Kapitel 1

Einführung

Die Optimierung oder Programmierung untersucht die Fragestellung: *Gesucht ist die optimale Lösung eines Problems unter irgendwelchen Bedingungen.* Die mathematische Formulierung ist: Gegeben seien Funktionen $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : S \to \mathbb{R}$, $i = 1, \ldots, m$, $S \subset \mathbb{R}^n$, suche

$$f(\mathbf{x}) \to \text{Extremum unter den Bedingungen } g_i(\mathbf{x}) \geq 0,\ i = 1, \ldots, m.$$

Sind alle Funktionen linear, so hat man ein Problem der linearen Optimierung oder linearen Programmierung, siehe Teil I.

Bei Optimierungsproblemen müssen folgende Fragestellungen untersucht werden:

- Wie lauten notwendige und hinreichende Bedingungen für die Existenz von Lösungen?
- Wie kann man Lösungen mit möglichst geringem Aufwand berechnen? Was sind die effizientesten Algorithmen?

In der Einführung werden einige typische Beispiele von Optimierungsproblemen angegeben.

Beispiel 1.1 Rundreiseproblem. Gegeben sind n verschiedene Orte O_i, $i = 1, \ldots, n$. Die Entfernung zwischen den Orten O_i und O_j sei a_{ij}. Anstelle der Entfernung können auch andere Parameter wie Kosten oder Zeit genommen werden. Man nimmt im allgemeinen auch $a_{ij} \neq a_{ji}$ an. Das Rundreiseproblem oder auch Traveling–Salesman–Problem kann nun wie folgt formuliert werden:
Ein Reisender, der in einem Ort startet, möchte alle restlichen Orte genau einmal besuchen und zum Ausgangsort zurückkehren. In welcher Reihenfolge hat er die Orte zu besuchen, damit die Gesamtlänge des Reiseweges minimal wird? □

Beispiel 1.2 Landwirtschaft, Anbauoptimierung. Es stehen 100 ha Ackerland zur Verfügung, die mit Kartoffeln x_1 ha und Getreide x_2 ha bestellt werden sollen. Ein Teil der Anbaufläche kann auch brach bleiben. Die Betriebskosten sind wie folgt (GE = Geldeinheit):

	Kartoffeln	Getreide	insgesamt verfügbar
Anbaukosten GE/ha	10	20	1100 GE
Arbeitstage/ha	1	4	160 Tage
Reingewinn GE/ha	40	120	

Bei welcher Bewirtschaftung erzielt man den größten Gewinn?

Die mathematische Formulierung des Problems ist wie folgt:

$$
\begin{array}{lrcl}
\text{zu maximierende Funktion:} & z = 40x_1 + 120x_2 & \to & \text{max}, \\
\text{zur Verfügung stehendes Geld:} & 10x_1 + 20x_2 & \le & 1100, \\
\text{zur Verfügung stehende Zeit:} & x_1 + 4x_2 & \le & 160, \\
\text{zur Verfügung stehende Fläche:} & x_1 + x_2 & \le & 100, \\
\text{keine negativen Anbauflächen:} & x_1, x_2 & \ge & 0.
\end{array}
$$

Diese Aufgabe kann man graphisch lösen.

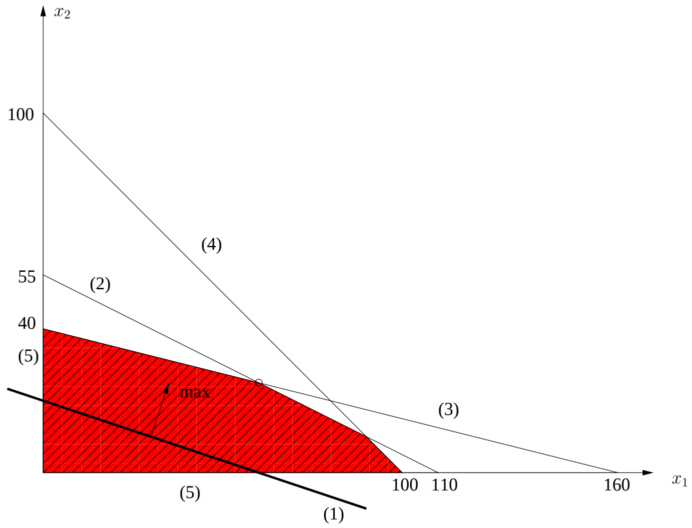


Abbildung 1.1: Darstellung der Nebenbedingungen und der Zielfunktion zum Beispiel 1.2.

Die Nebenbedingungen beschreiben Halbebenen und der Durchschnitt der Halbebenen ist gerade die Menge der Paare (x_1, x_2), in denen man das Maximum sucht. Zur graphischen Darstellung der Zielfunktion z wähle man sich einen beliebigen Punkt (x_1, x_2) und berechne die Gerade z durch diesen Punkt. In diesem Beispiel soll die Zielfunktion maximiert werden, das heißt, der Zielfunktionswert steigt, wenn man die Gerade orthogonal zu ihrem Anstieg nach oben verschiebt. Der letzte Punkt, der alle Nebenbedingungen erfüllt und der auf einer parallelen Geraden zur dargestellten Geraden liegt, ist der mit einem Kreis gekennzeichnete Punkt $(x_1, x_2) = (60, 25)$. Die Lösung dieses linearen Optimierungsproblems ist demzufolge $x_1 = 60$ ha, $x_2 = 25$ ha und der Gewinn ist 5400 GE. □

Beispiel 1.3 Ernährungsprogramm. Es stehen die folgenden drei Nahrungsmittel zur Verfügung (alle Angaben jeweils für 100 Gramm):

	Eiweiß	Fett	Kohlenhyd.	Wasser	Preis
1. Weißbrot	8	1	49	42	10
2. Wurst	12	20	0	68	80
3. Milch	3	3	5	89	7

Ein Ernährungsprogramm wird nur zugelassen, wenn es folgende Mindestanforderungen erfüllt: Eiweiß: 90 g, Fett: 80 g, Kohlenhydrate: 500 g, Wasser 2500 g. Ziel ist es, das kostengünstigste Ernährungsprogramm zu finden, welches diese Anforderungen erfüllt. Das zugehörige Optimierungsproblem lautet:

$$
\begin{aligned}
z = 10x_1 + 80x_2 + 7x_3 &\quad \to \quad \min \\
8x_1 + 12x_2 + 3x_3 &\quad \geq \quad 90 \\
x_1 + 20x_2 + 3x_3 &\quad \geq \quad 80 \\
49x_1 + 5x_3 &\quad \geq \quad 500 \\
42x_1 + 68x_2 + 89x_3 &\quad \geq \quad 2500 \\
x_1, x_2, x_3 &\quad \geq \quad 0,
\end{aligned}
$$

wobei die Maßeinheit für x_1, x_2, x_3 hier 100 g ist.

Die (gerundete) Lösung lautet: $x_1 = 7.71$, $x_2 = 0$, $x_3 = 24.45$, also 771 g Weißbrot und 2445 g Milch, das heißt vegetarisch. Die Kosten sind rund 248 GE. □

Beispiel 1.4 Rucksackproblem. Ein Wanderer kann in seinem Rucksack ein Gesamtgewicht von N tragen. Er hat n Gegenstände, die er mitnehmen möchte und jeder Gegenstand hat einen gewissen Nutzen n_i, $i = 1, \ldots, n$. Das Gesamtgewicht aller Gegenstände übersteigt das zulässige Maximalgewicht. Das Optimierungsproblem des Wanderers besteht nun darin, eine Teilmenge von Gegenständen mit maximalem Nutzen zu finden, so dass das Gesamtgewicht dieser Teilmenge höchstens N ist. Dabei kann als zusätzliche Nebenbedingung auftreten, dass gewisse Lösungskomponenten ganzzahlig sein müssen, zum Beispiel die Anzahl der Paar Schuhe, die er mitnehmen soll. □

Beispiel 1.5 Zuordnungsproblem. In einer Firma stehen zur Fertigung von n Produkten n Maschinen zur Verfügung. Jede Maschine eignet sich zur Herstellung jedes Produktes unterschiedlich gut. Es ergeben sich je nach Zuordnung verschiedene Arbeitszeiten. Jeder Maschine soll genau ein Produkt zugeordnet werden. Das Optimierungsproblem besteht darin, die Gesamtfertigungszeit der Produkte zu minimieren. □

Bemerkung 1.6 *Operations Research.* In der Fachliteratur werden Optimierungsaufgaben oft unter dem Begriff Operations Research (Optimalplanung) geführt. □

Literaturempfehlungen sind:

- Jarre und Stoer [JS04],
- Borgwardt [Bor01],
- Elster, Reinhardt, Schäuble, Donath [ERSD77],
- vor allem über das Gebiet der linearen Optimierung gibt es auch eine Reihe älterer Lehrbücher, die man verwenden kann.

Teil I

Lineare Optimierung

Kapitel 1

Grundlagen

Definition 1.1 Lineares Optimierungsproblem, lineares Programm. Eine Aufgabenstellung wird lineares Optimierungsproblem oder lineares Programm genannt, wenn das Extremum einer linearen Funktion

$$z = \sum_{i=1}^{n} c_i x_i = \mathbf{c}^T \mathbf{x} \tag{1.1}$$

zu bestimmen ist, über der durch das lineare Ungleichungssystem

$$\begin{aligned} \sum_{j=1}^{n} a_{ij} x_j \quad &\le (>) \quad b_i, \quad i = 1, \ldots, m & (1.2) \\ x_j \quad &\ge \quad 0, \quad j = 1, \ldots, n, & (1.3) \end{aligned}$$

definierten Punktmenge. $\square$

Seien $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Dann wird die Bezeichnung

$$\mathbf{x} \ge (>) \mathbf{y} \quad \iff \quad x_i \ge (>) y_i, \ \forall\ i = 1, \ldots, n,$$

verwendet.

Definition 1.2 Zulässiger Bereich. Die Menge $\mathcal{M}$ aller Punkte, die das Ungleichungssystem (1.2) – (1.3) erfüllen, heißt zulässiger Bereich. $\square$

Beispiel 1.3 Der zulässige Bereich, der durch lineare Nebenbedingungen beschrieben ist, ist der Durchschnitt von Halbräumen. Für $n = 2$ sind das Halbebenen und ein Beispiel ist in Abbildung 1.1 zu sehen. $\square$

Der zulässige Bereich ist nicht notwendig beschränkt. Er kann auch leer sein.

Definition 1.4 Konvexität. Eine Punktmenge $\mathcal{M}$ heißt konvex, wenn mit beliebigem $x_1, x_2 \in \mathcal{M}$ auch alle Punkte der Gestalt

$$\lambda x_1 + (1 - \lambda) x_2, \ 0 \le \lambda \le 1,$$

zu $\mathcal{M}$ gehören. $\square$

Für den $\mathbb{R}^n$ bedeutet Konvexität, dass mit zwei Punkten $\mathbf{x}_1, \mathbf{x}_2$ aus $\mathcal{M}$ auch ihre Verbindungsstrecke in $\mathcal{M}$ liegt.

Satz 1.5 *Die durch das lineare Ungleichungssystem (1.2) – (1.3) definierte Punktmenge ist konvex.*

Beweis: Seien $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{M}$ gegeben. Dann gelten

$$A\mathbf{x}_1 \leq \mathbf{b},\ \mathbf{x}_1 \geq \mathbf{0},\ A\mathbf{x}_2 \leq \mathbf{b},\ \mathbf{x}_2 \geq \mathbf{0}.$$

Mit $\lambda \in [0,1]$ gelten

$$\lambda A\mathbf{x}_1 \leq \lambda\mathbf{b}, \quad (1-\lambda)A\mathbf{x}_2 \leq (1-\lambda)\mathbf{b}.$$

Addition und Linearität der Matrizenmultiplikation ergibt

$$A\left(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2\right) \leq \mathbf{b}.$$

Analog folgt

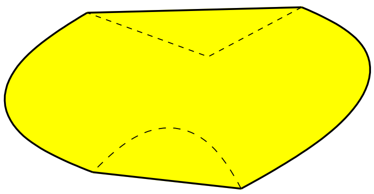
$$\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \geq \mathbf{0}.$$

■

Der Durchschnitt beliebig vieler konvexer Mengen ist wieder konvex. *Übungsaufgabe*

Definition 1.6 Konvexe Hülle. Die kleinstmögliche konvexe Menge $\overline{\mathcal{M}}$, die eine vorgegebene Punktmenge $\mathcal{P}$ enthält, heißt deren konvexe Hülle. □

Beispiel 1.7 Die dick umrandete Menge ist die konvexe Hülle der gestrichelt umrandeten Menge.



□

Beispiel 1.8 Die Menge

$$\mathcal{M} = \left\{ \frac{1}{n} \ :\ n \in \mathbb{N} \right\}$$

ist nicht konvex, da sie aus diskreten Punkten besteht. Ihre konvexe Hülle ist $(0,1]$. □

Definition 1.9 Konvexe Linearkombination. Gegeben seien q Punkte $\mathbf{x}_1, \ldots, \mathbf{x}_q$. Betrachtet werden alle Punkte der Gestalt

$$\mathbf{x} = \sum_{i=1}^{q} \lambda_i \mathbf{x}_i, \quad 0 \leq \lambda_i \leq 1, \quad \sum_{i=1}^{q} \lambda_i = 1. \tag{1.4}$$

Dann heißt die mit (1.4) erklärte Menge konvexe Linearkombination der Punkte $\mathbf{x}_1, \ldots, \mathbf{x}_q$. □

Welche Punkte einer konvexen Menge sollen ausgezeichnet werden?

Definition 1.10 Eckpunkt oder Extrempunkt einer konvexen Menge. Gegeben sei eine konvexe Menge $\mathcal{M}$. Der Punkt $\mathbf{x} \in \mathcal{M}$ heißt Eckpunkt oder Extrempunkt von $\mathcal{M}$, wenn aus $\mathbf{x} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$, $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{M}$, $0 < \lambda < 1$, folgt $\mathbf{x} = \mathbf{x}_1 = \mathbf{x}_2$. □

Man sagt, $\mathbf{x}$ lässt sich nicht als *echte* konvexe Linearkombination von Punkten $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{M}$ darstellen.

Beispiel 1.11 Bei einem Viereck im $\mathbb{R}^2$ sind die Eckpunkte gerade die vier Ecken. Bei einer Kugel im $\mathbb{R}^n$, $n \geq 1$, sind alle Randpunkte Eckpunkte. □

Definition 1.12 Konvexes Polyeder. Eine beschränkte konvexe Menge $\mathcal{M} \neq \emptyset$ mit endlich vielen Eckpunkten heißt konvexes Polyeder. $\square$

Beispiel 1.13 Konvexe Polyeder in $\mathbb{R}^n, n = 1, 2, 3$. Ein konvexes Polyeder in $\mathbb{R}^1$ ist ein abgeschlossenes Intervall. In $\mathbb{R}^2$ und $\mathbb{R}^3$ kann man sich konvexe Polyeder noch gut vorstellen. Ein Beispiel in $\mathbb{R}^2$ findet man in Abbildung 1.1. $\square$

Satz 1.14 *Sei $\mathcal{M}$ eine konvexe, abgeschlossene und beschränkte Menge, $\mathcal{P}$ sei die Menge der Eckpunkte von $\mathcal{M}$. Dann ist $\mathcal{M}$ die konvexe Hülle von $\mathcal{P}$.*

Beweis: Literatur. Beweisidee mit trennenden Hyperebenen siehe [ERSD77, Satz 2.48]. $\blacksquare$

Satz 1.15 *Ist der Lösungsbereich*

$$\mathcal{M} = \{\mathbf{x} \ : \ A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$$

beschränkt, so ist er ein konvexes Polyeder.

Beweis: Siehe später, Folgerung 3.7. $\blacksquare$

Kapitel 2

Geometrische Deutung des Linearen Programms

In diesem Abschnitt wird gezeigt, dass sich das anschauliche Merkmal des zweidimensionalen linearen Programms aus Beispiel 1.2, nämlich dass das Optimum auf dem Rand angenommen wird, verallgemeinern lässt.

Historie zur Untersuchung linearer Optimierungsprobleme:

- 1939 Leonid V. Kantorovitch (1912 – 1986); Methode der Auflösungskoeffizienten
- 1941 Frank L. Hitchcock, Transportproblem
- 1949 George Dantzig (1914 – 2005), Simplexmethode
- 1984 Narendra Karmarkar (geb. 1957), Innere–Punkt–Methoden für lineare Programme

Definition 2.1 Lineares Optimierungsproblem in 1. Normalform, lineares Programm in Normalform. Gesucht werden die Werte der n Variablen $x_1, \ldots, x_n$ so, dass die lineare Funktion

$$z = \sum_{j=1}^{n} c_j x_j = \mathbf{c}^T \mathbf{x} \tag{2.1}$$

die sogenannte Zielfunktion, unter den Nebenbedingungen

$$\begin{aligned} \sum_{j=1}^{n} a_{ij} x_j &\geq b_i, \quad i = 1, \ldots, m, \quad (A\mathbf{x} \geq \mathbf{b})\,, && (2.2) \\ x_j &\geq 0, \quad j = 1, \ldots, n, \quad (\mathbf{x} \geq \mathbf{0}), && (2.3) \end{aligned}$$

ein Minimum annimmt. Alle Koeffizienten sind reell. Das System (2.1) – (2.3) heißt lineares Optimierungsproblem oder lineares Programm in 1. Normalform. □

Bemerkung 2.2

1. Ob (2.1) in min– oder max–Form benutzt wird, ist im allgemeinen ohne Belang, in [JS04] wird beispielsweise die max–Form verwendet.
2. Die Relationen $\geq, =, \leq$ im System der Nebenbedingungen sind im wesentlichen äquivalent.
3. Fehlt zum Beispiel für einen Index k die Bedingung $x_k \geq 0$, so setzt man $x_k := \overline{x}_k - \hat{x}_k$ mit $\overline{x}_k, \hat{x}_k \geq 0$. Man erhöht damit die Anzahl der Variablen um Eins.

□

Definition 2.3 Zulässiger Punkt, zulässiger Bereich. Ein Punkt $\mathbf{x} = (x_1, \ldots, x_n)^T$ heißt zulässig, wenn er die Nebenbedingungen (2.2), (2.3) erfüllt. Die Gesamtheit aller zulässigen Punkte heißt zulässiger Bereich. $\Box$

Für die Lösung von (2.1) – (2.3) kommen nur zulässige Punkte in Betracht. Der zulässige Bereich ist konvex. Ist er beschränkt, so ist er ein konvexes Polyeder. Ist der zulässige Bereich nicht beschränkt ist, dann gilt:

- entweder ist (2.1) über diesen Bereich selbst nicht beschränkt, Beispiel: Minimiere $-2x_1 - x_2$ im Bereich $\{(x_1, x_2) \ : \ x_1 \ge 0, x_2 \ge 0\}$,
- oder (2.1) ist über dem unbeschränkten Bereich beschränkt. Dann kann man Zusatzbedingungen an den zulässigen Bereich stellen, die das Optimum nicht ändern, so dass der neue zulässige Bereich beschränkt ist. Beispiel: Minimiere $2x_1 + x_2$ im Bereich $\{(x_1, x_2) \ : \ x_1 \ge 0, x_2 \ge 0\}$.

Weitere Beispiele findet man in Beispiel 2.6.

Wenn von einem konvexen Polyeder gesprochen wird, ist ab sofort immer ein abgeschlossenes konvexes Polyeder gemeint.

Satz 2.4 ***Extremwertannahme.*** *Eine auf einem konvexen Polyeder definierte lineare Funktion $z = f(\mathbf{x})$ nimmt ihren kleinsten Wert in (mindestens) einem Eckpunkt an.*

Beweis: Seien $\mathbf{x}_1, \ldots \mathbf{x}_p$ die Eckpunkte des konvexen Polyeders $\mathcal{M}$. Die Funktion $f(\mathbf{x})$ nehme ihr Minimum in $\mathbf{x}_0 \in \mathcal{M}$ an, das heißt

$$f(\mathbf{x}_0) \le f(\mathbf{x}) \tag{2.4}$$

für alle Punkte $\mathbf{x}$ des konvexen Polyeders. Dass das Minimum angenommen wird, folgt nach dem Satz von Bolzano–Weierstraß (stetige Funktion in einem kompakten Gebiet nimmt ihre Extremwerte an). Ist $\mathbf{x}_0$ kein Eckpunkt, so existiert eine Darstellung (Satz 1.14)

$$\mathbf{x}_0 = \sum_{j=1}^{p} \lambda_j \mathbf{x}_j, \ 0 \le \lambda_j \le 1, \ \sum_{j=1}^{p} \lambda_j = 1.$$

Aus der Linearität von f folgt

$$f(\mathbf{x}_0) = f\left(\sum_{j=1}^{p} \lambda_j \mathbf{x}_j\right) = \sum_{j=1}^{p} \lambda_j f(\mathbf{x}_j).$$

Sei ein Index l definiert durch

$$f(\mathbf{x}_l) = \min_{j=1,\ldots,p} f(\mathbf{x}_j).$$

Dann folgt

$$f(\mathbf{x}_0) \ge f(\mathbf{x}_l) \sum_{j=1}^{p} \lambda_j = f(\mathbf{x}_l). \tag{2.5}$$

Wegen (2.4) und (2.5) wird das Minimum für $\mathbf{x}_l$ angenommen. $\blacksquare$

Folgerung 2.5 *Wird das Minimum in mehr als einem Eckpunkt des konvexen Polyeders angenommen, so wird es auf der konvexen Hülle dieser Eckpunkte angenommen.*

Beweis: Ohne Beschränkung der Allgemeinheit seien die Eckpunkte so numeriert, dass die Zielfunktion $f(\mathbf{x})$ ihr Minimum in den Eckpunkten $\mathbf{x}_1, \ldots, \mathbf{x}_q$ annehme. Die konvexe Hülle dieser Eckpunkte ist

$$\left\{ \tilde{\mathbf{x}} \ : \ \tilde{\mathbf{x}} = \sum_{i=1}^{q} \lambda_i \mathbf{x}_i, \ 0 \le \lambda_i \le 1, \sum_{i=1}^{q} \lambda_i = 1 \right\}.$$

Aus der Linearität von f folgt

$$f(\tilde{\mathbf{x}}) = f\left(\sum_{i=1}^{q} \lambda_i \mathbf{x}_i\right) = \sum_{i=1}^{q} \lambda_i f(\mathbf{x}_i) = f(\mathbf{x}_1) \sum_{i=1}^{q} \lambda_i = f(\mathbf{x}_1),$$

womit die Folgerung bewiesen ist. ■

Geometrische Interpretation

Die Gleichung $z = \mathbf{c}^T\mathbf{x} - d$ mit einer vorgegebenen Konstanten d ist die Gleichung einer Hyperebene in $\mathbb{R}^n$. Für $n = 3$, hat man beispielsweise die Normalform einer Ebenengleichung, wobei $\mathbf{c}$ ein Normalenvektor der Ebene ist.

Sei $z = \mathbf{c}^T\mathbf{x}$ die Zielfunktion. Es ist gerade

$$\mathbf{c} = \nabla z = \left(\frac{\partial z}{\partial x_1}, \ldots, \frac{\partial z}{\partial x_n}\right)^T.$$

Außerdem ist $\mathbf{c}$ orthogonal zu den Hyperebenen $\mathbf{c}^T\mathbf{x} = const$. Sei nämlich ein beliebiger Vektor einer Hyperebene gegeben, etwa zwischen den Punkten $\tilde{\mathbf{x}}$ und $\hat{\mathbf{x}}$, dann gilt

$$\mathbf{c}^T\tilde{\mathbf{x}} = const,\ \mathbf{c}^T\hat{\mathbf{x}} = const \implies \mathbf{c}^T(\tilde{\mathbf{x}} - \hat{\mathbf{x}}) = 0.$$

Aus der Menge $\{\mathbf{c}^T\mathbf{x} = const\}$ wählen wir diejenige Hyperebene, die einen vorgegebenen zulässigen Punkt $\mathbf{x}_0 \in \mathcal{M}$, nicht notwendig einen Eckpunkt, enthält: $\mathbf{c}^T\mathbf{x} = \mathbf{c}^T\mathbf{x}_0$. Wir definieren

$$g := \{\mathbf{x}\ :\ \mathbf{x} = \mathbf{x}_0 + t\mathbf{c}, t \in \mathbb{R}\}.$$

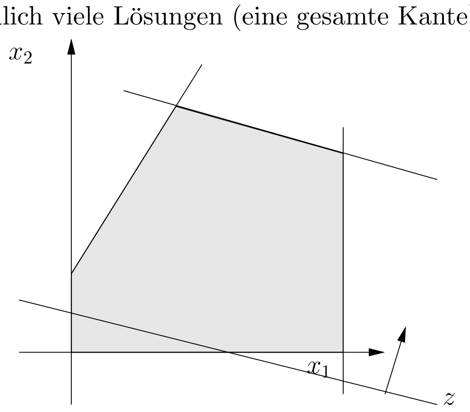
Diese Gerade enthält $\mathbf{x}_0$ und sie ist orthogonal zu $\mathbf{c}^T\mathbf{x} = const$. Für alle $\mathbf{x} \in g$ gilt bezüglich der Zielfunktion

$$z = \mathbf{c}^T\mathbf{x} = \mathbf{c}^T(\mathbf{x}_0 + t\mathbf{c}) = \mathbf{c}^T\mathbf{x}_0 + t\,\|\mathbf{c}\|_2^2 =: z_0 + t\,\|\mathbf{c}\|_2^2,$$

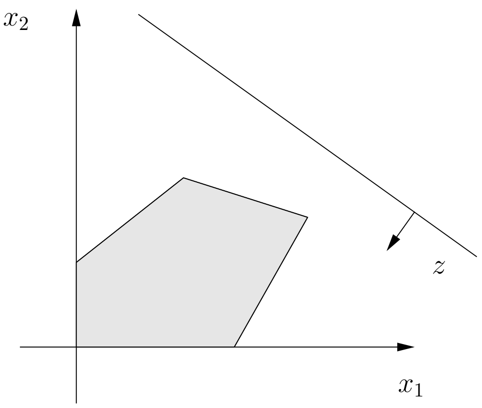
wobei z_0 der Startwert der Zielfunktion ist. Sei $t > 0$. Dann folgt $z > z_0$, das heißt, der Wert der Zielfunktion wächst streng monoton in Richtung von $\mathbf{c}$. Wenn man z zu maximieren hat, so verschiebe man die Hyperebene in Richtung ihres Gradienten. Also, ausgehend von $\mathbf{c}^T(\mathbf{x} - \mathbf{x}_0) = 0$ konstruiere man in Richtung von $\mathbf{c}$ eine Schar zu $\mathbf{c}^T(\mathbf{x} - \mathbf{x}_0) = 0$ paralleler Hyperebenen mit dem Ziel, diejenige Hyperebene aus der Schar zu finden, die $\{\mathbf{x}\ :\ A\mathbf{x} \le \mathbf{b},\ \mathbf{x} \ge \mathbf{0}\}$ berührt mit der Eigenschaft, dass $\{\mathbf{x}\ :\ A\mathbf{x} \le \mathbf{b},\ \mathbf{x} \ge \mathbf{0}\}$ ganz im negativen Halbraum der berührenden Hyperebene liegt. Berührung bedeutet, dass $\{\mathbf{x}\ :\ A\mathbf{x} \le \mathbf{b},\ \mathbf{x} \ge \mathbf{0}\} \cap \left\{\mathbf{c}^T\mathbf{x} = const\right\}$ eine Teilmenge des Randes des Polyeders ist, zum Beispiel ein Eckpunkt.

Beispiel 2.6 Beispiele für Situationen die in linearen Programmen auftreten können.

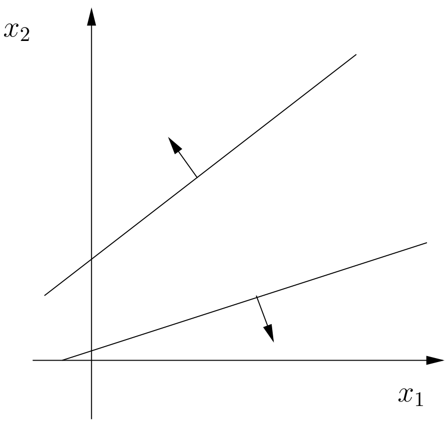
a) Es gibt unendlich viele Lösungen (eine gesamte Kante).



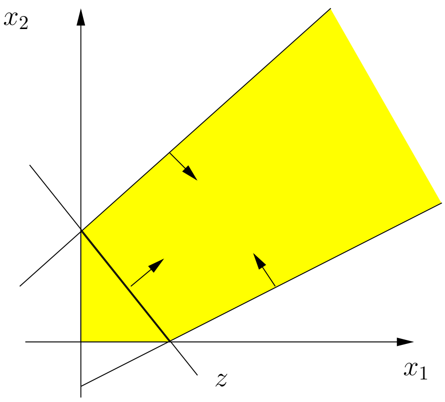
b) Es gibt überflüssig Nebenbedingungen. Die Zielfunktion nimmt ihren Extremwert in $(0, 0)$ an und die drei oberen Nebenbedingungen sind überflüssig.



c) Der zulässige Bereich ist leer.



d) Der Optimalwert ist nicht beschränkt.



Kapitel 3

Basislösungen eines linearen Programms in Normalform

Definition 3.1 Lineares Programm in 2. Normalform, einfache Normalform. Gegeben sei das lineare Programm

$$z = \mathbf{c}^T\mathbf{x} \to \min ! \tag{3.1}$$

unter den folgenden Bedingungen

$$A\mathbf{x} = \mathbf{b} \tag{3.2}$$

$$\mathbf{x} \geq \mathbf{0} \tag{3.3}$$

mit $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$ und $A \in \mathbb{R}^{m\times n}$. Dieses Problem wird lineares Programm in (2.) Normalform genannt. □

Bemerkung 3.2 Wenn man die lineare Ungleichung

$$\sum_{j=1}^{n'} a_{i^*j}x_j \leq b_{i^*}$$

gegeben hat, so kann man eine sogenannte Schlupfvariable einführen

$$\sum_{j=1}^{n'} a_{i^*j}x_j + x_{n'+1} = b_{i^*}, \quad x_{n'+1} \geq 0.$$

Mit Hilfe der Schlupfvariablen gelingt es aus dem linearen Programm in 1. Normalform ein lineares Programm in 2. Normalform zu machen. Diese sind äquivalent. Die Kosten der Einführung von Schlupfvariablen bestehen darin, dass man die Dimension des Lösungsvektors erhöht. □

Wir machen jetzt die folgenden Voraussetzungen:

1. $m < n$, das heißt weniger Nebenbedingungen als Unbekannte.
2. $\mathrm{rg}(A) = m$, das heißt, A hat vollen Zeilenrang, das heißt die Nebenbedingungen sind linear unabhängig.
3. $A\mathbf{x} = \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$ sei widerspruchsfrei, das heißt, der zulässige Bereich ist nicht leer.

Definition 3.3 Basislösung. Basislösungen des linearen Programms (3.1) – (3.3) sollen die Lösungsvektoren $\mathbf{x} = (x_{i1}, \ldots, x_{im}, 0, \ldots, 0)^T$ heißen, für die die m Variablen $x_{i1}, \ldots, x_{im}$ eine nicht singuläre Koeffizientenmatrix

$$A_{m,m} = (\mathbf{a}_{i1}, \ldots, \mathbf{a}_{im}) \in \mathbb{R}^{m\times m},$$

besitzen, wobei $(\mathbf{a}_j)$, $j = 1, \ldots, n$, die Spaltenvektoren von A bezeichne. □

Die ersten m Variablen einer Basislösung können beliebige reelle Zahlen sein.

Definition 3.4 zulässige Basislösung, ausgeartete (entartete) Basislösung, Basisvariable, Basisvektor. Gilt für eine Basislösung $\mathbf{x} = (x_{i1}, \ldots, x_{im}, 0, \ldots, 0)^T$, dass $x_{ij} \geq 0$ für alle $j = 1, \ldots, m$, dann heißt sie zulässig. Verschwindet sie in mindestens einer der Variablen $x_{i1}, \ldots, x_{im}$, so heißt sie ausgeartet oder entartet.

Die Komponenten einer Basislösung werden Basisvariable genannt, die zugehörigen Spaltenvektoren heißen Basisvektoren. Entsprechend spricht man von Nichtbasisvariablen und Nichtbasisvektoren. □

Beispiel 3.5

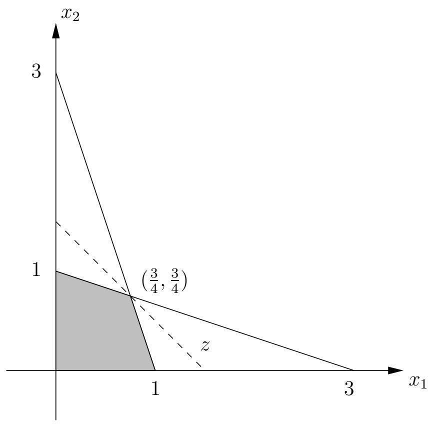
$$\begin{aligned} z = -x_1 - x_2 &\rightarrow \text{min !} \\ x_1 + 3x_2 + x_3 &= 3 \\ 3x_1 + x_2 + x_4 &= 3 \\ \mathbf{x} &\geq \mathbf{0}. \end{aligned}$$

Zulässige, nicht ausgeartete Basislösungen sind $(i1 = 3, i2 = 4)$

$$\mathbf{x} = (0, 0, 3, 3)^T, \; A_{2,2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; z = 0.$$

oder $(i1 = 1, i2 = 2)$

$$\mathbf{x} = (3/4, 3/4, 0, 0)^T, \; A_{2,2} \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}, z = -\frac{3}{2}.$$



Wir führen jetzt die weitere Nebenbedingung

$$x_1 + x_2 \leq \frac{3}{2}$$

ein. Die Nebenbedingungen des erweiterten linearen Programms haben die Gestalt

$$\begin{pmatrix} 1 & 3 & 1 & 0 & 0 \\ 3 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 1.5 \end{pmatrix}.$$

Dann ist eine ausgeartete zulässige Basislösung des erweiterten linearen Programms $(i1 = 1, i2 = 2, i3 = 5)$

$$(3/4, 3/4, 0, 0, 0)^T, \; A_{3,3} = \begin{pmatrix} 1 & 3 & 0 \\ 3 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \; z = -\frac{3}{2}.$$

Im Bild erkennt man, was Ausartung bedeutet. Die Ecke $(3/4, 3/4)$ des zulässigen Bereichs ist bereits durch die ersten beiden Nebenbedingungen bestimmt. Durch die neue Nebenbedingung ist diese Ecke nun wahlweise durch die ersten beiden, durch die erste und die dritte oder die zweite und die dritte Nebenbedingung bestimmt. Die Nebenbedingungen, die diese Ecke des zulässigen Bereichs bestimmen, sind nicht mehr eindeutig. □

Satz 3.6 *Ein Eckpunkt eines zulässigen Bereichs* $\mathcal{M} = \{\mathbf{x} \ : \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \ge \mathbf{0}\}$ *liegt genau dann vor, wenn seine Koordinaten eine zulässige Basislösung bilden.*

Beweis: *a) Aus Basislösung folgt Eckpunkt.* Sei $\mathbf{x} = (x_1, x_2, \ldots, x_m, 0, \ldots, 0)^T$ eine zulässige Basislösung, das heißt $x_i \ge 0$, $i = 1, \ldots m$,

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \ldots + \mathbf{a}_m x_m = \mathbf{b},$$

die Vektoren $\mathbf{a}_1, \ldots, \mathbf{a}_m$ sind linear unabhängig und die Nichtbasisvariablen $x_{m+1}, \ldots, x_n$ sind gleich Null.

Der Beweis wird indirekt geführt, indem angenommen wird, dass

$$\mathbf{x} = (x_1, \ldots, x_m, 0, \ldots, 0)$$

kein Eckpunkt ist. Dann gibt es Punkte $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{M}$ mit $\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 = \mathbf{x}$, $\mathbf{x}_1 \ne \mathbf{x}_2$, $0 < \lambda < 1$. Da die letzten $n - m$ Komponenten von $\mathbf{x}$ verschwinden, muss das auch für entsprechenden Komponenten von $\mathbf{x}_1$ und $\mathbf{x}_2$ gelten, da alle Komponenten dieser Vektoren nichtnegativ sind. Seien nun

$$\mathbf{x}_1 = \left(x_1^{(1)}, x_2^{(1)}, \ldots, x_m^{(1)}, 0, \ldots, 0\right)^T, \quad \mathbf{x}_2 = \left(x_1^{(2)}, x_2^{(2)}, \ldots, x_m^{(2)}, 0, \ldots, 0\right)^T.$$

Da diese Punkte zulässig sind, folgt

$$\begin{aligned} \mathbf{a}_1 x_1^{(1)} + \mathbf{a}_2 x_2^{(1)} + \ldots + \mathbf{a}_m x_m^{(1)} &= \mathbf{b}, \\ \mathbf{a}_1 x_1^{(2)} + \mathbf{a}_2 x_2^{(2)} + \ldots + \mathbf{a}_m x_m^{(2)} &= \mathbf{b}. \end{aligned}$$

Wegen der linearen Unabhängigkeit der Vektoren $\mathbf{a}_1, \ldots, \mathbf{a}_m$ folgt daraus $\mathbf{x}_1 = \mathbf{x}_2$, was im Widerspruch zur Annahme steht. Also ist $\mathbf{x}$ ein Eckpunkt.

b) Aus Eckpunkt folgt Basislösung. Sei $\mathbf{x}$ ein Eckpunkt des zulässigen Bereichs mit den positiven Koordinaten $x_1, \ldots, x_k$, das heißt, es gilt

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \ldots + \mathbf{a}_k x_k = \mathbf{b} \text{ mit } x_j > 0, \ j = 1, \ldots, k \le n.$$

Ohne Beschränkung der Allgemeinheit seien die verschwindenden Komponenten die hinteren. Es ist zu zeigen, dass $\mathbf{a}_1, \ldots, \mathbf{a}_k$ linear unabhängig sind.

Der Beweis ist wieder indirekt. Wir nehmen also an, dass es ein $\mathbf{y} \in \mathbb{R}^k$ gibt mit

$$\mathbf{a}_1 y_1 + \mathbf{a}_2 y_2 + \ldots + \mathbf{a}_k y_k = \mathbf{0}$$

und mindestens einem $y_i \ne 0$. Für jede reelle Zahl μ gelten damit

$$\sum_{j=1}^{k} \mathbf{a}_j (x_j + \mu y_j) = \mathbf{b} \quad \text{und} \quad \sum_{j=1}^{k} \mathbf{a}_j (x_j - \mu y_j) = \mathbf{b}.$$

Das bedeutet, die Punkte

$$\begin{aligned} \mathbf{x}_1 &= (x_1 + \mu y_1, \ldots, x_k + \mu y_k, 0, \ldots, 0), \\ \mathbf{x}_2 &= (x_1 - \mu y_1, \ldots, x_k - \mu y_k, 0, \ldots, 0) \end{aligned}$$

erfüllen die Nebenbedinungen (3.2). Falls man $\mu > 0$ hinreichend klein wählt, sind alle Komponenten dieser Punkte nichtnegativ und $\mathbf{x}_1, \mathbf{x}_2$ sind zulässig. Aus der Konstruktion von $\mathbf{x}_1, \mathbf{x}_2$ folgt, dass $\mathbf{x} = 0.5\mathbf{x}_1 + 0.5\mathbf{x}_2$ gilt. Das ist im Widerspruch zur Eckpunktannahme von $\mathbf{x}$. Diese Darstellung für den Eckpunkt $\mathbf{x}$ kann nur existieren, wenn $\mathbf{x}_1 = \mathbf{x}_2$. Da μ positiv ist, muss also $y_1 = \ldots = y_k = 0$ gelten. Also sind $\mathbf{a}_1, \ldots, \mathbf{a}_k$ linear unabhängig.

Die Basislösung verlangt jedoch m linear unabhängige Vektoren:

- Fall $k > m$. $m+1$ Vektoren des $\mathbb{R}^m$ sind stets linear abhänging. Dieser Fall kann also nicht eintreten.
- Fall $k = m$. In diesem Fall besitzt die zulässige Basislösung m positiven Komponenten, sie ist also nicht ausgeartet.
- Fall $k < m$. In diesem Fall hat man eine zulässige Basislösung mit weniger als m positiven Komponenten, also eine ausgeartete Basislösung. Aus den restlichen Spalten von A konstruiert man eine Menge von linear unabhängigen Vektoren $\mathbf{a}_1, \ldots, \mathbf{a}_m$, für welche offensichtlich

$$\mathbf{a}_1 x_1 + \ldots + \mathbf{a}_k x_k + \mathbf{a}_{k+1} 0 + \ldots \mathbf{a}_m 0 = \mathbf{b}$$

gilt. Diese Konstruktion ist möglich, da $\mathrm{rg}(A) = m$ ist.

■

Folgerung 3.7 ***Satz 1.15.*** *Ist der Lösungsbereich*

$$\mathcal{M} = \{\mathbf{x} \ : \ A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$$

beschränkt, so ist er ein konvexes Polyeder.

Beweis: Man kann nur endlich viele Mengen von m linear unabhängigen Spaltenvektoren der Matrix A bilden. (*Maximalanzahl ist Übungsaufgabe*) Mit dem eben bewiesenen Satz hat damit $\mathcal{M}$ nur endlich viele Ecken. ■

Eine weitere Folgerung des eben bewiesenen Satzes, zusammen mit Satz 2.4 ist wie folgt.

Folgerung 3.8 *Eine über einem konvexen Polyeder $\mathcal{M} = \{\mathbf{x} \ : \ A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ definierte lineare Funktion $z = \mathbf{c}^T\mathbf{x}$ nimmt ihr Minimum für wenigstens eine zulässige Basislösung an.*

Bemerkung 3.9 Naives Verfahren. Mit Hilfe der bisherigen Resultate können wir versuchen, ein Verfahren zur Lösung von

$$\min_{\mathbf{x} \in \mathbb{R}^n} \left\{ z = \mathbf{c}^T\mathbf{x} \ : \quad A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \right\}$$

zu konstruieren:

1. Aufstellung der $\begin{pmatrix} n \\ m \end{pmatrix}$ linearen Gleichungssysteme der Dimension m aus $A\mathbf{x} = \mathbf{b}$.
2. Ist die so generierte Matrix $A_{m,m}$ regulär?
3. Angabe der Lösung für reguläre $A_{m,m}$.
4. Auswahl der zulässigen Lösungen
5. Bestimmung der Lösung(en), die das Minimum liefern.

Diese Herangehensweise ist jedoch schon bei relativ kleiner Anzahl von Unbekannten und Nebenbedingungen viel zu aufwendig. Zum Beispiel hätte man bei $n = 20$, $m = 10$ schon 184 756 Gleichungssysteme aufzustellen und diese zu untersuchen. □

Das Ziel wird nun sein, ein Verfahren zu finden, welches einen cleveren Weg zum Optimum findet, unter Nutzung von zulässigen Basislösungen.

Kapitel 4

Hauptsatz und Optimalitätskriterium der Simplexmethode

In diesem Abschnitt wird das wichtigste Verfahren zur Lösung linearer Optimierungsprobleme eingeführt – die Simplexmethode. Es existiere für

$$\min_{\mathbf{x}\in\mathbb{R}^n} \left\{ z = \mathbf{c}^T\mathbf{x} \ : \quad A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \right\}$$

eine zulässige, nicht ausgeartete Basislösung.

Definition 4.1 Simplex. Ein Simplex ist die Menge aller Punkte $\mathbf{x} = (x_1, \ldots, x_n)^T$ mit

$$\sum_{i=1}^{n} x_i \leq 1, \quad x_i \geq 0, \ i = 1, \ldots, n.$$

□

Simplizes sind spezielle konvexe Polyeder. Für $n = 2$ ist das Dreieck mit den Eckpunkten $(0,0), (1,0), (0,1)$ ein Simplex.

Bemerkung 4.2 Geometrisches Prinzip der Simplexmethode. Das geometrische Prinzip der Simplexmethode ist wie folgt:

1. Man beginnt an einer Ecke des zulässigen Bereichs mit einer Startlösung $\mathbf{x}_1$ und dem Zielfunktionswert $z(\mathbf{x}_1)$.
2. Dann geht man entlang einer absteigenden Kante, dass heißt, bei welcher der Zielfunktionswert kleiner wird, $z(\mathbf{x}_1) > z(\mathbf{x}_2)$ zu einer sogenannten benachbarten Ecke $\mathbf{x}_2$.
3. Wiederhole Schritt 2 so lange, bis es keine absteigende Kante mehr gibt.

Dieses geometrische Prinzip muss in die algebraische Terminologie mit Basislösungen usw. transformiert werden. Wir werden später auch diskutieren, dass man Simplexschritte ausführen kann, bei denen der Zielfunktionswert gleich bleibt. In diesem Fall ist die Beschreibung des zweiten Schritts auch abzuändern, da man nicht zu einer benachbarten Ecke geht, sondern auf der gegebenen Ecke die Basis ändert. Diese Situation kann im Falle der Ausartung eintreten. Man nennt zwei Basislösungen benachbart, wenn sie sich nur in einem Basisvektor unterscheiden. □

Sei $\mathbf{x} = (x_1, \ldots, x_m, 0, \ldots, 0)^T$ eine erste zulässige Basislösung. Es gilt

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \ldots + \mathbf{a}_m x_m = \mathbf{b}. \tag{4.1}$$

Dabei sind $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ linear unabhängige Vektoren. Der Zielfunktionswert ist demzufolge

$$z_0 = c_1 x_1 + \ldots + c_m x_m. \tag{4.2}$$

Alle Nichtbasisvektoren $\mathbf{a}_{m+1}, \ldots, \mathbf{a}_n$ werden durch die Basis dargestellt

$$\mathbf{a}_j = x_{1j}\mathbf{a}_1 + \ldots + x_{mj}\mathbf{a}_m, \quad j = m+1, \ldots, n. \tag{4.3}$$

Mit diesen Darstellungskoeffizienten x_{ij}, die nichts mit den Werten $x_1, \ldots, x_m$ der Basislösung zu tun haben, werden die Hilfsgrößen

$$z_j = c_1 x_{1j} + c_2 x_{2j} + \ldots + c_m x_{mj}, \quad j = m+1, \ldots, n, \tag{4.4}$$

eingeführt.

Satz 4.3 ***Hauptsatz der Simplexmethode.*** *Sei z_0 der Wert der Zielfunktion für die zulässige Basislösung $\mathbf{x} = (x_1, \ldots, x_m, 0, \ldots, 0)^T$, $x_i > 0$, $i = 1, \ldots, m$. Gilt für ein k mit $m+1 \le k \le n$, dass $z_k - c_k > 0$, so existiert wenigstens eine zulässige Basislösung mit einem Zielfunktionswert kleiner als z_0.*

Beweis: Sei $\theta > 0$ vorerst beliebig gewählt. Man multipliziere (4.3) für $j = k$ mit θ und bilde (4.1)$-\theta$ (4.3):

$$\mathbf{a}_1 (x_1 - \theta x_{1k}) + \mathbf{a}_2 (x_2 - \theta x_{2k}) + \ldots + \mathbf{a}_m (x_m - \theta x_{mk}) + \theta \mathbf{a}_k = \mathbf{b}, \tag{4.5}$$

In der Gleichung (4.5) steht ein Vektor, der $A\mathbf{x} = \mathbf{b}$ erfüllt:

$$(x_1 - \theta x_{1k}, \ldots, x_m - \theta x_{mk}, 0, \ldots, \theta, \ldots, 0)^T. \tag{4.6}$$

Es wird in den Lemmata 4.4 und 4.7 gezeigt, dass man mit diesem Vektor eine Basislösung konstruieren kann. Für seinen Zielfunktionswert gilt

$$\begin{aligned} & c_1 (x_1 - \theta x_{1k}) + c_2 (x_2 - \theta x_{2k}) + \ldots + c_m (x_m - \theta x_{mk}) + \theta c_k \\ = \quad & z_0 - \theta z_k + \theta c_k = z_0 + \theta (c_k - z_k). \end{aligned} \tag{4.7}$$

Der zugehörige Zielfunktionswert ist kleiner als z_0 falls $\theta > 0$ und $z_k - c_k > 0$. ■

Unter der Annahme, dass der Hauptsatz bereits vollständig bewiesen ist, haben wir ein hinreichendes Kriterium um zu entscheiden, ob es eine zulässige Basislösung mit einem kleineren Zielfunktionswert gibt. Man benötigt jetzt noch eine Methode zur Konstruktion dieser zulässigen Basislösung. Diese erfolgt mit Hilfe von θ.

Lemma 4.4 **Wahl von θ.** *Sei*

$$\theta = \min_{i=1,\ldots,m, x_{ik}>0} \frac{x_i}{x_{ik}} =: \frac{x_l}{x_{lk}}. \tag{4.8}$$

Dann ist die Lösung (4.6) zulässig.

Beweis: Man hat

$$\begin{aligned} \mathbf{a}_1 \left(x_1 - \frac{x_l}{x_{lk}} x_{1k}\right) + \ldots + \mathbf{a}_{l-1} \left(x_{l-1} - \frac{x_l}{x_{lk}} x_{l-1,k}\right) + \mathbf{a}_l \underbrace{\left(x_l - \frac{x_l}{x_{lk}} x_{l,k}\right)}_{=0} & \\ + \mathbf{a}_{l+1} \left(x_{l+1} - \frac{x_l}{x_{lk}} x_{l+1,k}\right) + \ldots + \mathbf{a}_m \left(x_m - \frac{x_l}{x_{lk}} x_{mk}\right) + \frac{x_l}{x_{lk}} \mathbf{a}_k & = \mathbf{b}, \end{aligned}$$

und die neue Lösung

$$\hat{x}_i = x_i - \frac{x_l}{x_{lk}} x_{ik}, \quad i = 1, \ldots, m, \ i \ne l, \quad \hat{x}_k = \frac{x_l}{x_{lk}}. \tag{4.9}$$

Alle Komponenten sind auf Grund der Konstruktion nichtnegativ und bei Ausschluss der Entartung sogar positiv. Damit hat man eine zulässige Lösung erhalten. Man hat also die Komponente x_l aus der Basisliste gestrichen und durch die Komponente x_k ersetzt. ■

Bemerkung 4.5 Nicht nach unten beschränkte Zielfunktion. Damit die Wahl (4.8) von θ überhaupt funktioniert, brauchen wir ein $x_{ik} > 0$. Falls es kein solches x_{ik} gibt, dann folgt, dass die Zielfunktion nach unten unbeschränkt ist. Man kann nämlich in diesem Fall θ beliebig groß wählen, da stets $x_i - \theta x_{ik} \geq 0$. Aus (4.7) folgt dann, dass unter der Bedingung $c_k - z_k < 0$ die Zielfunktion unbeschränkt nach unten ist. Fazit: Falls für ein $z_k - c_k > 0$ alle $x_{ik} \leq 0$, dann ist die Zielfunktion nicht von unten beschränkt und man breche die Simplexmethode ab. □

Bemerkung 4.6 Basiszyklus. Wenn man Entartung ausschließt, dann wird das Minimum in (4.8) wird für genau einen Index l angenommen. Es gilt auch die Umkehrung, dass falls der Index l in (4.8) nicht eindeutig bestimmt ist, dann hat man Ausartung. Ausartung kann zur Folge haben, dass gilt $z(\mathbf{x}_i) = z(\mathbf{x}_{i+1}) = \ldots$. Das nennt man einen Basiszyklus. □

Es gilt also, (4.9) ist eine zulässige Lösung mit einem kleineren Zielfunktionswert als die ursprüngliche Lösung. Damit bleibt nur noch die Basiseigenschaft von $\{\mathbf{a}_1, \ldots, \mathbf{a}_{l-1}, \mathbf{a}_k, \mathbf{a}_{l+1}, \ldots, \mathbf{a}_m\}$ zu prüfen.

Lemma 4.7 *Sei* $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ *ein System linear unabhängiger Vektoren und sei*

$$\mathbf{w} = \sum_{i=1}^{m} \mu_i \mathbf{w}_i, \quad \mu_l \neq 0. \tag{4.10}$$

Dann ist auch $\{\mathbf{w}_1, \ldots, \mathbf{w}_{l-1}, \mathbf{w}, \mathbf{w}_{l+1}, \ldots \mathbf{w}_m\}$ *ein System linear unabhängiger Vektoren.*

Beweis: Indirekter Beweis. Sei $\{\mathbf{w}_1, \ldots, \mathbf{w}_{l-1}, \mathbf{w}, \mathbf{w}_{l+1}, \ldots \mathbf{w}_m\}$ kein System linear unabhängiger Vektoren. Dann gibt es Zahlen $\alpha_1, \ldots, \alpha_{l-1}, \alpha_{l+1}, \ldots, \alpha_m, \alpha$, von denen wenigstens eine ungleich Null ist, so dass

$$\sum_{i=1, i\neq l}^{m} \alpha_i \mathbf{w}_i + \alpha \mathbf{w} = \mathbf{0}.$$

In diese Gleichung wird (4.10) eingesetzt. Es folgt

$$\sum_{i=1, i\neq l}^{m} (\alpha_i + \alpha\mu_i)\, \mathbf{w}_i + \alpha\mu_l \mathbf{w}_l = \mathbf{0}.$$

Die Vektoren $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ sind linear unabhängig, das heißt, alle Koeffizienten in dieser Gleichung müssen Null sein. Wegen $\mu_l \neq 0$ folgt dann $\alpha = 0$ und daraus $\alpha_i = 0$ für alle i. Damit ist gezeigt, dass $\{\mathbf{w}_1, \ldots, \mathbf{w}_{l-1}, \mathbf{w}, \mathbf{w}_{l+1}, \ldots \mathbf{w}_m\}$ ein System linear unabhängiger Vektoren ist. ■

Folgerung 4.8 *Das im Hauptsatz der Simplexmethode, Satz 4.3, konstruierte System von Vektoren ist mit der Wahl von θ nach (4.8) ein System linear unabhängiger Vektoren.*

Beweis: Da in der Linearkombination (4.3) der Faktor vor $\mathbf{a}_l$ gleich $x_{lk} (= \mu_l)$ ist und wir $x_{lk} > 0$ bei der Definition von l vorausgesetzt haben, lässt sich Lemma 4.7 anwenden und die Basiseigenschaft von $\{\mathbf{a}_1, \ldots, \mathbf{a}_{l-1}, \mathbf{a}_k, \mathbf{a}_{l+1}, \ldots, \mathbf{a}_m\}$ ist gewährleistet. ■

Bemerkung 4.9 Im allgemeinen ist der Hauptsatz der Simplexmethode solange anzuwenden, wie noch wenigstens ein $z_k - c_k > 0$ ist. Dabei kann man im allgemeinen nicht erwarten, falls noch q Größen $z_j - c_j > 0$ existieren, dass man noch q Schritte auszuführen hat. Gilt für alle $z_j - c_j \leq 0$, j – Index von Nichtbasisvariablen, so ist man in dem Sinne fertig, dass der Hauptsatz nicht mehr anwendbar ist. Der Hauptsatz gibt aber bisher nur ein hinreichendes und kein notwendiges Kriterium für die Existenz einer Basislösung mit einem kleineren Zielfunktionswert. □

Im folgenden Satz wird gezeigt, dass das Kriterium auch notwendig ist.

Satz 4.10 ***Optimalitätskriterium.*** *Eine zulässige Basislösung* $\mathbf{x} \in \mathbb{R}^n$ *mit* $z_0 = \sum_{i=1}^m c_i x_i$ *ist optimale Basislösung, wenn für alle* $j = m+1, \ldots, n$*, gilt* $z_j - c_j \le 0$.

Beweis: Sei $\mathbf{x} = (x_1, \ldots, x_m, 0, \ldots, 0)^T$. Des weiteren sei $\mathbf{y} = (y_1, \ldots, y_n)^T$ eine beliebige zulässige Lösung

$$\begin{aligned} \mathbf{a}_1 y_1 + \mathbf{a}_2 y_2 + \ldots + \mathbf{a}_n y_n &= \mathbf{b}, \quad \mathbf{y} \ge \mathbf{0}, && (4.11) \\ \text{mit } \; z^* &= \sum_{i=1}^n c_i y_i. && (4.12) \end{aligned}$$

Zu zeigen ist, dass $z_0 \le z^*$ für alle $\mathbf{y}$.

Durch (4.3) ist jeder Nichtbasisvektor mit Hilfe der Basis dargestellt. Jetzt wird diese Darstellung auf die Basisvektoren ausgedehnt

$$\mathbf{a}_j = x_{1j}\mathbf{a}_1 + \ldots + x_{mj}\mathbf{a}_m, \quad j = 1, \ldots, n,$$

wobei

$$x_{ij} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \ne j \end{cases}, i = 1, \ldots, m. \tag{4.13}$$

Weiter gilt die Darstellung (4.4) für z_j, $j = m+1, \ldots, n$. Mit (4.13) hat man eine analoge Darstellung für $j = 1, \ldots, m$, die sich letztlich auf $z_j = c_j$ reduziert. Zusammen mit der Voraussetzung gilt jetzt $z_j \le c_j$, $j = 1, \ldots, n$. Mit (4.12) folgt nun

$$\sum_{i=1}^n z_i y_i \le z^*. \tag{4.14}$$

Nun wird in (4.11) die Darstellung aller Spaltenvektoren durch die ersten m Spaltenvektoren eingesetzt

$$y_1 \sum_{i=1}^m x_{i1}\mathbf{a}_i + y_2 \sum_{i=1}^m x_{i2}\mathbf{a}_i + \ldots + y_n \sum_{i=1}^m x_{in}\mathbf{a}_i = \mathbf{b}.$$

Durch Umordnung nach den Basisvektoren folgt

$$\mathbf{a}_1 \sum_{j=1}^n y_j x_{1j} + \mathbf{a}_2 \sum_{j=1}^n y_j x_{2j} + \ldots + \mathbf{a}_m \sum_{j=1}^n y_j x_{mj} = \mathbf{b}. \tag{4.15}$$

Analog schreibt man (4.14) mit Hilfe von (4.4) und der entsprechenden Darstellung für $j = 1, \ldots, m$, mit (4.13) ($z_j = c_j$, $j = 1, \ldots, m$)

$$c_1 \sum_{j=1}^n y_j x_{1j} + c_2 \sum_{j=1}^n y_j x_{2j} + \ldots + c_m \sum_{j=1}^n y_j x_{mj} \le z^* \tag{4.16}$$

Der Vektor $\mathbf{x}$ ist eine zulässige Basislösung, das heißt, es gilt

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \ldots + \mathbf{a}_m x_m = \mathbf{b}. \tag{4.17}$$

Da $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ eine Basis ist, ist die Darstellung von $\mathbf{b}$ mit Hilfe dieser Vektoren eindeutig. Damit folgt aus (4.15) und (4.17)

$$x_i = \sum_{j=1}^n y_j x_{ij}, \; i = 1, \ldots, m.$$

Setzt man dies in (4.16), so erhält man

$$z_0 = \sum_{i=1}^m c_i x_i \le z^*.$$

■