

Kapitel 5

Die Simplexmethode

Bemerkung 5.1 Simplextabelle und Simplexmethode. Es werden folgende Bezeichnungen verwendet:

- das untersuchte Problem ist $\min_{\mathbf{x} \in \mathbb{R}^n} \{z = \mathbf{c}^T \mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$,
- die erste zulässige Basislösung sei $\mathbf{x} = (x_1, x_2, \dots, x_m, 0, \dots, 0)^T$, $\mathbf{x} \geq \mathbf{0}$, mit $z_0 = \mathbf{c}^T \mathbf{x}$,
- die Basisvektoren sind $A_B = (\mathbf{a}_1, \dots, \mathbf{a}_m)$,
- die Nichtbasisvektoren sind $A_N = (\mathbf{a}_{m+1}, \dots, \mathbf{a}_n)$,
- die Darstellung der Nichtbasisvektoren durch die Basis ist

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

- die Hilfsgrößen z_j sind

$$z_j = c_1x_{1j} + c_2x_{2j} + \dots + c_mx_{mj}, \quad j = m+1, \dots, n.$$

Diese Größen werden in der sogenannten Simplextabelle eingetragen:

			$m+1$	$m+2$	\dots	k	\dots	n
i	c_i	x_i	c_{m+1}	c_{m+2}	\dots	c_k	\dots	c_n
1	c_1	x_1	$x_{1,m+1}$	$x_{1,m+2}$	\dots	$x_{1,k}$	\dots	$x_{1,n}$
2	c_2	x_2	$x_{2,m+1}$	$x_{2,m+2}$	\dots	$x_{2,k}$	\dots	$x_{2,n}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
l	c_l	x_l	$x_{l,m+1}$	$x_{l,m+2}$	\dots	$x_{l,k}$	\dots	$x_{l,n}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
m	c_m	x_m	$x_{m,m+1}$	$x_{m,m+2}$	\dots	$x_{m,k}$	\dots	$x_{m,n}$
z_0			$z_{m+1} - c_{m+1}$	$z_{m+2} - c_{m+2}$	\dots	$z_k - c_k$	\dots	$z_n - c_n$
Basisteil			Nichtbasisteil					

Bei der Simplexmethode folgt man jetzt im wesentlichen dem Beweis des Hauptsatzes. Sei $z_k - c_k > 0$. Gilt für mehrere Indizes $j \in \{m+1, \dots, n\}$, dass $z_j - c_j > 0$, so nehme man zum Beispiel einen Index, bei dem die Differenz maximal ist

$$z_k - c_k := \max_{j=m+1, \dots, n} z_j - c_j.$$

Dann liegt x_k als Nichtbasisvariable vor, die in die Basis soll. Nun bestimmt man

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

das heißt, x_l soll aus der Basis raus. □

Definition 5.2 Hauptspalte, Hauptzeile, Hauptelement, Pivotelement. Die Spalte k nennt man Hauptspalte, die Zeile l heißt Hauptzeile und das Element x_{lk} heißt Hauptelement oder Pivotelement. \square

Bemerkung 5.3 Berechnung der Einträge der neuen Simplextabelle. Die neue Basislösung sei

$$(\hat{x}_1, \dots, \hat{x}_{l-1}, \hat{x}_k, \hat{x}_{l+1}, \dots, \hat{x}_m, 0, \dots, 0)^T. \quad (5.1)$$

Nun müssen die Elemente der neuen Simplextabelle bestimmt werden:

1. Man benötigt insbesondere eine Darstellung von (5.1). Aus (4.9) erhält man

$$\hat{x}_i = x_i - \frac{x_l}{x_{lk}} x_{ik}, \quad i = 1, \dots, m; i \neq l; \quad \hat{x}_k = \frac{x_l}{x_{lk}}. \quad (5.2)$$

2. Aus (4.3) folgt für $j = k$

$$\begin{aligned} \mathbf{a}_l &= \frac{1}{x_{lk}} (\mathbf{a}_k - x_{1k} \mathbf{a}_1 - \dots - x_{l-1,k} \mathbf{a}_{l-1} - x_{l+1,k} \mathbf{a}_{l+1} - \dots - x_{mk} \mathbf{a}_m) \\ &= -\frac{x_{1k}}{x_{lk}} \mathbf{a}_1 - \dots - \frac{x_{l-1,k}}{x_{lk}} \mathbf{a}_{l-1} + \frac{\mathbf{a}_k}{x_{lk}} - \frac{x_{l+1,k}}{x_{lk}} \mathbf{a}_{l+1} - \dots - \frac{x_{mk}}{x_{lk}} \mathbf{a}_m. \end{aligned} \quad (5.3)$$

Damit haben wir eine Darstellung des neuen Nichtbasisvektors \mathbf{a}_l durch die neue Basis und die neuen Elemente der alten Hauptspalte sind

$$\hat{x}_{kl} = \frac{1}{x_{lk}}, \quad \hat{x}_{il} = -\frac{x_{ik}}{x_{lk}}, \quad i = 1, \dots, m, \quad i \neq k. \quad (5.4)$$

3. Für den Rest erhält man, beispielhaft an \mathbf{a}_n gezeigt, die folgende Darstellung, wobei man in der ersten Gleichung die alte Basisdarstellung (4.3) nutzt:

$$\begin{aligned} \mathbf{a}_n &= x_{1n} \mathbf{a}_1 + \dots + x_{l-1,n} \mathbf{a}_{l-1} + x_{l+1,n} \mathbf{a}_{l+1} + \dots + x_{mn} \mathbf{a}_n + x_{ln} \underbrace{\mathbf{a}_l}_{(5.3)} \\ &= \left(x_{1n} - \frac{x_{1k} x_{ln}}{x_{lk}} \right) \mathbf{a}_1 + \dots + \left(x_{l-1,n} - \frac{x_{l-1,k} x_{ln}}{x_{lk}} \right) \mathbf{a}_{l-1} + \frac{x_{ln}}{x_{lk}} \mathbf{a}_k \\ &\quad + \dots + \left(x_{mn} - \frac{x_{mk} x_{ln}}{x_{lk}} \right) \mathbf{a}_m. \end{aligned}$$

Man erhält also die folgenden Koeffizienten für die neue Basisdarstellung

$$\begin{aligned} \hat{x}_{kj} &= \frac{x_{lj}}{x_{lk}}, \quad j = m+1, \dots, n, \quad j \neq k, \quad (5.5) \\ \hat{x}_{ij} &= x_{ij} - \underbrace{\frac{x_{lj}}{x_{lk}} x_{ik}}_{\hat{x}_{kj}}, \quad i = 1, \dots, m, \quad i \neq k, \quad j = m+1, \dots, n, \quad j \neq l. \end{aligned} \quad (5.6)$$

4. Die Elemente $z_0, z_{m+1} - c_{m+1}, \dots, z_n - c_n$ transformieren sich ebenfalls nach den obigen Regeln. *Übungsaufgabe*

Damit sind alle Elemente der neuen Simplextabelle berechnet. Zur Berechnung von \hat{x}_{ij} benötigt man die im Rechteck angeordneten Elemente x_{ij}, x_{lj}, x_{lk} und x_{ik} der alten Simplextabelle. Deshalb spricht man auch von der Rechteckregel. \square

Bemerkung 5.4 Basisform der Simplexmethode.

1. Normalform des linearen Programms herstellen.
2. Erste zulässige Basislösung angeben.
3. Simplextabelle zu dieser Basislösung erstellen.
4. Existieren Bewertungen $z_j - c_j > 0$? Wenn ja, gehe zu 6.
5. Sind alle Bewertungen $z_j - c_j < 0$?
 - Wenn ja, einzige Optimallösung gefunden, *Simplexmethode beendet*.
 - Wenn nicht, dann gibt es außer negativen Bewertungen $z_j - c_j$ nur noch verschwindende. Das Optimum nicht eindeutig. Man hat ein Optimum gefunden, *beende Simplexmethode*.
6. Wähle die Hauptspalte, also die Spalte, zu der das größte $z_j - c_j > 0$, $j = k$ gehört.
7. Falls $x_{ik} \leq 0$ für alle $i = 1, \dots, m$, so ist die Zielfunktion nach unten nicht beschränkt, *beende Simplexmethode*.
8. Bestimme θ zur Festlegung der Hauptzeile und des Pivotelements.
9. Basistransformation:
 - 9.1 Ersetze das Pivoelement durch seinen Kehrwert, siehe (5.4).
 - 9.2 Multipliziere die übrigen Elemente der Hauptzeile mit diesem Kehrwert, einschließlich x_i , siehe (5.2) und (5.5).
 - 9.3 Multipliziere die übrigen Elemente der Hauptspalte mit dem negativen Kehrwert, siehe (5.4).
 - 9.4 Vermindere die nicht in einer Hauptreihe stehenden Elemente, einschließlich der übrigen Werte von x_i und der letzten Zeile, um das Produkt der zugehörigen Hauptreihenelemente (Rechteckregel). Dabei nimmt man für das Pivoelement schon den neuen Wert und für die übrigen Elemente die alten Werte, siehe (5.2) und (5.6).
10. Gehe zu 4.

Beispiel 5.5 Wir betrachten das lineare Programm

$$z = -3x_1 - 2x_2 - 4x_3 - x_4 \rightarrow \min !$$
$$\begin{pmatrix} 2 & 2 & 3 & 0 & 1 & 0 & 0 \\ 1 & 3 & 0 & 2 & 0 & 1 & 0 \\ 1 & 1 & 5 & 2 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} 700 \\ 400 \\ 500 \end{pmatrix}$$
$$\mathbf{x} \geq \mathbf{0}.$$

Bekannt sei eine erste zulässige Basislösung $x_1 = 350$, $x_4 = 25$, $x_7 = 100$, die den Zielfunktionswert $z = -1075$ besitzt. Die Basisvektoren sind demzufolge

$$\mathbf{a}_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{a}_4 = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}, \quad \mathbf{a}_7 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Gesucht ist nun die Darstellung der Nichtbasisvektoren durch die Basisvektoren. Setze $A_B = (\mathbf{a}_1, \mathbf{a}_4, \mathbf{a}_7)$ und $A_N = (\mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_5, \mathbf{a}_6)$. Dann ist die Matrix X der Simplexkoeffizienten gesucht, für die gilt

$$A_N = A_B X \implies X = A_B^{-1} A_N.$$

Man erhält hier

$$X = \begin{pmatrix} 1 & 3/2 & 1/2 & 0 \\ 1 & -3/4 & -1/4 & 1/2 \\ -2 & 5 & 0 & -1 \end{pmatrix}.$$

Daraus ergibt sich

$$\begin{aligned} z_2 &= c_1x_{12} + c_4x_{42} + c_7x_{72} = (-3)1 + (-1)1 + 0(-2) = -4, \\ z_3 &= -9/2 + 3/4 + 0 = -15/4, \\ z_5 &= -3/2 + 1/4 + 0 = -5/4, \\ z_6 &= 0 - 1/2 + 0 = -1/2 \end{aligned}$$

und somit

$$z_2 - c_2 = -2, \quad z_3 - c_3 = 1/4, \quad z_5 - c_5 = -5/4, \quad z_6 - c_6 = -1/2.$$

Damit erhält man folgende Simplextabelle:

i	c_i	x_i	2	3	5	6
			-2	-4	0	0
1	-3	350	1	3/2	1/2	0
4	-1	25	1	-3/4	-1/4	1/2
7	0	100	-2	5	0	-1
		-1075	-2	1/4	-5/4	-1/2

Es gibt nur einen Index k mit $z_k - c_k > 0$, nämlich $k = 3$. Damit ist die Hauptspalte bestimmt (Schritt 6). Zur Bestimmung der Hauptzeile (Schritt 8) berechnet man θ :

$$\theta = \min_{x_{i3} > 0, i \in \{1,4,7\}} \left(\frac{x_i}{x_{i3}} \right) = \min \left\{ \frac{350}{3/2}, \frac{100}{5} \right\} = 20$$

für $i = 7$. Damit ist der Hauptzeilenindex $l = 7$ und das Pivotelement $x_{73} = 5$. Nun führt man die Basistransformation aus (Schritt 9):

i	c_i	x_i	2	7	5	6
			-2	0	0	0
1	-3	320	8/5	-3/10	1/2	3/10
4	-1	40	7/10	3/20	-1/4	7/20
3	-4	20	-2/5	1/5	0	-1/5
		-1080	-19/10	-1/20	-5/4	-9/20

Den neuen Wert für x_1 erhält man beispielsweise aus

$$x_1 = 350 - \frac{3}{2}100\frac{1}{5} = 350 - 30 = 320.$$

Da in der neuen Simplextabelle alle Werte $z_j - c_j < 0$, $j \in \{2, 5, 6, 7\}$, hat man die einzige Optimallösung bestimmt: $\mathbf{x} = (320, 0, 20, 40, 0, 0, 0)^T$. \square

Bemerkung 5.6 Angenommen, man hat in einer Simplextabelle mehrere $z_j - c_j > 0$. Zu einer dieser Spalten mögen nur Koeffizienten $x_{ij} \leq 0$ gehören. Dann ist die Zielfunktion unbeschränkt. \square

Beispiel 5.7 Zur Ausartung. Wir betrachten das lineare Programm

$$\begin{aligned} z &= -x_1 \rightarrow \min ! \\ \begin{pmatrix} 1 & 1 & 1 & 0 \\ 4 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} &= \begin{pmatrix} 1 \\ 4 \end{pmatrix} \\ \mathbf{x} &\geq \mathbf{0}. \end{aligned}$$

Eine zulässige Basislösung, die gleichzeitig ein Optimum ist, ist $\mathbf{x} = (1, 0, 0, 0)^T$. Wir nehmen als Basisvariablen x_1 und x_2 . Da x_2 verschwindet, ist die Basislösung ausgeartet. Man hat

$$A_B = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}, \quad A_N = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

und erhält die Simplextabelle

i	c_i	x_i	3	4
1	-1	1	-1/3	1/3
2	0	0	4/3	-1/3
		-1	1/3	-1/3

Die Simplexmethode sagt uns an dieser Stelle nicht, dass das Optimum bereits erreicht ist! Gemäß Simplexmethode muss x_3 in die Basis anstelle von x_2 eingeführt werden. Man erhält die Simplextabelle

i	c_i	x_i	2	4
1	-1	1	1/4	1/4
3	0	0	3/4	-1/4
		-1	-1/4	-1/4

Damit ist das Optimalitätskriterium der Simplexmethode erfüllt und diese wird beendet. Man hat für das Optimum $\mathbf{x} = (1, 0, 0, 0)^T$ mit diesen beiden Simplextabellen zwei unterschiedliche Basisdarstellungen. Der Zielfunktionswert hat sich im Simplexschritt nicht verändert, es wurde lediglich die Basis gewechselt. \square

Kapitel 6

Bestimmung einer ersten zulässigen Basislösung

Ein Problem, was man für die Durchführung der Simplexmethode lösen muss, ist die Bestimmung einer ersten zulässigen Basislösung. Wie gut das geht, hängt auch vom konkreten Problem ab.

Bemerkung 6.1 1. Fall. Liegt

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

vor und gilt $\mathbf{b} \geq \mathbf{0}$. Dann führt man Schlupfvariablen ein und setzt $\mathbf{x} = (0, \dots, 0, \mathbf{b}^T)^T$. \square

Bemerkung 6.2 2. Fall, Engpassmethode. Liegt das lineare Programm in der Gestalt

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

vor mit $A = (a_{ij}), i = 1, \dots, m; j = 1, \dots, n, \mathbf{b} = (b_1, \dots, b_m)^T, a_{ij} \geq 0, b_i \geq 0$ für alle i, j . Dann kann man mit einer sogenannten Engpassmethode zur ersten zulässigen Basislösung gelangen:

1. Ordne die Variablen nach wachsenden Zielfunktionskoeffizienten c_i , Beispiel

$$z = -10x_1 - 6x_2 - 4x_3 - 3x_4 - 5x_5 \rightarrow \min !$$
$$\begin{pmatrix} 2 & 0 & 4 & 0 & 2 & 1 & 0 & 0 & 0 \\ 2 & 3 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 5 & 0 & 0 & 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_9 \end{pmatrix} = \begin{pmatrix} 8 \\ 6 \\ 20 \\ 8 \end{pmatrix}$$
$$\mathbf{x} \geq \mathbf{0}.$$

Dann ist die Ordnung x_1, x_2, x_5, x_3, x_4 .

2. In der festgelegten Reihenfolge werden die Variablen mit dem größtmöglichen Wert genommen, so dass die Nebenbedingungen erfüllt sind. Im Beispiel beginnt man mit $x_1 = 3$
3. Man setzt diesen Wert ein und entfernt die Variable damit aus den Nebenbedingungen. Im Beispiel ergibt sich

$$\begin{pmatrix} 0 & 4 & 0 & 2 & 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_9 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 5 \\ 8 \end{pmatrix}.$$

Aus der zweiten Gleichung folgt $x_2 = x_3 = x_7 = 0$, welche Werte man auch gleich einsetzen kann. Damit vereinfacht sich das System der Nebenbedingungen zu

$$\begin{pmatrix} 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 3 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_4 \\ x_5 \\ x_6 \\ x_8 \\ x_9 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 5 \\ 8 \end{pmatrix}. \quad (6.1)$$

4. Gehe zu 2.

Im Beispiel betrachtet man als nächstes x_5 , da ja bereits $x_2 = 0$ gilt. Der maximale Wert von x_5 , so dass (6.1) erfüllt ist, beträgt $x_5 = 1$. Einsetzen ergibt

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_4 \\ x_6 \\ x_8 \\ x_9 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 8 \end{pmatrix}. \quad (6.2)$$

Damit folgt $x_6 = 0$. Da ja auch schon $x_3 = 0$ gilt, wird nun x_4 betrachtet. Der maximale Wert von x_4 , so dass (6.2) erfüllt ist, ist $x_4 = 2$. Man erhält

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_8 \\ x_9 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

Nun bestimmt man die letzten beiden Werte und erhält als erste zulässige Basislösung $\mathbf{x} = (3, 0, 0, 2, 1, 0, 0, 0, 2)^T$. \square

Bemerkung 6.3 Hat man bei der Engpassmethode nicht genügend Variablen, dann führt man künstliche Variablen ein. \square

Bemerkung 6.4 Anderes Ordnungsprinzip der Variablen im Fall, dass die Koeffizienten von unterschiedlicher Größenordnung sind. Wir betrachten das lineare Programm

$$z = 10x_1 + 20x_2 + 30x_3 + 40x_4 + 50x_5 \rightarrow \min !$$

$$\begin{pmatrix} 1 & 10 & 50 & 1 & 0 \\ 2 & 20 & 50 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_5 \end{pmatrix} = \begin{pmatrix} 100 \\ 101 \end{pmatrix}$$

$$\mathbf{x} \geq \mathbf{0}.$$

Nach dem obigen Ordnungsprinzip hat man die Reihenfolge x_1, x_2, x_3, x_4, x_5 und erhält mit der Engpassmethode die erste zulässige Basislösung *Übungsaufgabe*

$$x_1 = \frac{101}{2}, \quad x_4 = \frac{99}{2} \implies z = 2485.$$

Man erhält jedoch mit einer anderen Basislösung einen schon viel kleineren Zielfunktionswert

$$x_3 = 2, \quad x_5 = 1 \implies z = 110.$$

In diesem Fall ist das Ordnungsprinzip

$$\min_{j, c_j \neq 0} \left\{ c_j \min_{i, a_{ij} \neq 0} \left\{ \frac{b_i}{a_{ij}} \right\} \right\}$$

günstiger. Bei diesem Ordnungsprinzip wird auch die Größe der Matrixeinträge und der rechten Seite beachtet. Im Beispiel kann man x_3 wegen der großen Matrixeinträge nur relativ klein wählen, wenn man die Nebenbedingungen nicht verletzen will.

Im Gegensatz dazu kann man x_1 sehr groß wählen ohne die Nebenbedingungen zu verletzen. Der etwas höhere Koeffizient vor x_3 in der Zielfunktion führt wegen des viel kleineren möglichen Wertes dieser Variablen letztlich auf einen kleineren Beitrag als $10x_1$ mit großem x_1 . \square

Bemerkung 6.5 3. Fall, Bestimmung mit Hilfe der Simplexmethode. Die erste zulässige Basislösung soll jetzt

- ohne spezielle Voraussetzungen und
- mit Hilfe der Simplexmethode

bestimmt werden. Dazu betrachten wir

$$\sum_{j=1}^n c_j x_j = \mathbf{c}^T \mathbf{x} \rightarrow \min ! \quad (6.3)$$

$$A\mathbf{x} = \mathbf{b}, \quad (6.4)$$

$$\mathbf{x} \geq \mathbf{0}. \quad (6.5)$$

und nehmen $\mathbf{b} \geq \mathbf{0}$ an. Das kann immer durch Multiplikation der entsprechenden Gleichungen mit einer negativen Zahl erreicht werden. Dem Problem (6.3) – (6.5) wird die Hilfsaufgabe

$$\sum_{i=1}^m x_{n+i} \rightarrow \min ! \quad (6.6)$$

$$\sum_{j=1}^n a_{ij} x_j + x_{n+i} = b_i, \quad i = 1, \dots, m, \quad (6.7)$$

$$x_j \geq 0, \quad j = 1, \dots, m+n \quad (6.8)$$

zugeordnet. Die Variablen x_{n+1}, \dots, x_{n+m} heißen künstliche Variablen. Zur Bestimmung der ersten zulässigen Basislösung von (6.3) – (6.5) wird eine Zweiphasenmethode verwendet:

- 1. Phase. Wähle als erste zulässige Basislösung für (6.6) – (6.8)

$$x_i = 0, \quad i = 1, \dots, n, \quad x_{n+i} = b_i, \quad i = 1, \dots, m.$$

- 2. Phase. Löse (6.6) – (6.8) mit der Simplexmethode.

Es stellt sich nun die Frage, ob man auf diesem Wege schließlich eine erste zulässige Basislösung für (6.3) – (6.5) erhält. \square

Im nächsten Satz wird gezeigt, dass eine Lösung von (6.6) – (6.8) existiert, falls man Ausartung ausschließt.

Lemma 6.6 *Unter der Annahme, dass (6.6) – (6.8) keine ausgearteten Basislösungen besitzt, liefert die Simplexmethode nach endlich vielen Schritten eine optimale Lösung des linearen Programms (6.6) – (6.8).*

Beweis: Da Ausartung per Annahme ausgeschlossen ist, kann kein Basiszyklus auftreten. Somit verringert die Simplexmethode in jedem Schritt den Zielfunktionswert. Es ist dann nur noch die Beschränktheit von unten der Zielfunktion (6.6) über (6.7) bis (6.8) zu zeigen. Das ist offensichtlich, da (6.6) eine Summe nichtnegativer reeller Zahlen ist, die durch Null nach unten beschränkt ist. \blacksquare

Nun wird eine Bedingung angegeben, mit welcher man aus dem Optimum des Hilfsproblems (6.6) – (6.8) eine erste zulässige Basislösung von (6.3) – (6.5) erhält.

Satz 6.7 *Sei $\tilde{\mathbf{x}}_0$ eine Optimallösung der künstlichen Aufgabe (6.6) – (6.8) mit dem zugehörigen Zielfunktionswert \tilde{z}_0 . Gilt $\tilde{z}_0 = 0$, so sind die ersten n Komponenten von $\tilde{\mathbf{x}}_0$ eine zulässige Basislösung der Aufgabe (6.3) – (6.5). Gilt jedoch $\tilde{z}_0 > 0$, so besitzt (6.3) – (6.5) keine zulässige Basislösung.*

Beweis: Aus $\tilde{z}_0 = 0$ folgt $x_{n+i} = 0, i = 1, \dots, m$, das heißt im Optimum verschwinden alle künstlichen Variablen. Also hat $\tilde{\mathbf{x}}_0$ die Gestalt

$$\tilde{\mathbf{x}}_0 = (x_1, \dots, x_n, \underbrace{0, \dots, 0}_m)^T.$$

Da $\tilde{\mathbf{x}}_0$ mit der Simplexmethode konstruiert wurde, folgt dass $\tilde{\mathbf{x}}_0$ eine zulässige Basislösung von (6.3) – (6.5) ist.

Sei nun $\tilde{z}_0 > 0$. Der Beweis wird indirekt geführt, das heißt, wir nehmen an, dass (6.3) – (6.5) die zulässige Basislösung $\bar{\mathbf{x}} = (\bar{x}_1, \dots, \bar{x}_n)^T$ besitzt. Dann besitzt jedoch (6.6) – (6.8) die zulässige Basislösung $(\bar{x}_1, \dots, \bar{x}_n, 0, \dots, 0)^T$ mit dem zugehörigen Zielfunktionswert (6.6) $\bar{z} = 0$. Das ist im Widerspruch zur Annahme dass \tilde{z}_0 der minimale Wert ist. ■

Bemerkung 6.8 4. Fall, M–Methode. Die M–Methode. Es wird das lineare Programm (6.3) – (6.5) betrachtet und diesem die folgende Hilfsaufgabe zugeordnet

$$\sum_{j=1}^n c_j x_j + M \sum_{i=1}^m x_{n+i} \rightarrow \min ! \quad (6.9)$$

$$\sum_{j=1}^n a_{ij} x_j + x_{n+i} = b_i \quad i = 1, \dots, m, \quad (6.10)$$

$$\mathbf{x} \geq \mathbf{0}. \quad (6.11)$$

Bei dieser Aufgabe muss der Straffaktor $M > 0$ hinreichend groß gewählt werden, damit im Optimum die Variablen x_{n+1}, \dots, x_{n+m} verschwinden. Die Schwierigkeit besteht darin, dass man im allgemeinen nicht von vornherein festlegen kann, wie groß M zu wählen ist. □

Möglich sind Aussagen folgender Gestalt:

Satz 6.9 *Es existiert ein $M_0 > 0$, so dass für alle $M > M_0$ aus der Lösbarkeit von (6.3) – (6.5) die Lösbarkeit von (6.9) – (6.11) mit $x_{n+1} = \dots = x_{n+m} = 0$ folgt.*

Beweis: Siehe Literatur. ■

Der Vorteil der M–Methode im Vergleich zur Herangehensweise von Fall 3 wird mit folgendem Satz beschrieben.

Satz 6.10 *Falls (6.9) – (6.11) eine Lösung*

$$\tilde{\mathbf{x}} = (x_1, \dots, x_n, \underbrace{0, \dots, 0}_m)^T$$

besitzt, so ist $\mathbf{x} = (x_1, \dots, x_n)^T$ bereits eine Optimallösung von (6.3) – (6.5).

Beweis: Das sieht man durch Einsetzen von $\tilde{\mathbf{x}}$ in (6.9) – (6.11). ■

Kapitel 7

Zur Ausartung

Bemerkung 7.1 Nach Definition 3.4 liegt Ausartung dann vor, wenn mindestens eine der Variablen $x_i, i = 1 \dots, m$, einer zulässigen Basislösung verschwindet. Das dahinterliegende Problem ist, dass die Zuordnung Ecke – zulässige Basislösung nicht eindeutig ist. Eine Ecke des Polyeders kann Basislösung zu verschiedenen Basen sein. Das kann aber nur bei ausgearteten Basislösungen auftreten. \square

Beispiel 7.2 Betrachte das lineare Programm mit

$$z = x_1 + x_2 - x_3 \rightarrow \min !, \quad A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Der einzige Extrempunkt ist $\mathbf{x} = (0, 0, 1)^T$. Zulässige Basen sind

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ und } \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Der Grund für die Nichteindeutigkeit der Basis besteht darin, dass es zu viele Nebenbedingungen gibt, die den Extrempunkt bestimmen. In diesem Beispiel ist er durch die beiden Gleichungen

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ und } \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

gleichermaßen gegeben. Das haben wir bereits in den Beispielen 3.5 (2. Teil) und 5.7 gesehen. \square

Bemerkung 7.3

- In der Praxis stellt sich heraus, dass die meisten zu lösenden linearen Programme ausgeartet sind.
- In der Simplexmethode ist es möglich, dass im Falle der Ausartung der zulässigen Basislösung nur ein Basiswechsel stattfindet, siehe Beispiel 5.7. Das kann zu einem unendlichen Zyklus werden, einem sogenannten Basiszyklus. Es ist jedoch möglich, Ausartung prinzipiell auszuschließen, beispielsweise mit der Methode der ε -Störung, beziehungsweise einen Basiszyklus zu umgehen, mit der lexikographischen Simplexmethode. \square

7.1 Die Methode der ε -Störung

Bemerkung 7.4 Herangehensweise. Wir betrachten das lineare Programm

$$\min_{\mathbf{x} \in \mathbb{R}^n} \{ \mathbf{c}^T \mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \}. \quad (7.1)$$

Sei $\mathbf{x} = (x_1, \dots, x_m, 0, \dots, 0)^T$ eine zulässige Basislösung mit den Basisvektoren $\mathbf{a}_1, \dots, \mathbf{a}_m$:

$$\begin{aligned} \mathbf{a}_1 x_1 + \dots + \mathbf{a}_m x_m &= \mathbf{b}, \\ \mathbf{a}_1 x_{1j} + \dots + \mathbf{a}_m x_{mj} &= \mathbf{a}_j, \quad j = 1, \dots, n. \end{aligned} \quad (7.2)$$

Sei $\varepsilon > 0$ vorgegeben und sei $A_B = (\mathbf{a}_1, \dots, \mathbf{a}_m)$ die Matrix der Basisvektoren. Dann betrachtet man anstelle (7.1) ein lineares Programm mit gestörten Nebenbedingungen

$$\begin{aligned} \mathbf{c}^T \mathbf{x} &\rightarrow \min ! \\ A_B \mathbf{x} + \sum_{j=1}^n \varepsilon^j (7.2) &= \mathbf{b} \implies \end{aligned} \quad (7.3)$$

$$A_B \mathbf{x} + \sum_{j=1}^n \varepsilon^j \mathbf{a}_j = \mathbf{a}_1 \left(x_1 + \sum_{j=1}^n x_{1j} \varepsilon^j \right) + \dots + \mathbf{a}_m \left(x_m + \sum_{j=1}^n x_{mj} \varepsilon^j \right) = \mathbf{b}.$$

Mit den Nebenbedingungen (7.3) hat man für hinreichend kleines ε die zulässige Basislösung

$$x_i^{(\varepsilon)} := x_i + \sum_{j=1}^n x_{ij} \varepsilon^j = x_i + \varepsilon^i + \sum_{j=m+1}^n x_{ij} \varepsilon^j,$$

da für $i = 1, \dots, m$, gilt $x_{ij} = \delta_{ij}$. Die Eigenschaft der Basislösung folgt daraus, dass die Basis nicht geändert wurde und die Nebenbedingung in (7.3) erfüllt ist. Die Zulässigkeit folgt für hinreichend kleines ε aus $\varepsilon^i > 0$ und $\varepsilon^i \gg \varepsilon^j$ für $i < j$

$$\varepsilon^i > \left| \sum_{j=m+1}^n x_{ij} \varepsilon^j \right| \implies x_i^{(\varepsilon)} > 0.$$

Die Basislösung $\mathbf{x}^{(\varepsilon)}$ ist also für hinreichend kleine ε nicht ausgeartet. Der zugehörige Zielfunktionswert ist

$$\begin{aligned} z_0^{(\varepsilon)} &= \sum_{i=1}^m c_i x_i + \sum_{i=1}^m c_i \left(\sum_{j=1}^n x_{ij} \varepsilon^j \right) = \sum_{i=1}^m c_i x_i + \sum_{j=1}^n \left(\sum_{i=1}^m c_i x_{ij} \right) \varepsilon^j \\ &= \sum_{i=1}^m c_i x_i + \sum_{j=1}^n z_j \varepsilon^j. \end{aligned}$$

Im Bild 7.1 wird die Störung der Nebenbedingungen graphisch veranschaulicht. Im dicken Punkt schneiden sich drei Geraden. Das führt dazu, dass die Zuordnung Ecke – Basislösung nicht eindeutig ist. Man hat Ausartung. Durch die Störung der Nebenbedingungen (durchgezogene Geraden) erreicht man, dass es nur noch Schnittpunkte mit genau zwei Geraden gibt. \square

Bemerkung 7.5 Berechnung von θ . In der Simplexmethode benötigt man die Größe θ , siehe (4.8). Sei $z_k - c_k > 0$. Dann berechnet sich θ in der Methode der ε -Störung durch

$$\theta = \min_{i=1, \dots, m; x_{ik} > 0} \frac{x_i^{(\varepsilon)}}{x_{ik}} = \min_{i=1, \dots, m; x_{ik} > 0} \frac{x_i + \varepsilon^i + \sum_{j=m+1}^n x_{ij} \varepsilon^j}{x_{ik}}. \quad (7.4)$$

\square

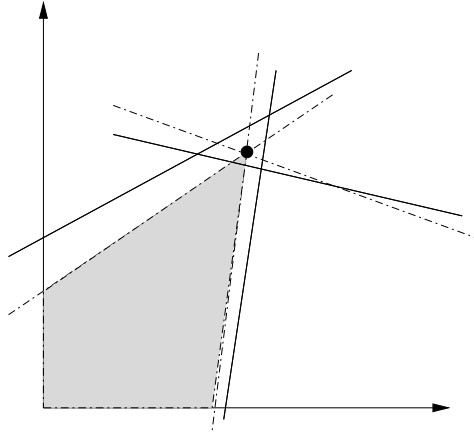


Abbildung 7.1: Veranschaulichung der Störung der Nebenbedingungen.

Satz 7.6 Sei $\mathbf{x} = (x_1, \dots, x_m, 0, \dots, 0)^T$ eine zulässige Basislösung der Originalaufgabe (7.1). Falls

$$\theta = \min_{i=1, \dots, m; x_{ik} > 0} \frac{x_i}{x_{ik}} = 0$$

gilt (Ausartung), dann gibt es ein $\bar{\varepsilon} > 0$ dergestalt, dass

$$\theta = \min_{i=1, \dots, m; x_{ik} > 0} \frac{x_i^{(\varepsilon)}}{x_{ik}} = \frac{x_l^{(\varepsilon)}}{x_{lk}} > 0 \quad \forall \varepsilon \in (0, \bar{\varepsilon}) \quad (7.5)$$

und der Index l ist im gestörten Problem (7.3) eindeutig bestimmt.

Beweis: Aus Bemerkung 7.4 folgt $x_i^{(\varepsilon)} > 0$, $i = 1, \dots, m$, und damit $\theta > 0$ in (7.5) für $\varepsilon \in (0, \bar{\varepsilon})$.

Die Eindeutigkeit von l wird indirekt bewiesen. Sei l also nicht eindeutig bestimmt, das heißt es gibt zwei Indizes $l_1 \neq l_2$, $1 \leq l_1, l_2 \leq m$, mit

$$\frac{x_{l_1} + \varepsilon^{l_1} + \sum_{j=m+1}^n x_{l_1 j} \varepsilon^j}{x_{l_1 k}} = \frac{x_{l_2} + \varepsilon^{l_2} + \sum_{j=m+1}^n x_{l_2 j} \varepsilon^j}{x_{l_2 k}}$$

für alle $\varepsilon \in (0, \bar{\varepsilon})$. Die beiden Terme sind Polynome in ε . Diese sind genau dann gleich für alle $\varepsilon \in (0, \bar{\varepsilon})$, wenn sie in allen Koeffizienten übereinstimmen. Insbesondere müssen die Koeffizienten vor den Termen mit Potenz l_1 und l_2 gleich sein. Ist $l_1 \neq l_2$, so ist für den linken Term der Koeffizient vor ε^{l_1} ungleich Null und für den rechten Term gleich Null. Für den Koeffizienten vor ε^{l_2} gilt sinngemäß das gleiche. Das heißt, diese Koeffizienten können nur dann gleich sein, wenn $l_1 = l_2$, im Widerspruch zur Annahme. ■

Prinzipiell kann diese Manipulation in jedem Simplexschritt durchgeführt werden und man kann damit sichern, dass l stets eindeutig bestimmt ist. Diese Vorgehensweise ist für jeden Eckpunkt des zulässigen Bereichs ausgeführt zu denken. Da die Anzahl der Eckpunkte endlich ist, erhält man folgenden Satz.

Satz 7.7 Zu jedem linearen Optimierungsproblem existiert bei geeigneter Wahl von $\varepsilon \in (0, \varepsilon^*)$ ein gestörtes lineares Optimierungsproblem, so dass dieses keine ausgeartete zulässige Basislösung besitzt. Für $\varepsilon \rightarrow 0$ konvergiert das Optimum des gestörten Problems (7.3) zum Optimum des Originalproblems (7.1).

Bemerkung 7.8 Praktische Umsetzung der Methode der ε -Störung. Trotz dieser schönen Theorie macht man das alles bei praktischen Problemen nicht. Für

diese wird vorgeschlagen: Falls in einer zulässigen Basislösung wenigstens ein Wert $x_i = 0$ bestimmt wurde, so kann $\theta = 0$ sein. Wähle dann

$$l = \min_{x_{ik} > 0} \{i : x_i = 0\},$$

wobei i über alle Basisvariablen läuft und k der Index der festgelegten Hauptspalte ist, und transformiere mit diesem Index l . Theoretisch besteht die Gefahr eines Zyklus, in der Praxis ist das aber eher unwahrscheinlich. \square

7.2 Die lexikographische Simplexmethode

Bemerkung 7.9 Idee. Bei der lexikographischen Simplexmethode erfolgt die Auswahl der zu tauschenden Basisvektoren so, dass keine Wiederholungen auftreten können. Mit dieser Vorgehensweise wird nicht die Ausartung behoben, sondern es wird verhindert, dass Basiszyklen auftreten. \square

Definition 7.10 Lexikopositiver Vektor. Ein Vektor $\mathbf{x} \in \mathbb{R}^n$ heißt lexikopositiv, falls $\mathbf{x} = (0, \dots, 0, x_i, x_{i+1}, \dots, x_n)^T$ mit $i \geq 1$ und $x_i > 0$. Das heißt, die erste von Null verschwindende Komponente ist positiv. Die Schreibweise ist

$$\mathbf{x} >_l \mathbf{0}.$$

Sei $\mathbf{y} \in \mathbb{R}^n$. Dann ist $\mathbf{y} >_l \mathbf{x}$ genau dann, wenn $\mathbf{y} - \mathbf{x} >_l \mathbf{0}$. \square

Bemerkung 7.11 Die lexikographische Simplexmethode. Wir betrachten das lineare Programm (7.1) mit $\text{rg}(A) = m$. Sei

$$\mathbf{x}_B = (x_1, \dots, x_m, 0, \dots, 0)^T$$

eine zulässige Basislösung. Die zugehörige Matrix der Basisvektoren sei $A_B \in \mathbb{R}^{m \times m}$ und die der Nichtbasisvektoren A_N . Dann sind die Zeilen der Matrix

$$(\bar{\mathbf{b}}, \bar{A}) := A_B^{-1}(\mathbf{b}, A) = A_B^{-1}(\mathbf{b}, A_B, A_N) \in \mathbb{R}^{m \times (n+1)}$$

lexikopositiv, da

$$A_B^{-1}(\mathbf{b}, A) = (\mathbf{x}_B, I_m, \bar{\mathbf{a}}_{m+1}, \dots, \bar{\mathbf{a}}_n),$$

$\mathbf{x}_B \geq \mathbf{0}$ und I_m die Einheitsmatrix des $\mathbb{R}^{m \times m}$ ist. Falls die Basisvariablen nicht die ersten m Variablen sind, dann ordnet man sie nach vorn.

Anstelle von (4.8) wird bei der lexikographischen Simplexmethode der Index l durch

$$\theta = \min_{>l; i=1, \dots, m, x_{ik} > 0} \frac{\mathbf{e}_i^T(\bar{\mathbf{b}}, \bar{A})}{x_{ik}} =: \frac{\mathbf{e}_l^T(\bar{\mathbf{b}}, \bar{A})}{x_{lk}}$$

bestimmt, wobei $\mathbf{e}_i \in \mathbb{R}^m$ der Einheitsvektor ist, der in der i -ten Komponente eine Eins hat. Das heißt, das Minimum wird bezüglich der lexikographischen Ordnung genommen. Das obige Symbol bedeutet, dass man sich wie üblich alle Einträge mit $x_{ik} > 0$ ansieht, die zugehörigen Vektoren $\mathbf{e}_i^T(\bar{\mathbf{b}}, \bar{A})$ bildet, durch x_{ik} dividiert und von den so erhaltenen Vektoren den lexikographisch kleinsten nimmt, und von diesem die erste von Null verschiedene Komponente, um l zu bestimmen. Es gilt, siehe beispielsweise [JS04]:

- Falls l in der allgemeinen Simplexmethode (4.8) eindeutig bestimmt ist, erhält man bei der lexikographischen Simplexmethode den gleichen Index.
- Die lexikographische Simplexmethode definiert einen eindeutigen Index l . Man kann zeigen, dass eine Nichteindeutigkeit im Widerspruch zu $\text{rg}(A) = m$ steht.

- Das Ergebnis eines lexikographischen Simplexschrittes ist wiederum eine lexikopositive Basis. Die Basislösung ist also insbesondere zulässig.
- Bei der neuen Basislösung ist entweder der Zielfunktionswert kleiner oder die Differenz der Koeffizienten der Zielfunktion der neuen und der alten Basis ist lexikopositiv. Im ersten Fall hat man die Ecke verlassen. Im zweiten Fall kann es bei weiteren lexikographischen Simplexschritten nicht passieren, dass die alte Basis noch einmal verwendet wird. Ein Basiszyklus ist ausgeschlossen.

Bei der lexikographischen Simplexmethode werden also ausgehend von einer lexikopositiven Startlösung weitere lexikopositive Lösungen erzeugt. Dieses Verfahren ist endlich. Es bricht entweder mit einer Lösung des Optimierungsproblems ab, oder es wird gefunden, dass die Zielfunktion nicht nach unten beschränkt ist. Die Anzahl der Schritte kann $n!$ nicht übersteigen. Diese Schranke ist allerdings für größere Werte von n für die Praxis bedeutungslos. \square

Kapitel 8

Zur Effizienz der Simplexmethode

Bemerkung 8.1 Fragestellung. Die Simplexmethode ist ein Verfahren zur Bestimmung der Lösung des linearen Programms

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

Für ihre praktische Anwendung muss untersucht werden, wie teuer die Berechnung des Optimums ist. Dazu unterscheidet man zwei Situationen:

- Analyse des schlimmsten Falls, der bei der Lösung eines linearen Programms mit der Simplexmethode auftreten kann, *worst case* Modell,
- Analyse des in der Praxis zu erwartenden normalen Falls, der bei der Lösung eines linearen Programms mit der Simplexmethode auftreten kann, *real world* Modell.

Die zweite Situation ist an sich interessanter. Das Problem besteht darin, ein real world Modell aufzustellen. Diese Frage ist bis heute teilweise ungeklärt. In der praktischen Anwendung der Simplexmethode hat man jedoch die Erfahrung gewonnen, dass sie im Normalfall hervorragend funktioniert. Wir werden hier nur das worst case Modell untersuchen. \square

8.1 Maße für die Effizienz

Das Grundanliegen besteht darin, den Aufwand zur Abarbeitung eines numerischen Verfahrens in Abhängigkeit vom Umfang der Eingangsdaten abzuschätzen.

Definition 8.2 compl(A,B). Die Komplexität $\text{compl}(A,B)$ eines Algorithmus A zur Lösung von Aufgaben B (eines Problemkreises P) ist die Anzahl der elementaren Operationen auf einem Computer oder die benötigte Rechenzeit in Abhängigkeit vom Umfang der Eingangsdaten. \square

Der Wunsch ist natürlich, einen effizienten Algorithmus für jedes praxisrelevante Optimierungsproblem zu konstruieren. Das geht aber im allgemeinen nicht, da schwierige und auch unlösbare Probleme existieren.

Definition 8.3 P(d). Bezeichne P ein Problem und d den Umfang seiner Eingangsdaten. Dann beschreibt $P(d)$ die Menge aller Aufgaben P mit gleichem Umfang d der Eingangsdaten. \square

Definition 8.4 Worst case Komplexität eines Algorithmus. Die worst case Komplexität eines Algorithmus A zur Lösung eines Problems P ist gegeben durch

$$\text{w-compl}(A, P) = \sup_{B \in P(d)} \text{compl}(A, B).$$

Man kann entsprechend eine *average case Komplexität* von A bezüglich P erklären

$$\text{a-compl}(A, P) = \text{Erwartungswert}_{B \in P(d)} \text{compl}(A, B).$$

□

Definition 8.5 Worst case Komplexität eines Problems. Sei A_P die Menge aller Algorithmen zur Lösung eines Problems P . Die worst case Komplexität von P ist erklärt durch

$$\text{w-compl}(P) = \min_{A \in A_P} \text{w-compl}(A, P).$$

□

Die Komplexität wird im allgemeinen als Funktion der Menge der Eingangsdaten in der Form $\mathcal{O}(f(d))$ angegeben.

Beispiel 8.6 Matrizenmultiplikation. Gegeben seien zwei Matrizen $A, B \in \mathbb{R}^{n \times n}$ und gesucht ist das Produkt $C = AB$. Ein Eintrag von C berechnet sich wie folgt

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj},$$

benötigt also n Multiplikationen und $(n - 1)$ Additionen, das heißt $\mathcal{O}(n)$ Operationen. Die Anzahl der zu berechnenden Einträge von C ist n^2 . Somit hat man insgesamt $\mathcal{O}(n^3)$ Operationen durchzuführen.

Die Frage ist, ob der Aufwand von $\mathcal{O}(n^3)$ optimal ist. Da man insgesamt n^2 Größen zu berechnen hat, kann der minimale Aufwand der Matrizenmultiplikation nicht kleiner als $\mathcal{O}(n^2)$ sein. Man kennt heute Verfahren, deren Aufwand für große n wie $\mathcal{O}(n^{2.38})$ ist. (SIAM News 38, Vol. 9, 2005) □

Definition 8.7 Gutartiges Problem. Probleme mit polynomialer Komplexität $\mathcal{O}(d^z)$, $z \in \mathbb{R}$, heißen gutartig, Probleme mit exponentieller Komplexität $\mathcal{O}(z^d)$, $z \in \mathbb{R}$, $z > 1$, bössartig. □

8.2 Zur worst case Komplexität der Simplexmethode

Bemerkung 8.8 Konstruktion eines Beispielproblems. In diesem Abschnitt wird ein Beispiel konstruiert, bei welchem der Aufwand der Simplexmethode exponentiell wächst, wobei der Aufwand in der Anzahl der Schritte gemessen wird. Das bedeutet, dass die Simplexmethode theoretisch ein sehr ineffizientes Verfahren sein kann. Dieser Fall ist in der Praxis glücklicherweise faktisch nicht zu beobachten.

Bei der worst case Komplexität wird der schlechteste Fall betrachtet. Für die Simplexmethode bedeutet das, dass die schlechteste Wahl der Hauptspalte bezüglich der Anzahl der zulässigen Basislösungen betrachtet wird, die man beim Transformationsprozess erzeugt.

Wie betrachten den n -dimensionalen Einheitswürfel $[0, 1]^n$. Dieser hat 2^n Ecken. Die Koordinaten des Einheitswürfels werden nun gestört

$$\begin{aligned} \varepsilon &\leq x_1 \leq 1 && \text{mit } 0 < \varepsilon < 1/2, \\ \varepsilon x_{j-1} &\leq x_j \leq 1 - \varepsilon x_{j-1} \leq 1 && j = 2, \dots, n. \end{aligned}$$

Über diesem gestörten Würfel wird folgendes lineares Programm definiert:

$$\begin{aligned}
 -x_n &\rightarrow \min ! \\
 x_1 - r_1 &= \varepsilon \\
 x_1 + s_1 &= 1 \\
 x_j - \varepsilon x_{j-1} - r_j &= 0 \quad j = 2, \dots, n, \\
 x_j + \varepsilon x_{j-1} + s_j &= 1 \quad j = 2, \dots, n, \\
 x_j, r_j, s_j &\geq 0 \quad j = 1, \dots, n.
 \end{aligned} \tag{8.1}$$

Ordnet man die Unbekannten gemäß $(x_1, \dots, x_n, r_1, \dots, r_n, s_1, \dots, s_n)$, dann hat die Matrix der Nebenbedingungen die Gestalt

$$A = \begin{pmatrix}
 1 & 0 & 0 & \dots & 0 & -1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\
 -\varepsilon & 1 & 0 & \dots & 0 & 0 & -1 & \dots & 0 & 0 & 0 & \dots & 0 \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 0 & 0 & 0 & \dots & 1 & 0 & 0 & \dots & -1 & 0 & 0 & \dots & 0 \\
 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\
 \varepsilon & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 0 & 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1
 \end{pmatrix} \in \mathbb{R}^{2n \times 3n}.$$

□

Lemma 8.9 Die Menge der zulässigen Basislösungen von (8.1) ist die Klasse der Untermengen von

$$(x_1, \dots, x_n, r_1, \dots, r_n, s_1, \dots, s_n),$$

bei denen alle $x_j > 0$, $j = 1, \dots, n$, und entweder $r_j > 0$ oder $s_j > 0$ für jedes $j = 1, \dots, n$. Alle Basislösungen sind nicht ausgeartet.

Beweis: Es wird erst die Zulässigkeit untersucht, dann die Basislösungseigenschaft.

Zuerst wird gezeigt, dass für zulässige Lösungen alle x_j , $j = 1, \dots, n$ positiv sein müssen. Ist $x_1 = 0$, dann folgt aus der ersten Nebenbedingung $r_1 = -\varepsilon < 0$. Ein Vektor mit $x_1 = 0$ kann also nicht zulässig sein. Der Beweis erfolgt nun durch Induktion. Seien $x_j > 0$ für $j = 1, \dots, k-1$ und $x_k = 0$. Dann folgt aus den Nebenbedingungen

$$0 = x_k = r_k + \underbrace{\varepsilon x_{k-1}}_{>0} \implies r_k < 0,$$

im Widerspruch zur letzten Nebenbedingung. Also sind alle x_j positiv, sie können damit keine Nichtbasisvariablen sein.

Nun werden die r_j, s_j betrachtet. Gelte für ein j , dass $r_j = s_j = 0$. Für $j = 1$ folgt dann aus den ersten beiden Nebenbedingungen von (8.1) $x_1 = \varepsilon = 1$. Das steht im Widerspruch zur Wahl von ε . Sei $j > 1$. Dann gelten

$$\begin{aligned}
 x_j - \varepsilon x_{j-1} &= 0, \\
 x_j + \varepsilon x_{j-1} &= 1.
 \end{aligned}$$

Subtraktion dieser Gleichungen ergibt

$$2\varepsilon x_{j-1} = 1.$$

Da jedoch $\varepsilon < 1/2$ und $x_{j-1} \leq 1$ ist die linke Seite echt kleiner als 1. Damit sind r_j oder s_j für jedes $j = 1, \dots, n$, positiv. Da man genau $2n$ Basisvariablen hat und bereits n davon durch die x_j gegeben sind, ist entweder r_j oder s_j für jedes $j = 1, \dots, n$, positiv.

Die Baseigenschaft der soeben konstruierten Menge sieht man durch Umordnung der Zeilen der Matrix $A_{2n, 2n}$. Für jeden Index j vertauscht man die Zeilen j und $j+n$ falls

$r_j > 0, s_j = 0$. Damit wird die Matrix auf Dreiecksform gebracht, wobei in der Diagonalen ± 1 stehen. Ihre Determinante ist somit ebenfalls ± 1 .

Die Basislösungen sind nicht ausgeartet, weil die $2n$ Basisvariablen alle positiv sind. ■

Bemerkung 8.10 Die Menge der Indizes j für die die Basislösungen $r_j > 0$ erfüllen wird mit S bezeichnet. Sind dies beispielsweise die Indizes $1, 3, 7$, so ist $S = \{1, 3, 7\}$. Die zugehörigen Basislösungen werden als $\mathbf{x}^{(S)}$ geschrieben, im Beispiel $\mathbf{x}^{\{1,3,7\}}$. Der Wert x_j in $\mathbf{x}^{(S)}$ wird mit $x_j^{(S)}$ bezeichnet. Wegen der Zielfunktion betrachten wir jetzt insbesondere den letzten Index n . □

Lemma 8.11 Seien $n \in S$ und $n \notin S'$. Dann ist $x_n^{(S)} > x_n^{(S')}$. Falls außerdem $S' = S \setminus \{n\}$ gilt, folgt $x_n^{(S')} = 1 - x_n^{(S)}$.

Beweis: Sei $n \in S$, das heißt $r_n > 0, s_n = 0$. Dann folgt aus

$$x_n^{(S)} + \varepsilon x_{n-1}^{(S)} + s_n = 1 \implies x_n^{(S)} = 1 - \varepsilon x_{n-1}^{(S)} > 1/2$$

wegen $x_{n-1}^{(S)} \leq 1, \varepsilon < 1/2$.

Andererseits gilt für $n \notin S'$, dass $r_n = 0$. Mit denselben Argumenten folgt

$$x_n^{(S')} - \varepsilon x_{n-1}^{(S')} - r_n = 0 \implies x_n^{(S')} = \varepsilon x_{n-1}^{(S')} < 1/2.$$

Die Mengen in der zweiten Aussage des Lemmas unterscheiden sich nur dadurch, dass in S gilt $r_n > 0, s_n = 0$ und in S' gilt $r_n = 0, s_n > 0$. Da alle anderen Indizes in S und S' gleich sind und die Nebenbedingungen für die Indizes kleiner n nicht von x_n, r_n, s_n abhängen, gilt insbesondere

$$x_{n-1}^{(S)} = x_{n-1}^{(S')}.$$

Da $r_n = 0$ für S' und $s_n = 0$ für S ist, folgt

$$x_n^{(S')} = \varepsilon x_{n-1}^{(S')} = 1 - (1 - \varepsilon x_{n-1}^{(S')}) = 1 - (1 - \varepsilon x_{n-1}^{(S)}) = 1 - x_n^{(S)}.$$

■

Lemma 8.12 Die Untermengen von $\{1, 2, \dots, n\}$ seien so geordnet, dass

$$x_n^{(S_1)} \leq x_n^{(S_2)} \leq \dots \leq x_n^{(S_{2^n})}$$

gilt. Diese Ungleichungen sind scharf, das heißt es gilt $<$, und die zulässigen Basislösungen $x^{(S_j)}$ und $x^{(S_{j+1})}$ sind benachbart für $j = 1, \dots, 2^n - 1$, das heißt, sie unterscheiden sich nur in einem Basisvektor.

Beweis: Der Beweis erfolgt durch Induktion über die Dimension n .

Induktionsanfang. $n = 1$. Man hat zwei Eckpunkte. Aus

$$x_1 - r_1 = \varepsilon, \quad x_1 + s_1 = 1$$

folgt

$$(x_1, r_1, s_1) = (\varepsilon, 0, 1 - \varepsilon) \vee (1, 1 - \varepsilon, 0).$$

Diese Punkte sind natürlich benachbart und die Schärfe der Ungleichung wurde bereits im letzten Lemma bewiesen ($x_n^{(S)} > x_n^{(S')}$).

Induktionsannahme. Für einen n -Würfel sei alles korrekt. Die entsprechende Numerierung sei S_1, \dots, S_{2^n} .

Induktionsschritt. Man betrachtet jetzt $\{1, 2, \dots, n+1\}$. Offensichtlich gelten $S_j \subset \{1, 2, \dots, n+1\}$ und $n+1 \notin S_j, j = 1, \dots, 2^n$. Damit ist $r_{n+1}^{(S_j)} = 0$. Aus der entsprechenden Nebenbedingung folgt

$$x_{n+1}^{(S_j)} = \varepsilon x_n^{(S_j)}.$$

Nach Induktionsannahme ist

$$x_n^{(S_1)} < x_n^{(S_2)} < \dots < x_n^{(S_{2^n})},$$

woraus nun folgt (Durchmultiplizieren mit ε)

$$x_{n+1}^{(S_1)} < x_{n+1}^{(S_2)} < \dots < x_{n+1}^{(S_{2^n})}. \quad (8.2)$$

Die Reihenfolge dieser Untermengen bleibt also erhalten.

Nun betrachten wir die Lösungen mit $r_{n+1}^{(S_j)} > 0$. Sei $S'_j = S_j \cup \{n+1\}$. Nach Lemma 8.11 ist

$$x_{n+1}^{(S_j)} = 1 - x_{n+1}^{(S'_j)} \implies x_{n+1}^{(S'_j)} = 1 - x_{n+1}^{(S_j)} \quad j = 1, \dots, 2^n, \quad (8.3)$$

und speziell

$$x_{n+1}^{(S'_{2^n})} > x_{n+1}^{(S_{2^n})}. \quad (8.4)$$

Aus (8.2), (8.3) und (8.4) resultiert

$$x_{n+1}^{(S_1)} < \dots < x_{n+1}^{(S_{2^n})} < x_{n+1}^{(S'_{2^n})} < \dots < x_{n+1}^{(S'_1)}.$$

Nun muss noch die Nachbarschaft der Basislösungen bewiesen werden. Nach Induktionsannahme, sind $\mathbf{x}^{(S_1)}, \dots, \mathbf{x}^{(S_{2^n})}$ in n Dimensionen benachbart. In der $(n+1)$ -sten Dimension, erhalten diese Basislösungen alle noch den Spaltenvektor von s_{n+1} . Sie bleiben damit benachbart. Die Basislösungen $\mathbf{x}^{(S_{2^n})}$ und $\mathbf{x}^{(S'_{2^n})}$ unterscheiden sich nur im Basisvektor von s_{n+1} beziehungsweise r_{n+1} . Sie sind also auch benachbart. Die Basislösungen $\mathbf{x}^{(S'_{2^n})}, \dots, \mathbf{x}^{(S'_1)}$ sind benachbart, weil $\mathbf{x}^{(S_1)}, \dots, \mathbf{x}^{(S_{2^n})}$ benachbart sind. ■

Satz 8.13 Für jedes $n \geq 1$ existiert ein lineares Programm, bestehend aus $2n$ Gleichungen und $3n$ Variablen, so dass die Simplexmethode $2^n - 1$ Iterationsschritte braucht, um das Optimum zu bestimmen.

Die Struktur dieses linearen Programms kann so eingerichtet werden, dass alle Koeffizienten (zum Beispiel) ≤ 4 sind.

Beweis: Der erste Teil des Satzes wird durch das angegebene Beispiel (8.1) bewiesen. In Lemma 8.12 sind die 2^n verschiedenen zulässigen Basislösungen streng geordnet. Die Simplexmethode wird so angewendet, dass mit jeder Transformation nur die jeweils geringste Verbesserung des Zielfunktionswertes erreicht wird. Bei $\mathbf{x}_n^{(1)}$ beginnend, sind somit $2^n - 1$ Transformationen nötig.

Für den zweiten Teil des Satzes wähle man $\varepsilon = 1/4$. ■

Beispiel 8.14 Wir betrachten das in diesem Abschnitt studierte Problem (8.1) für $n = 3$.

$$\begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -\varepsilon & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -\varepsilon & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \varepsilon & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & \varepsilon & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -x_3 \rightarrow \min ! \\ x_1 \\ x_2 \\ x_3 \\ r_1 \\ r_2 \\ r_3 \\ s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} \varepsilon \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

$$\mathbf{x}, \mathbf{r}, \mathbf{s} \geq \mathbf{0}.$$

Die Ecken des Würfels, des gestörten Würfels und die Zielfunktionswerte sind wie folgt

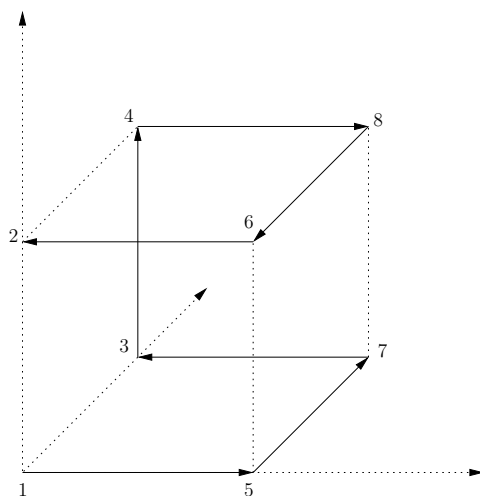
Nr.	Würfel	gest. Würfel	z	z für $\varepsilon = 1/4$	Reihenfolge
1	$(0, 0, 0)$	$(\varepsilon, \varepsilon^2, \varepsilon^3)$	$-\varepsilon^3$	-0.015625	8
2	$(0, 0, 1)$	$(\varepsilon, \varepsilon^2, 1 - \varepsilon^3)$	$-1 + \varepsilon^3$	-0.984375	1
3	$(0, 1, 0)$	$(\varepsilon, 1 - \varepsilon^2, \varepsilon - \varepsilon^3)$	$-\varepsilon + \varepsilon^3$	-0.234375	5
4	$(0, 1, 1)$	$(\varepsilon, 1 - \varepsilon^3, 1 - \varepsilon + \varepsilon^3)$	$-1 + \varepsilon - \varepsilon^3$	-0.765625	4
5	$(1, 0, 0)$	$(1, \varepsilon, \varepsilon^2)$	$-\varepsilon^2$	-0.0625	7
6	$(1, 0, 1)$	$(1, \varepsilon, 1 - \varepsilon^2)$	$-1 + \varepsilon^2$	-0.9375	2
7	$(1, 1, 0)$	$(1, 1 - \varepsilon, \varepsilon - \varepsilon^2)$	$-\varepsilon + \varepsilon^2$	-0.1875	6
8	$(1, 1, 1)$	$(1, 1 - \varepsilon, 1 - \varepsilon + \varepsilon^2)$	$-1 + \varepsilon - \varepsilon^2$	-0.8125	3

Man beginnt mit der Startlösung

$$\mathbf{x} = \begin{pmatrix} \varepsilon \\ \varepsilon^2 \\ \varepsilon^3 \end{pmatrix} \implies \mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{s} = \begin{pmatrix} 1 - \varepsilon \\ 1 - \varepsilon^2 \\ 1 - \varepsilon^3 \end{pmatrix}.$$

Das Transformationsprinzip der Simplexmethode wählt jeweils die kleinste Verbesserung der Zielfunktion. Die Eckpunkte des gestörten Würfels werden in folgender Reihenfolge durchgegangen:

$$1 \Rightarrow 5 \Rightarrow 7 \Rightarrow 3 \Rightarrow 4 \Rightarrow 8 \Rightarrow 6 \Rightarrow 2.$$



□

Die Zusammenfassung dieses Kapitels ist wie folgt.

Satz 8.15 Die Simplexmethode besitzt als worst case Komplexität mindestens $\mathcal{O} = (2^n)$.

Bemerkung 8.16 Diese tritt aber praktisch nicht auf. In der Praxis hat die Simplexmethode eine polynomiale Komplexität. Erfahrungsgemäß liegt die Anzahl der Iterationen etwa bei $3n$. Im Kapitel 12 werden wir noch Verfahren kennenlernen, bei denen man beweisen kann, dass sie von polynomialer Komplexität sind, die sogenannten Innere-Punkt-Methoden. □