+1)\langle C \rangle - n^3 + nT).$$

Damit folgt aus der unteren Volumenschranke

$$\text{Vol}(E_N) < 2^{-(n+1)\langle C \rangle + n^3} \leq \text{Vol}(P).$$

□

9.7 Polynomialität der Ellipsoid-Methode

Wir beschreiben die Methode für ein Polytop $P = \{x \in \mathbb{R}^n \mid Cx \leq d\}$ mit ganzzahligen Daten und $n \geq 2$. Für $n = 1$ genügt die Bestimmung eines zulässigen Intervalls. Erfüllte Ungleichungen mit Nullnormalen werden vorab entfernt; eine verletzte Ungleichung dieser Form zeigt Unzulässigkeit. Die etwas größere Iterationsschranke ist auch bei einer leeren zulässigen Menge positiv.

1. Initialisierung:

- (a) $k := 0$,
- (b) $N := (2n + 2)((n + 1)\langle C \rangle + n(\langle C \rangle + \langle d \rangle))$,
- (c) $A_0 := R^2 I$ mit $R := \sqrt{n} 2^{\langle C \rangle + \langle d \rangle - n^2}$,
- (d) $a_0 := 0$.

2. Allgemeiner Schritt:

- (a) Ist $k = N$, STOP: $\overset{\circ}{P} = \emptyset$. Ist zugesichert, dass P entweder leer oder volldimensional ist, folgt $P = \emptyset$.
- (b) Ist $a_k \in P$, gib a_k aus und STOP.
- (c) Andernfalls wähle eine verletzte Ungleichung $ca_k > \gamma$ und setze

$$\begin{aligned}
 b &:= \frac{A_k c^T}{\sqrt{c^T A_k c}}, \\
 a_{k+1} &:= a_k - \frac{b}{n+1}, \\
 A_{k+1} &:= \frac{n^2}{n^2 - 1} \left(A_k - \frac{2}{n+1} b b^T \right), \\
 k &:= k + 1.
 \end{aligned}$$

- (d) Gehe zu Schritt 2.

Die Korrektheit folgt aus den bisherigen Lemmata. Die Initialisierung und der Zulässigkeitstest benötigen jeweils $O(mn)$ arithmetische Operationen. Die Berechnung von b benötigt $O(n^2)$ Operationen und eine Wurzel; die Updates von a_k und A_k benötigen $O(n)$ bzw. $O(n^2)$ Operationen. Es gibt höchstens $N = O(n^2(\langle C \rangle + \langle d \rangle))$ Updates. Die grobe Gesamtschranke $O(mn^4(\langle C \rangle + \langle d \rangle))$ ist somit polynomiell. Hier zählen Wurzeln idealisiert als einzelne Operationen. Für polynomielle Bitlaufzeit sind zusätzlich rationale Approximationen und polynomiell beschränkte Kodierungslängen erforderlich; darauf gehen wir am Ende ein.

9.8 Nicht-Volldimensionale Polytope

Sei $P = \{x \in \mathbb{R}^n \mid Cx \leq d\}$ ein Polytop mit $C \in \mathbb{Z}^{m \times n}$ und $d \in \mathbb{Z}^m$. Definiere

$$P(\epsilon) := \{x \in \mathbb{R}^n \mid Cx \leq d + \epsilon e\}, \quad (9.21)$$

wobei $e = (1, \dots, 1)^T$ und

$$\epsilon := \frac{2^{-(\langle C \rangle + \langle d \rangle + 1)}}{n+1}. \quad (9.22)$$

Lemma 9.34. Es gilt $P = \emptyset$ genau dann, wenn $P(\epsilon) = \emptyset$. Ist P nichtleer, so ist $P(\epsilon)$ volldimensional.

Beweis. Wegen $P \subseteq P(\epsilon)$ genügt für die erste Aussage die Richtung $P = \emptyset \Rightarrow P(\epsilon) = \emptyset$. Ist $P = \emptyset$, ist $\min\{0x \mid Cx \leq d\}$ unzulässig. Sein duales Problem

$$\max\{-\lambda d \mid \lambda C = 0, \lambda \geq 0\}$$

ist wegen der zulässigen Lösung $\lambda = 0$ und Satz 6.7 unbeschränkt. Daher ist das System

$$\lambda C = 0, \quad \lambda d = -1, \quad \lambda \geq 0$$

nichtleer. Entferne abhängige Gleichungen; der verbleibende Rang sei $r \leq n+1$. Wähle eine zulässige Basislösung λ .

Setze $T := \langle C \rangle + \langle d \rangle$. Nach der Cramerschen Regel sind die Basiskomponenten Quotienten von Determinanten ganzzahliger $r \times r$ -Matrizen. Der Nenner hat Betrag mindestens 1. Im Zähler wird eine Spalte durch die rechte Seite ersetzt, deren einzige mögliche Nichtnullkomponente -1 ist. Entwicklung nach dieser Spalte liefert eine Unterdeterminante von $(C, d)^T$. Mit Lemma 9.13 folgt

$$0 \leq \lambda_i \leq 2^T.$$

Für $r = 1$ ist der Zähler 0 oder ± 1 , so dass dieselbe Schranke gilt. Höchstens $r \leq n+1$ Komponenten sind positiv. Somit erhalten wir

$$\lambda(d + \epsilon e) = -1 + \epsilon \sum_i \lambda_i \leq -1 + \epsilon(n+1)2^T = -\frac{1}{2} < 0.$$

Ein Punkt $x \in P(\epsilon)$ würde dagegen $0 = \lambda Cx \leq \lambda(d + \epsilon e) < 0$ erfüllen. Also ist $P(\epsilon) = \emptyset$.

Sei nun $x \in P$. Mit

$$\bar{\epsilon} := \frac{\epsilon}{\max\{1, \max_i \sum_j |c_{ij}|\}}$$

liegt der Würfel $S = \{y \mid |y_j - x_j| \leq \bar{\epsilon}, j = 1, \dots, n\}$ in $P(\epsilon)$, denn für jede Zeile gilt

$$c_i y \leq c_i x + \bar{\epsilon} \sum_j |c_{ij}| \leq d_i + \epsilon.$$

Wegen $\text{Vol}(S) = (2\bar{\epsilon})^n > 0$ ist $P(\epsilon)$ volldimensional. □

Ist P nichtleer und beschränkt, gilt $Cv \leq 0$ nur für $v = 0$. Eine Störung der rechten Seite ändert diese Bedingung nicht, so dass auch $P(\epsilon)$ beschränkt ist. Ist P leer, ist nach dem Lemma auch $P(\epsilon)$ leer. Vor der Anwendung der Ellipsoid-Methode auf $P(\epsilon)$ beseitigen wir die Nenner der rechten Seite und berechnen Radius und Iterationsschranke aus dem neuen ganzzahligen System. Dessen Kodierungslänge bleibt polynomiell.

9.9 Bestimmung eines exakten zulässigen Punktes

Ein Punkt in $P(\epsilon)$ muss nicht in P liegen. Bisher haben wir also eine exakte **Entscheidung** über Zulässigkeit erhalten. Daraus gewinnen wir wie folgt einen Punkt des ursprünglichen Polytops.

Sei $P \neq \emptyset$. Beginne mit $F = P$ und betrachte nacheinander jede ursprüngliche Ungleichung $c_i x \leq d_i$. Teste, ob

$$F \cap \{x \mid c_i x = d_i\} \neq \emptyset.$$

Falls ja, behalte die Gleichung in der Beschreibung von F ; andernfalls bleibt F unverändert. Gleichungen werden beim Aufruf des Zulässigkeitsverfahrens als zwei Ungleichungen dargestellt.

Am Ende ist F weiterhin ein nichtleeres Polytop. Jede abgelehnte Gleichung gehört zu einer Ungleichung, die auf dem gesamten endgültigen F strikt gilt. Wähle eine Ecke von F . Ihre aktiven Nebenbedingungen haben Rang n und stammen daher ausschließlich aus den beibehaltenen Gleichungen. Diese bestimmen folglich einen eindeutigen Punkt von P , den wir durch Gaußelimination berechnen. Es sind höchstens m weitere Zulässigkeitstests an Systemen polynomialer Kodierungslänge erforderlich.

Ein allgemeines rationales Zulässigkeitssystem wird durch Aufspalten freier Variablen und Einführen von Schlupfvariablen in Standardform gebracht. Nach Entfernen redundanter Gleichungen liefert Satz 9.17 Schranken, die bei Zulässigkeit mindestens eine zulässige Basislösung erhalten. Damit reduziert sich das Problem auf den Polytopfall. Zusammen mit der primal-dualen Reduktion folgt ein polynomialer Algorithmus zur linearen Optimierung, sofern die Updates mit der nachstehend beschriebenen Genauigkeit ausgeführt werden.

9.10 Implementation

Die Initialisierung benutzt eine Wurzel für R , benötigt zur Darstellung des Anfangsellipsoids aber nur den rationalen Wert R^2 . Die Wurzel im allgemeinen Schritt erfordert mehr Sorgfalt. Einfaches Abschneiden von Nachkommastellen kann die Invariante verletzen, dass das neue Ellipsoid das Polytop enthält. Stattdessen verwendet man ein geringfügig größeres Ellipsoid mit rationalen Daten polynomialer Kodierungslänge. Dabei muss der zusätzliche Volumenanstieg gegen die Genauigkeit abgewogen werden. Eine solche Implementierung ist möglich; die technischen Details werden hier nicht bewiesen.

Kapitel 10

Lineare Ganzzahlige Optimierung

In diesem letzten Abschnitt beschäftigen wir uns (sehr oberflächlich) mit der linearen Optimierung über Teilmengen $S \subseteq \mathbb{Z}^n$. Der fundamentale Unterschied zur linearen Optimierung besteht in der Forderung dass eine zulässige Lösung $x \in S$ nur **ganzzahlige** Einträge haben darf. Diese Annahme “zerstört” alle wesentlichen topologischen Eigenschaften der Lösungsmenge, die wir in vorherigen Abschnitten benutzt hatten um z.B. die Dualitätstheorie zu entwickeln und auch die Korrektheit des Simplexverfahrens zu zeigen.

Allgemein betrachten wir folgende Problemklasse.

$$\begin{aligned} &\text{minimiere } cx \\ &Ax \leq b \\ &x \in \mathbb{Z}_+^n, \end{aligned} \tag{IP}$$

wobei $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$, $c^T \in \mathbb{Q}^n$. Die zulässige Menge kann folgendermassen geschrieben werden

$$S := \{x \in \mathbb{R}_+^n : Ax \leq b\} \cap \mathbb{Z}_+^n.$$

Wir schreiben $z_{IP} := \inf\{cx \mid x \in S\}$, mit Wert $+\infty$ bei $S = \emptyset$ und $-\infty$ bei Unbeschränktheit nach unten. Ein endlicher Wert wird angenommen: Nach Multiplikation mit einem gemeinsamen Nenner von c sind alle Zielfunktionswerte ganzzahlig. In diesem Fall schreiben wir $z_{IP} = \min\{cx \mid x \in S\}$. Auch im Folgenden hat ein leeres Minimierungsproblem den Wert $+\infty$.

Beispiel 10.1. Wir betrachten ein Beispiel, das wir auch schon in der Einführung zur linearen Optimierung gesehen hatten. Die Menge S sei gegeben durch

$$S := \{x \in \mathbb{R}^n \mid Ax \leq b, x \geq 0\} \cap \mathbb{Z}^n,$$

wobei $Ax \leq b, x \geq 0$ wie unten definiert ist.

$$\begin{aligned}
2x_1 + x_2 &\leq 8 \\
4x_1 + 5x_2 &\leq 20 \\
6x_1 + 15x_2 &\leq 45 \\
x_1, x_2 &\geq 0
\end{aligned}$$

Die Menge S lässt sich nun leicht angeben, siehe unten.

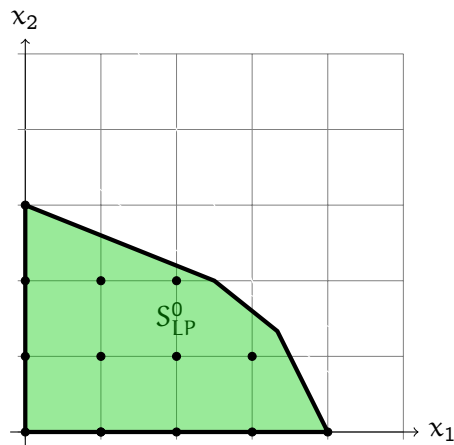


Abbildung 10.1: Die Schattierung zeigt die LP-Relaxierung; ihre Gitterpunkte bilden S .

10.1 Totale duale Ganzzahligkeit und ganzzahlige Programme

Bei bestimmten ganzzahligen Programmen liefert bereits die LP-Relaxierung eine ganzzahlige Optimallösung. Wir geben nun eine hinreichende Bedingung an, die auf dem dualen Problem beruht. Wie bisher bezeichnet c einen Zeilenvektor.

10.1.1 Total dual ganzzahlige Systeme

Betrachte ein rationales System $Ax \leq b$ mit $A \in \mathbb{Q}^{m \times n}$ und $b \in \mathbb{Q}^m$ sowie das primal-duale Paar

$$\begin{aligned}
(P_c) \quad & \max\{cx \mid Ax \leq b\}, \\
(D_c) \quad & \min\{\lambda b \mid \lambda A = c, \lambda \geq 0\}.
\end{aligned} \tag{10.1}$$

Hier ist $x \in \mathbb{R}^n$ frei. Eine Bedingung $x \geq 0$ muss bei der folgenden Definition durch die Zeilen $-Ix \leq 0$ in das System aufgenommen werden.

Definition 10.2. Das System $Ax \leq b$ heißt total dual ganzzahlig (englisch **totally dual integral**, TDI), wenn für jeden Vektor $c \in \mathbb{Z}^n$, für den (P_c) ein endliches Optimum besitzt, (D_c) eine Optimallösung $\lambda \in \mathbb{Z}_+^m$ hat.

Gefordert ist die Existenz einer ganzzahligen dualen Optimallösung, nicht die Ganzzahligkeit jedes dual zulässigen Punktes. TDI ist eine Eigenschaft des gesamten Ungleichungssystems, einschließlich der rechten Seite b und der gewählten Darstellung. Die Bezeichnung “TDI-Matrix” kann daher irreführend sein.

Beispiel 10.3. Die beiden Systeme

$$-x \leq 0, \quad x \leq 1 \quad \text{und} \quad -x \leq 0, \quad 2x \leq 2$$

beschreiben das Intervall $[0, 1]$. Das erste ist TDI: Für ganzzahliges $c \geq 0$ wähle den Multiplikator c für $x \leq 1$, für $c < 0$ den Multiplikator $-c$ für $-x \leq 0$. Das zweite System ist nicht TDI. Für $c = 1$ lautet sein duales Problem

$$\min\{2\lambda_2 \mid -\lambda_1 + 2\lambda_2 = 1, \lambda_1, \lambda_2 \geq 0\}.$$

Seine eindeutige Optimallösung ist $(0, 1/2)$. Selbst ein Intervall mit ganzzahligen Endpunkten kann also eine Darstellung besitzen, die nicht TDI ist.

10.1.2 Ganzzahlige rechte Seiten und ganzzahlige Optima

Satz 10.4 (Ganzzahligkeitsfolgerung aus TDI). Sei $Ax \leq b$ TDI und $b \in \mathbb{Z}^m$. Dann ist jede Ecke von $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ ganzzahlig. Ist P nichtleer und spitz, besitzt jede lineare Zielfunktion mit endlichem Optimum auf P eine ganzzahlige optimale Ecke.

Beweis. Für jeden ganzzahligen Zeilenvektor h mit endlichem Maximum auf P liefern TDI und starke Dualität

$$\max_{x \in P} hx = \lambda b \in \mathbb{Z}$$

mit einem ganzzahligen optimalen Multiplikator λ . Sei v eine Ecke. Wähle n linear unabhängige aktive Zeilen und multipliziere jede mit einer positiven ganzen Zahl, so dass ganzzahlige Zeilen $a'_1, \dots, a'_n$ entstehen. Diese Skalierung dient nur der Konstruktion von Zielfunktionen; das ursprüngliche TDI-System bleibt erhalten. Setze $h := \sum_i \alpha_i a'_i$. Da alle Koeffizienten positiv sind, maximiert v die Funktion hx eindeutig auf P . Für eine Koordinate j schreibe $e_j^T = \sum_i \alpha_i a'_i$. Wähle eine ganze Zahl N mit $N + \alpha_i > 0$ für alle i . Dann maximiert v auch $(Nh + e_j^T)x$ eindeutig. Beide Zielfunktionsvektoren sind ganzzahlig, also gilt

$$v_j = (Nh + e_j^T)v - Nhv \in \mathbb{Z}.$$

Dies gilt für jedes j . Die letzte Aussage folgt durch Wahl einer optimalen Ecke des spitzen Polyeders. $\square$

Dass ein spitzes rationales Polyeder bei beschränkter Zielfunktion eine optimale Ecke besitzt, folgt durch Umformung in Standardform und Satz 4.19, anschließend durch Korollar 4.17. Ist insbesondere das **vollständige** System $Ax \leq b, x \geq 0$ TDI, b ganzzahlig

und das Optimum der Relaxierung endlich, so gilt

$$\min\{cx \mid Ax \leq b, x \in \mathbb{Z}_+^n\} = \min\{cx \mid Ax \leq b, x \geq 0\}. \quad (10.2)$$

Dies gilt auch für nichtganzzahliges c , da die Ganzzahligkeit der Ecken unabhängig von der Zielfunktion ist. Ein beliebiger Optimalpunkt des LPs kann bei mehreren optimalen Ecken dennoch fraktional sein; man wählt daher eine optimale Ecke.

Bemerkung 10.5. Die Voraussetzung an b ist wesentlich. Das System $-x \leq 0, x \leq 1/2$ ist nach demselben Multiplikatorargument TDI. Sein Maximum für $c = 1$ ist jedoch $1/2$, während das zugehörige IP den Wert 0 hat.

10.1.3 Zusammenhang mit total unimodularen Matrizen

Definition 10.6. Eine Matrix A heißt total unimodular (TU), wenn die Determinante jeder quadratischen Untermatrix in $\{-1, 0, 1\}$ liegt.

Insbesondere liegen ihre Einträge in $\{-1, 0, 1\}$. Transponieren, Vorzeichenwechsel einer Zeile sowie Hinzufügen von Zeilen aus I oder $-I$ erhalten totale Unimodularität.

Satz 10.7. Ist A TU, so ist $Ax \leq b$ für jede rationale rechte Seite b TDI. Bei ganzzahligem b sind somit alle Ecken von $\{x \mid Ax \leq b\}$ ganzzahlig.

Beweis. Sei c ganzzahlig und das primale Optimum endlich. Die dual zulässige Menge

$$Q = \{\lambda^T \geq 0 \mid A^T \lambda^T = c^T\}$$

ist in Standardform und besitzt eine optimale Basislösung. Entferne abhängige Gleichungen. Die Basiskomponenten ergeben sich durch Lösen einer regulären quadratischen Untermatrix von A^T mit ganzzahliger rechter Seite. Ihre Determinante ist 1 oder -1 ; die Cramersche Regel liefert daher ganzzahlige Komponenten. Dies zeigt TDI. Die Folgerung ergibt sich aus dem vorigen Satz. $\square$

Die Umkehrung gilt nicht. Das System $-x \leq 0, x \leq 1, 2x \leq 2$ ist TDI: Die ersten beiden Zeilen liefern bereits ganzzahlige duale Optima, und der dritte Multiplikator kann null sein. Wegen des Eintrags 2 ist die Matrix aber nicht TU. Ersetzt man nur die letzte rechte Seite durch 1, geht TDI verloren. Für $c = 1$ ist das Maximum $1/2$, dual erreicht durch den Multiplikator $1/2$ für $2x \leq 1$. Kein ganzzahliger Multiplikatorvektor hat diesen Zielfunktionswert. Dieselbe Matrix kann also in TDI- und Nicht-TDI-Systemen auftreten.

10.1.4 Anwendung: Bipartites Matching

Sei $G = (U \cup W, E)$ ein bipartiter Graph. Ein Matching ist eine Kantenmenge, in der keine zwei Kanten einen Endknoten gemeinsam haben. Mit Kantengewichten w_e lautet

das Problem eines Matchings maximalen Gewichts

$$\max \left\{ \sum_{e \in E} w_e x_e \mid \sum_{e \ni v} x_e \leq 1 \ (v \in U \cup W), \quad x_e \in \mathbb{Z}_+ \ (e \in E) \right\}. \quad (10.3)$$

Die Gradbedingungen implizieren $x_e \leq 1$. Die ganzzahligen zulässigen Punkte sind daher genau die Inzidenzvektoren von Matchings.

Sei M die Knoten-Kanten-Inzidenzmatrix. Multipliziere die Zeilen aus W mit -1 . Jede Spalte hat dann einen Eintrag $+1$ und einen Eintrag -1 . Betrachte eine quadratische Untermatrix. Hat eine Spalte höchstens einen Nichtnulleintrag, entwickle die Determinante nach dieser Spalte und benutze Induktion. Andernfalls enthält jede Spalte beide Vorzeichen, die Zeilensumme ist null und die Determinante verschwindet. Damit ist M TU, ebenso die vollständige Matrix $(M^T, -I)^T$. Die Relaxierung von (10.3) hat somit eine ganzzahlige optimale Ecke.

Bei ganzzahligen Gewichten liefert TDI auch ganzzahlige optimale duale Variablen in

$$\min \left\{ \sum_{v \in U \cup W} y_v \mid y_u + y_v \geq w_{uv} \ (uv \in E), \quad y \geq 0 \right\}. \quad (10.4)$$

Die Multiplikatoren der Zeilen $-x \leq 0$ entsprechen dabei den nichtnegativen Schlupfvariablen dieser dualen Ungleichungen. Die Gleichheit der beiden ganzzahligen Optima ist eine ganzzahlige Min-Max-Beziehung.

Beispiel 10.8. Seien $U = \{u_1, u_2\}$, $W = \{v_1, v_2\}$ und alle vier Kanten vorhanden, mit Gewichten

	v_1	v_2
u_1	5	2
u_2	4	6

Das Matching $\{u_1v_1, u_2v_2\}$ hat Gewicht 11. Die duale Belegung

$$y_{u_1} = 5, \quad y_{u_2} = 4, \quad y_{v_1} = 0, \quad y_{v_2} = 2$$

ist zulässig: Die vier Summen an den Endknoten sind 5, 7, 4, 6. Ihr Wert ist ebenfalls 11. Schwache Dualität beweist daher, dass das angegebene Matching sowohl für das LP als auch für das IP optimal ist.

Bei einem explizit gegebenen rationalen TDI-System mit ganzzahliger rechter Seite kann lineare Optimierung somit die Suche nach ganzzahligen Punkten ersetzen. TDI allein liefert jedoch weder einen Erkennungsalgorithmus noch eine kurze Darstellung eines Systems mit exponentiell vielen Ungleichungen. Fehlen diese strukturellen Voraussetzungen, benötigt man die allgemeinen Verfahren der folgenden Abschnitte.

Bemerkung 10.9. Der Ganzzahligkeitssatz geht auf Edmonds und Giles zurück. Eine ausführliche Darstellung findet sich in A. Schrijver, [Theory of Linear and Integer Programming](#), Kapitel 22. Vorlesungsdarstellungen bieten M. Goemans, Total Dual Integrality, und B. Gärtner, Discrete Optimization. Zur totalen Unimodularität siehe auch Goemans' Vorlesungsnotizen zur polyedrischen Kombinatorik.

Wir lernen nun zwei grundlegende exakte Verfahren kennen: Branch & Bound und Schnittebenen. Eine fehlende polynomielle Laufzeitgarantie macht ein korrektes exaktes Verfahren nicht zu einer Heuristik.

10.2 Die Branch & Bound Methode

Definition 10.10. Eine Familie von Mengen $\{S^i, i = 1, \dots, k\}$ ist eine [Aufteilung](#) von S , falls $\cup_{i=1}^k S^i = S$. Eine Aufteilung ist eine [Partition](#), falls zusätzlich $S^i \cap S^j = \emptyset$ für alle $i \neq j$ gilt.

Lemma 10.11. Sei

$$z_{IP}^i = \min\{cx : x \in S^i\}, \quad (IP^i)$$

wobei $\{S^i, i = 1, \dots, k\}$ eine Aufteilung von S ist. Dann gilt,

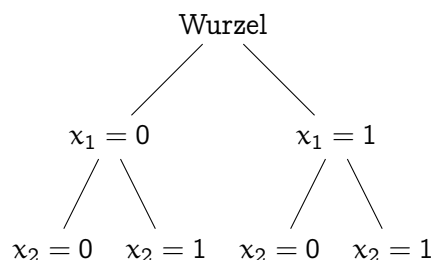
$$z_{IP} = \min\{z_{IP}^i : i = 1, \dots, k\}.$$

Beispiel 10.12. Wir betrachten folgendes Optimierungsproblem mit [binären](#) Variablen.

$$\text{minimiere } cx \quad (IP)$$

$$Ax \leq b$$

$$x \in \{0, 1\}^n.$$



Die Branch & Bound Methode berechnet iterativ Aufteilungsmengen S^i und versucht dann das auf S^i reduzierte IP zu lösen. Man kann sich die Vorgehensweise der Methode an einen [Enumerationsbaum](#) veranschaulichen. Wir starten mit dem Wurzelknoten S und bestimmen eine Aufteilung von S in Teilmengen S^i . Jede Aufteilungsmenge S^i wird mit einem Knoten assoziiert und mit einer Kante zur Wurzel S verbunden. Dieses Prinzip

wird nun iterativ und top-down auf jeden Kindknoten S^i angewendet. Falls eine optimale Lösung auf S^i gefunden wird, so weiss man, dass keine weitere Aufteilung von S^i nötig ist. Man spricht davon, dass im assoziierten **Enumerationsbaum** der entsprechende Knoten **abgeschnitten** wird.

Lemma 10.13. Der Enumerationsbaum kann an einem Knoten S^i abgeschnitten werden, falls eine der folgenden drei Bedingungen erfüllt ist:

1. Unzulässigkeit: $S^i = \emptyset$.
2. Optimalität: Eine Optimallösung für IP^i ist bekannt. Vor dem Abschneiden wird die bisher beste bekannte Lösung gegebenenfalls durch diese Lösung ersetzt.
3. Optimalwertdominanz: $z_{IP}^i \geq \bar{z}_{IP}$, wobei $\bar{z}_{IP}$ der Wert einer beliebigen bekannten zulässigen Lösung des IPs ist.

Wir werden nun Lemma 10.13 dazu benutzen, einen exakten Algorithmus zu entwerfen. Dabei werden wir allerdings nicht die Unterprobleme IP^i explizit lösen, sondern **Relaxierungen** verwenden.

Definition 10.14. Eine Relaxierung des Problems IP^i ist ein Optimierungsproblem

$$\min\{z_R^i(x) | x \in S_R^i\}, \quad (RP^i)$$

für das gilt:

$$S^i \subseteq S_R^i \text{ und } z_R^i(x) \leq cx \text{ für alle } x \in S^i.$$

Lemma 10.15. Der Enumerationsbaum kann am Knoten S^i abgeschnitten werden, falls eine der folgenden drei Bedingungen erfüllt ist:

1. RP^i ist unzulässig.
2. Eine Optimallösung x_R^i für RP^i erfüllt $x_R^i \in S^i$ und $z_R^i(x_R^i) = cx_R^i$. Vor dem Abschneiden wird die bisher beste bekannte Lösung gegebenenfalls aktualisiert.
3. Werte-Dominanz: $z_R^i \geq \bar{z}_{IP}$, wobei $\bar{z}_{IP}$ ein beliebiger Zielfunktionswert einer zulässigen Lösung des IPs ist.

Beweis. 1. impliziert $S^i = \emptyset$. 2. impliziert, dass x_R^i das Problem IP^i löst. 3. impliziert $z_{IP}^i \geq z_R^i \geq \bar{z}_{IP}$, so dass keine bessere Lösung in diesem Teilproblem liegen kann. $\square$

10.3 LP-Basierte Relaxierungen

Wir setzen voraus, dass die LP-Relaxierung am Wurzelknoten beschränkt ist. Jedes nicht-leere Teilproblem hat dann ein endliches Optimum. Vor der ersten ganzzahligen Lösung

setzen wir $\bar{z}_{IP} = +\infty$.

Wir entwickeln nun einen Enumerationsbaum bei dem eine LP-basierte Relaxierung benutzt wird.

Für den Wurzelknoten betrachten wir als zulässige Menge $S_{LP}^0 = \{x \in \mathbb{R}_+^n : Ax \leq b\}$ und definieren $z_R(x) = cx$.

Angenommen die LP Relaxierung am Knoten i ist gegeben durch

$$z_{LP}^i = \min\{cx : x \in S_{LP}^i\}, \quad (LP^i)$$

wobei $S_{LP}^i = \{x \in \mathbb{R}_+^n : A^i x \leq b^i\}$. Falls LP^i eine Optimallösung besitzt, bezeichnen wir diese mit x^i . Die Bedingungen um den Knoten i abzuschneiden ergeben sich durch:

1. $S_{LP}^i = \emptyset$ (unzulässig).
2. Optimale Lösung, falls $x^i \in \mathbb{Z}_+^n$. Aktualisiere die bisher beste bekannte Lösung, falls $cx^i < \bar{z}_{IP}$, und schneide dann den Knoten ab.
3. Werte-Dominanz: $z_{LP}^i \geq \bar{z}_{IP}$, wobei $\bar{z}_{IP}$ ein beliebiger Zielfunktionswert einer zulässigen Lösung des IPs ist.

Wir müssen nun noch erklären, wie wir die einzelnen Knoten des Enumerationsbaums konstruieren, bzw. wie die Unterteilung konstruiert wird. Das wichtigste Verfahren ist die sog. **Variablen-Dichotomie**: falls x^i fraktional ist, wählen wir eine fraktionale Komponente x_j^i und fügen die Bedingungen

$$x_j \leq \lfloor x_j^i \rfloor \text{ and } x_j \geq \lfloor x_j^i \rfloor + 1.$$

hinzu.

Satz 10.16. Sei $P = \{x \in \mathbb{R}_+^n : Ax \leq b\}$ ein nichtleeres Polytop. Dann ist ein Enumerationsbaum, der durch Variablen-Dichotomie erstellt wurde, stets endlich. Sei

$$\omega_j = \lceil \max\{x_j : x \in P\} \rceil.$$

Dann gilt, dass kein Pfad des Baums mehr als $\sum_{j=1}^n \omega_j$ Kanten enthält.

Beweis. Entlang eines Pfades liegen für jede Variable ganzzahlige Schranken $\ell_j \leq x_j \leq u_j$ vor, anfangs $\ell_j = 0$, $u_j = \omega_j$. Eine Verzweigung an einer fraktionalen Komponente erhöht die untere oder senkt die obere Schranke um mindestens 1. Solange der Knoten nichtleer ist, gilt $\ell_j \leq u_j$. Somit kann die Summe $\sum_j (u_j - \ell_j)$ nur endlich oft abnehmen, und jeder Pfad hat höchstens $\sum_j \omega_j$ Kanten. $\square$

Beispiel 10.17.

$$\begin{aligned} &\text{minimiere } x_1 - 2x_2 \\ &-4x_1 + 6x_2 \leq 9 \end{aligned} \quad (IP)$$

$$\begin{aligned} x_1 + x_2 &\leq 4 \\ x &\in \mathbb{Z}_+^2, \end{aligned}$$

Die LP-Lösung des Wurzelknotens S_{LP}^0 ergibt $cx^0 = -7/2$ mit $x^0 = (3/2, 5/2)$. Mit $x_2^0 = 5/2$ können wir S^0 unterteilen in $S^1 = \{x \in S^0 : x_2 \leq 2\}$ und $S^2 = \{x \in S^0 : x_2 \geq 3\}$. Wir lösen die nachfolgenden Relaxierungen S_{LP}^1 und S_{LP}^2 und erhalten $cx^1 = -13/4$ für $x^1 = (3/4, 2)$, und $S^2 = \emptyset$. Somit kann S^2 abgeschnitten werden. Wir unterteilen S^1 in $S^{11} = \{x \in S^1 : x_1 \leq 0\}$ und $S^{12} = \{x \in S^1 : x_1 \geq 1\}$. Die zugehörigen optimalen Lösungen der LP Relaxierungen sind $cx^{11} = -3$ für $x^3 = x^{11} = (0, 3/2)$ und $cx^{12} = -3$ für $x^4 = x^{12} = (1, 2)$. Weil nun $x^4 = x^{12}$ ganzzahlig ist können wir $\bar{z}_{IP} = -3$ setzen und S^{11} abschneiden.

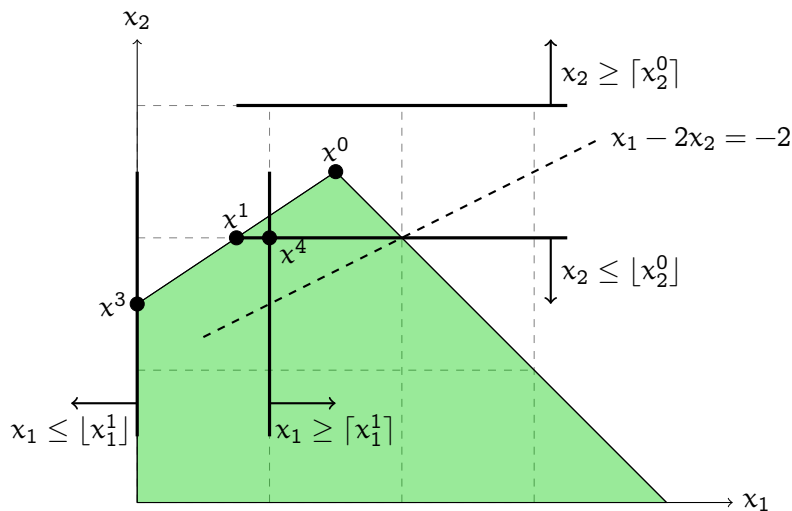


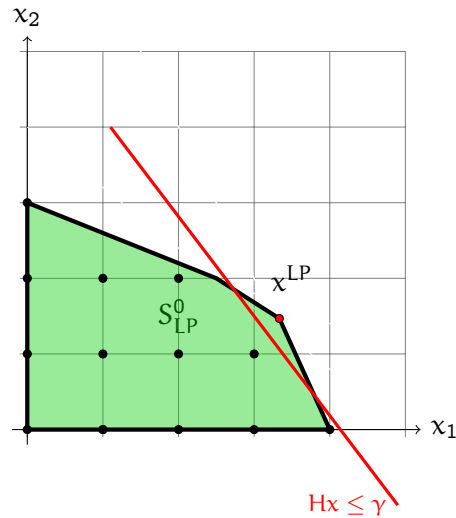
Abbildung 10.2: Beispiel des LP-basierten Branch & Bound Verfahrens.

10.4 Schnittebenenverfahren

Auch hier sei die anfängliche LP-Relaxierung beschränkt.

Wir wollen abschliessend noch das **Schnittebenenverfahren** beschreiben. Die Idee ist folgende: Zu einem ganzzahligen Problem (IP) lösen wir die zugehörige LP Relaxierung und überprüfen, ob die gefundene Ecke (also eine optimale Basis-Lösung) ganzzahlig ist. Falls ja, sind wir fertig, falls nicht, gilt es, eine gültige Ungleichung zu finden die die Basis-Lösung **abschneidet**.

Wir erhalten folgenden Algorithmus für das Problem (IP): $\min\{cx \mid x \in S\}$ wobei $S := \{x \in \mathbb{R}^n \mid Ax \leq b, x \in \mathbb{Z}_+^n\}$:

Abbildung 10.3: Beispiel einer Schnittebene $Hx \leq \gamma$.

1. Setze $k = 0$. Setze P^0 als relaxierte zulässige Menge von S indem die Ganzzahligkeit weggelassen wird.

Iterativ:

- (a) Ist $P^k = \emptyset$, STOP: Das IP ist unzulässig. Andernfalls berechne eine optimale Ecke x^k der LP-Relaxierung bezüglich P^k .
- (b) Falls $x^k \in \mathbb{Z}^n$, STOP: optimale Lösung gefunden.
- (c) Falls $x^k \notin \mathbb{Z}^n$, finde eine trennende Hyperebene $Hx \leq \gamma$ mit $x^k \notin P^k \cap \{x \in \mathbb{R}^n | Hx \leq \gamma\} \supseteq S$.
- (d) $P^{k+1} \leftarrow P^k \cap \{x \in \mathbb{R}^n | Hx \leq \gamma\}$
- (e) $k \leftarrow k + 1$

Eine fraktionale Ecke von P^k liegt nicht in der konvexen Hülle der ganzzahligen Punkte von P^k : Sonst wäre sie eine echte konvexe Kombination anderer zulässiger Punkte. Daher existiert eine gültige trennende Ungleichung. Ein beliebiger fraktionaler Optimalpunkt muss diese Eigenschaft nicht haben. Die Wahl beliebiger gültiger Schnitte allein garantiert noch keine endliche Terminierung; dazu benötigt man eine geeignete Schnittregel.