

Kapitel 8

Zur Effizienz der Simplexmethode

Die Simplexmethode ist ein Verfahren zur Bestimmung der Lösung des linearen Programms

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

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 model 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.

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 8.1 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 8.2 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 8.3 Worst case Komplexität eines Algorithmus. Die worst case Komplexität eines Algorithmus A zur Lösung eines Problems P ist gegeben durch

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

□

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 8.4 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 8.5 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 8.6 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

In diesem Abschnitt wird ein Beispiel konstruiert, bei welchem der Aufwand der Simplexmethode exponentiell wächst. 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:

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

Die Matrix der Nebenbedingungen hat 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 8.7 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 Basiseigenschaft 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 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 8.8 *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

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

■

Lemma 8.9 *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 8.8 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 8.10 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 8.9 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 8.11 Wir betrachten das in diesem Abschnitt studierte Problem (8.1) für

$n = 3$.

$$\begin{array}{c}
 -x_3 \rightarrow \min ! \\
 \begin{pmatrix}
 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
 -\varepsilon & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
 0 & -\varepsilon & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
 \varepsilon & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
 0 & \varepsilon & 1 & 0 & 0 & 0 & 0 & 0 & 1
 \end{pmatrix}
 \begin{pmatrix}
 x_1 \\
 x_2 \\
 x_3 \\
 r_1 \\
 r_2 \\
 r_3 \\
 s_1 \\
 s_2 \\
 s_3
 \end{pmatrix}
 =
 \begin{pmatrix}
 \varepsilon \\
 0 \\
 0 \\
 1 \\
 1 \\
 1
 \end{pmatrix}, \\
 \mathbf{x}, \mathbf{r}, \mathbf{s} \geq \mathbf{0}.
 \end{array}$$

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

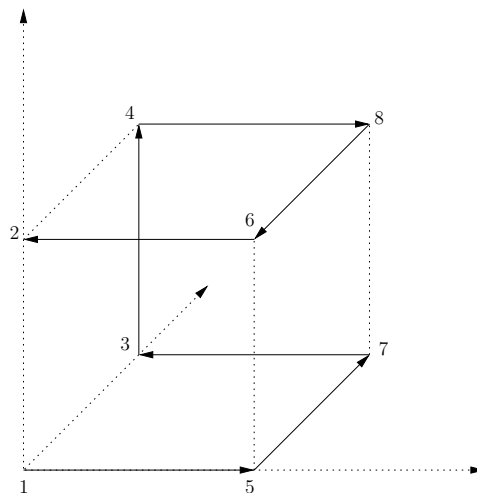
Nr.	Würfel	gest. Würfel	z	z für $\varepsilon = 1/4$	Reihenfolge
1	(0, 0, 0)	$(\varepsilon, \varepsilon^2, \varepsilon^3)$	$-\varepsilon^3$	-0.015625	8
2	(0, 0, 1)	$(\varepsilon, \varepsilon^2, 1 - \varepsilon^3)$	$-1 + \varepsilon^3$	-0.984375	1
3	(0, 1, 0)	$(\varepsilon, 1 - \varepsilon^2, \varepsilon - \varepsilon^3)$	$-\varepsilon + \varepsilon^3$	-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 8.12 *Die Simplexmethode besitzt als worst case Komplexität mindestens $\mathcal{O} = (2^n)$.*

Diese tritt aber praktisch nicht auf. In der Praxis hat die Simplexmethode eine polynomiale Komplexität.

Bemerkung 8.13 *Innere-Punkt-Methoden.* Seit 1984 hat sich eine weitere Klasse von Verfahren zur Lösung linearer Programme etabliert, sogenannte Innere-Punkt-Methoden. Diese arbeiten mit Techniken der nichtlinearen Optimierung (Newton-Verfahren). Es ist derzeit ungeklärt, welche Herangehensweise zur Lösung linearer Programme, Simplexmethode oder Innere-Punkt-Methoden, wirklich effizienter ist. Ein Vorteil von Innere-Punkt-Methoden liegt in ihren theoretischen Eigenschaften, da man zeigen kann dass ihre worst case Komplexität zur Berechnung einer Näherungslösung mit einer vorgegebenen Toleranz sich polynomial verhält. Einige Innere-Punkt-Methoden sind:

- Ellipsoid-Methode ($\mathcal{O}(n^6)$),
- Algorithmus von Karmarkar (1984) ($\mathcal{O}(n^{7/2})$),
- Verbesserung des Algorithmus von Karmarkar (Qi/Shi) ($\mathcal{O}(n^3)$),
- Kurz-Schritt-Algorithmus ($\sim \mathcal{O}(n^{1/2})$), siehe [JS04].

□

Kapitel 9

Dualitätssätze der linearen Optimierung

Sei

$$\begin{aligned} z = \mathbf{c}^T \mathbf{x} &\rightarrow \min ! \\ \mathbf{A} \mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0} \end{aligned} \tag{9.1}$$

mit $\mathbf{c}, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ ein lineares Programm.

Definition 9.1 *Duales lineares Programm.* Das lineare Programm

$$\begin{aligned} \tilde{z} = \mathbf{b}^T \mathbf{y} &\rightarrow \max ! \\ \mathbf{A}^T \mathbf{y} &\leq \mathbf{c} \end{aligned} \tag{9.2}$$

mit $\mathbf{y} \in \mathbb{R}^m$ wird das zu (9.1) duale lineare Programm genannt. Man nennt (9.1) primal. \square

Die Ziele dieses Abschnitts bestehen darin, die Existenz zulässiger Lösungen des dualen linearen Programms, die Relationen der zulässigen Lösungen des primalen und dualen linearen Programms, die Relationen zwischen den Optimallösungen und die Verbesserung der numerischen Verfahren zu untersuchen. Ein duales Analogon zur Simplexmethode soll entwickelt werden.

Satz 9.2 *Ist \mathbf{x} eine zulässige Lösung von (9.1) und ist \mathbf{y} eine zulässige Lösung von (9.2), dann gilt $z(\mathbf{x}) \geq \tilde{z}(\mathbf{y})$.*

Beweis: Es gilt $\mathbf{x} \geq \mathbf{0}$ und $\mathbf{c}^T \geq \mathbf{y}^T A$. Damit folgt

$$z(\mathbf{x}) = \mathbf{c}^T \mathbf{x} \geq \mathbf{y}^T A \mathbf{x} = \mathbf{y}^T \mathbf{b} = \tilde{z}(\mathbf{y}).$$

■

Man nennt diesen Satz auch schwachen Dualitätssatz.

Folgerung 9.3 *Ist \mathbf{x}_0 eine zulässige Lösung von (9.1) und \mathbf{y}_0 eine zulässige Lösung von (9.2) und gilt $z(\mathbf{x}_0) = \tilde{z}(\mathbf{y}_0)$, dann ist \mathbf{x}_0 eine Optimallösung von (9.1) und \mathbf{y}_0 eine Optimallösung von (9.2).*

Satz 9.4 Starker Dualitätssatz. *Das primale Problem (9.1) besitzt genau dann eine endliche Optimallösung, wenn das duale Problem (9.2) eine endliche Optimallösung besitzt. In diesem Fall gilt $z_{\min} = \tilde{z}_{\max}$.*

Beweis: 1.) Es existiere ein endliches Minimum des primalen Problems (9.1) und dieses Minimum werde von $\mathbf{x}_0 = (x_1^{(0)}, \dots, x_m^{(0)}, 0, \dots, 0)^T$ angenommen. Dann sind erklärt:

- Die zugehörigen Basisvektoren seien $\mathbf{a}_1, \dots, \mathbf{a}_m$.
- Mit $X = (x_{ij})_{i=1, \dots, m, j=1, \dots, n}$ werden die Darstellungskoeffizienten für alle Spalten von A bezüglich dieser Basisvektoren bezeichnet.
- $\mathbf{z} = (z_1, \dots, z_n)^T$ sei der Vektor, der durch $z_j = \sum_{i=1}^m c_i x_{ij}$, $j = 1, \dots, n$, erzeugt wird.
- $A_0 = (\mathbf{a}_1, \dots, \mathbf{a}_m)$,
- $\mathbf{c}_0 = (c_1, \dots, c_m)^T$.

Wegen des Optimalitätskriteriums der Simplexmethode gilt für \mathbf{x}_0

$$\mathbf{z} \leq \mathbf{c} \quad (z_k - c_k \leq 0, \quad k = 1, \dots, n). \quad (9.3)$$

Aus der Nebenbedingung und der Definition der Darstellungskoeffizienten folgt

$$A_0 \mathbf{x}_0 = \mathbf{b}, \quad A_0 X = A.$$

Daraus ergibt sich

$$\mathbf{x}_0 = A_0^{-1} \mathbf{b}, \quad X = A_0^{-1} A. \quad (9.4)$$

Weiter erhalten wir aus der Definition von \mathbf{z}

$$\mathbf{c}_0^T X = \mathbf{z}^T \leq \mathbf{c}^T. \quad (9.5)$$

Jetzt setzen wir $\mathbf{y}_0 = A_0^{-T} \mathbf{c}_0$ und zeigen, dass \mathbf{y}_0 eine Optimallösung des dualen Problems (9.2) ist. Die Zulässigkeit von \mathbf{y}_0 folgt aus (9.4) und (9.5)

$$\mathbf{y}_0^T A = \mathbf{c}_0^T A_0^{-1} A = \mathbf{c}_0^T X \leq \mathbf{c}^T.$$

Die Lösung \mathbf{y}_0 ist optimal wegen

$$\tilde{z}(\mathbf{y}_0) = \mathbf{b}^T \mathbf{y}_0 = \mathbf{y}_0^T \mathbf{b} = \mathbf{c}_0^T A_0^{-1} \mathbf{b} = \mathbf{c}_0^T \mathbf{x}_0 = z(\mathbf{x}_0),$$

wobei (9.4) verwendet wurde. Damit liefert \mathbf{y}_0 einen Zielfunktionswert, der mit dem von \mathbf{x}_0 übereinstimmt. Da \mathbf{x}_0 Optimum des primalen Problems ist, ist \mathbf{y}_0 wegen Folgerung 9.3 Optimum des dualen Problems.

2) Das Ziel besteht darin, diesen Teil des Beweises auf den ersten Teil zurückzuführen, indem gezeigt wird, dass das duale Problem des dualen Problems (9.2) gerade das primale Problem (9.1) ist. Dazu wird das duale Problem so umgeformt, dass es die Gestalt eines primalen Problems annimmt. Bildet man dann aus dieser Form das duale Problem, erhält man die Behauptung.

Es existiere ein endliches Maximum \tilde{z}_{\max} des dualen Problems (9.2). Wir setzen für einen beliebigen Vektor $\mathbf{y} \in \mathbb{R}^m$, der die Nebenbedingungen von (9.2) erfüllt, $\mathbf{y} = \mathbf{y}_1 - \mathbf{y}_2$, $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^m$, $\mathbf{y}_1, \mathbf{y}_2 \geq \mathbf{0}$. Aus den Nebenbedingungen von (9.2) erhält man

$$A^T (\mathbf{y}_1 - \mathbf{y}_2) + \mathbf{y}_3 = \mathbf{c} \quad \in \mathbb{R}^n,$$

mit den Schlupfvariablenvektor $\mathbf{y}_3 \in \mathbb{R}^n$, $\mathbf{y}_3 \geq \mathbf{0}$. Daraus bilden wir folgendes zu (9.2) äquivalentes Problem, wobei das Vorzeichen der Zielfunktion geändert wird

$$\begin{aligned} \mathbf{b}^T (\mathbf{y}_2 - \mathbf{y}_1) &\rightarrow \min ! \\ -A^T \mathbf{y}_1 + A^T \mathbf{y}_2 - \mathbf{y}_3 &= -\mathbf{c} \\ \mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3 &\geq \mathbf{0}. \end{aligned}$$

Setzt man

$$\begin{aligned} \mathbf{w} &= \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \end{pmatrix} \in \mathbb{R}^{2m+n}, \quad \mathbf{d} = \begin{pmatrix} -\mathbf{b} \\ \mathbf{b} \\ \mathbf{0}_n \end{pmatrix} \in \mathbb{R}^{2m+n}, \\ \mathcal{A} &= (-A^T, A^T, -I_n) \in \mathbb{R}^{n \times (2m+n)}, \end{aligned}$$

so kann man dieses System in der Form

$$\begin{aligned} \mathbf{d}^T \mathbf{w} &\rightarrow \min ! \\ \mathcal{A} \mathbf{w} &= -\mathbf{c} \\ \mathbf{w} &\geq \mathbf{0} \end{aligned} \tag{9.6}$$

schreiben.

Einerseits ist dieses Problem zum dualen Problem (9.2) äquivalent. Damit besitzt (9.6) nach Voraussetzung ein endliches Optimum $z_{\min}^{(\mathbf{w})}$, für welches gilt $z_{\min}^{(\mathbf{w})} = \tilde{z}_{\max}$.

Andererseits besitzt das Problem (9.6) die gleiche Gestalt wie das primale Problem (9.1). Die duale Aufgabe zu (9.6) hat nun folgende Gestalt: Gesucht ist $\mathbf{x} \in \mathbb{R}^n$ mit

$$-\mathbf{c}^T \mathbf{x} \rightarrow \max !, \quad \mathcal{A}^T \mathbf{x} \leq \mathbf{d},$$

das heißt

$$\begin{aligned} \mathbf{c}^T \mathbf{x} &\rightarrow \min ! \\ -A \mathbf{x} &\leq -\mathbf{b} \\ A \mathbf{x} &\leq \mathbf{b} \\ -\mathbf{x} &\leq \mathbf{0}. \end{aligned}$$

Aus den ersten beiden Nebenbedingungen folgt $A \mathbf{x} = \mathbf{b}$. Damit ist gezeigt, dass (9.1) das duale Problem zu (9.2) ist. Aus dem ersten Teil des Beweises wissen wir, dass die duale Aufgabe zu (9.6) ein endliches Maximum besitzt und dass dieses Maximum mit $z_{\min}^{(\mathbf{w})} = \tilde{z}_{\max}$ übereinstimmt. ■

Jetzt wird der Fall betrachtet, dass die Zielfunktion der primalen Aufgabe nach unten nicht beschränkt ist.

Satz 9.5 *Ist die Zielfunktion $z = \mathbf{c}^T \mathbf{x}$ der primalen Aufgabe (9.1) auf der Menge der zulässigen Lösungen nach unten unbeschränkt, dann besitzt die zugehörige duale Aufgabe (9.2) keine zulässige Lösung. Analog gilt, dass im Falle dass die Zielfunktion der dualen Aufgabe auf der Menge der zulässigen Lösungen nicht nach oben beschränkt ist, die primale Aufgabe keine zulässige Lösung besitzt.*

Beweis: Indirekter Beweis. Sei die Zielfunktion des primalen Problems nicht nach unten beschränkt und sei \mathbf{y} eine zulässige Lösung des dualen Problems, das heißt es gilt $A^T \mathbf{y} \leq \mathbf{c}^T$. Aus Satz 9.2 folgt dann aber $z(\mathbf{x}) \geq \tilde{z}(\mathbf{y})$ und die Zielfunktion wäre nach unten beschränkt.

Die zweite Aussage folgt aus der ersten Aussage und daraus, dass das primale Problem (9.1) das duale Problem des dualen Problems (9.2) ist. ■

Folgerung 9.6 *Eine zulässige Lösung \mathbf{x}_0 des primalen Problems (9.1) ist genau dann optimal, wenn eine zulässige Lösung \mathbf{y}_0 des dualen Problems (9.2) existiert, mit $\mathbf{c}^T \mathbf{x}_0 = \mathbf{b}^T \mathbf{y}_0$. Eine analoge Aussage gilt, wenn man vom dualen Problem ausgeht.*

Satz 9.7 Komplementaritätssatz. Es sei $\mathbf{x}_0 = (x_1^{(0)}, \dots, x_m^{(0)}, 0, \dots, 0)^T$ eine zulässige Basislösung des primalen Problems (9.1). Dann ist \mathbf{x}_0 genau dann optimal, wenn es eine zulässige Lösung \mathbf{y} des dualen Problems (9.2) mit folgenden Eigenschaften gibt:

- 1) für alle Indizes $i \in \{1, \dots, m\}$ mit $x_i^{(0)} > 0$ gilt $\mathbf{a}_i^T \mathbf{y} = c_i$,
- 2) für alle Indizes $j \in \{1, \dots, n\}$ mit $\mathbf{a}_j^T \mathbf{y} < c_j$ gilt $x_j^{(0)} = 0$.

Beweis: i) Sei \mathbf{x}_0 optimal. Nach Folgerung 9.6 gibt es dann ein \mathbf{y}_0 mit $\mathbf{c}^T \mathbf{x}_0 = \mathbf{b}^T \mathbf{y}_0$. Einsetzen von $A\mathbf{x}_0 = \mathbf{b}$ ergibt

$$\mathbf{c}^T \mathbf{x}_0 = \underbrace{\mathbf{x}_0^T A^T \mathbf{y}_0}_{\in \mathbb{R}} = \mathbf{y}_0^T A \mathbf{x}_0 \iff (\mathbf{c}^T - \mathbf{y}_0^T A) \mathbf{x}_0 = 0.$$

Komponentenweise lautet diese Gleichung

$$\sum_{j=1}^n (c_j - \mathbf{y}_0^T \mathbf{a}_j) x_j^{(0)} = 0.$$

Da \mathbf{y}_0 eine zulässige Lösung des dualen Problems ist, sind alle Faktoren nichtnegativ. Damit die Summe Null wird, müssen alle Summanden verschwinden und wenigstens jeweils einer der Faktoren Null sein. Ist $x_j^{(0)} > 0$, muss $\mathbf{a}_j^T \mathbf{y}_0 = c_j$ sein. Ist $\mathbf{a}_j^T \mathbf{y} < c_j$, so muss $x_j^{(0)} = 0$ sein. Die Optimallösung \mathbf{y}_0 des dualen Problems erfüllt also die Bedingungen 1) und 2).

ii) Es gibt einen Vektor \mathbf{y} der die Bedingungen 1) und 2) erfüllt. Wir nehmen an, \mathbf{x}_0 sei nicht optimal. Dann gilt $\mathbf{c}^T \mathbf{x}_0 > \mathbf{b}^T \mathbf{y}$. Analog zum ersten Teil erhält man

$$\sum_{j=1}^n (c_j - \mathbf{y}^T \mathbf{a}_j) x_j^{(0)} > 0.$$

Aus den Bedingungen 1), 2) folgt jedoch, dass die Summe verschwindet. Damit ist die Annahme falsch und \mathbf{x}_0 ist optimal. ■

Bemerkung 9.8 Mit Hilfe der Dualität ist die Möglichkeit der Bestimmung von Schranken für eine zulässige (optimale) Lösung gegeben. Es gilt

$$z(\mathbf{x}) \geq z(\mathbf{x}_0) = \tilde{z}(\mathbf{y}_0) \geq \tilde{z}(\mathbf{y}),$$

wobei \mathbf{x} eine zulässige Lösung des primalen Problems (9.1), \mathbf{x}_0 eine Optimallösung von (9.1), \mathbf{y} eine zulässige Lösung des dualen Problems (9.2) und \mathbf{y}_0 eine Optimallösung des dualen Problems ist. □

Ein Spezialfall des dualen linearen Programms ist das symmetrische duale lineare Programm. Gegeben sei das lineare Programm

$$\begin{aligned} z = \mathbf{c}^T \mathbf{x} &\rightarrow \min ! \\ A\mathbf{x} &\geq \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0}. \end{aligned} \tag{9.7}$$

Aus (9.7) wird

$$\begin{aligned} \tilde{z} = \mathbf{b}^T \mathbf{y} &\rightarrow \max ! \\ A^T \mathbf{y} &\leq \mathbf{c} \\ \mathbf{y} &\geq \mathbf{0}. \end{aligned} \tag{9.8}$$

konstruiert. Man hat eine Nichtnegativitätsbedingung an die zulässigen Lösungen des dualen Programms.

Satz 9.9 Die linearen Programm (9.7) und (9.8) sind duale lineare Programme im Sinne von Definition 9.1.

Beweis: Aus (9.7) konstruieren wir das lineare Programm in Normalform

$$z = \mathbf{c}^T \mathbf{x} \rightarrow \min !, \quad \mathbf{A}\mathbf{x} - \mathbf{v} = \mathbf{b}, \quad \mathbf{x}, \mathbf{v} \geq \mathbf{0}.$$

Aus Definition 9.1 ergibt sich das folgende duale lineare Programm

$$\tilde{z} = \mathbf{b}^T \mathbf{y} \rightarrow \max !, \quad \mathbf{A}^T \mathbf{y} \leq \mathbf{c}, \quad -\mathbf{I}_m \mathbf{y} \leq \mathbf{0}.$$

Die letzte Bedingung ist äquivalent zu $\mathbf{y} \geq \mathbf{0}$. ■

Satz 9.10 Komplementaritätssatz. Sind \mathbf{x}_0 eine zulässige Lösung von (9.7) und \mathbf{y}_0 eine zulässige Lösung von (9.8), so sind sie genau dann optimal, wenn die folgenden Relationen erfüllt sind:

$$\mathbf{y}_0^T (\mathbf{A}\mathbf{x}_0 - \mathbf{b}) = 0, \tag{9.9}$$

$$(\mathbf{y}_0^T \mathbf{A} - \mathbf{c}^T) \mathbf{x}_0 = 0. \tag{9.10}$$

Beweis: 1) Seien \mathbf{x}_0 und \mathbf{y}_0 optimal. Dann folgt aus der Nebenbedingung von (9.8), aus der Nichtnegativität von \mathbf{x}_0 und aus Folgerung 9.6

$$\mathbf{y}_0^T (\mathbf{A}\mathbf{x}_0 - \mathbf{b}) = \mathbf{y}_0^T \mathbf{A}\mathbf{x}_0 - \mathbf{y}_0^T \mathbf{b} \leq \mathbf{c}^T \mathbf{x}_0 - \mathbf{y}_0^T \mathbf{b} = 0.$$

Andererseits gilt mit $\mathbf{y}_0 \geq \mathbf{0}$ und der Nebenbedingung von (9.7)

$$\mathbf{y}_0^T (\mathbf{A}\mathbf{x}_0 - \mathbf{b}) \geq 0.$$

Aus beiden Ungleichungen zusammen folgt (9.9). Die Beziehung (9.10) beweist man analog.

2) Gelten jetzt (9.9) und (9.10). Daraus folgt

$$\mathbf{c}^T \mathbf{x}_0 = \mathbf{y}_0^T \mathbf{A}\mathbf{x}_0 = \mathbf{y}_0^T \mathbf{b}.$$

Nach Folgerung 9.6 sind \mathbf{x}_0 und \mathbf{y}_0 optimal. ■

Beispiel für Anwendungen des Dualitätsprinzips werden in den Übungsaufgaben behandelt.