

Kapitel 8

Zur Effizienz der Simplexmethode

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

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

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 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

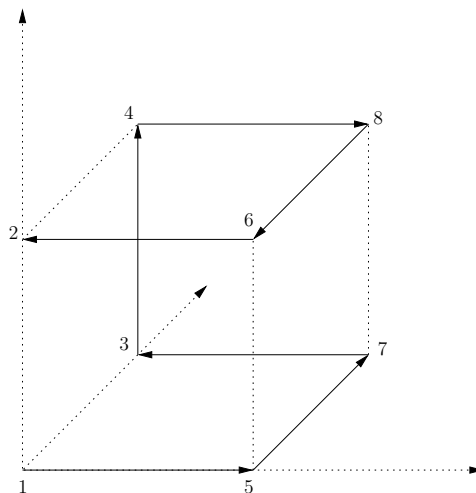
Nr.	Würfel	gest. Würfel	z	z für $\varepsilon = 1/4$	Reihenfolge
1	(0, 0, 0)	$(\varepsilon, \varepsilon^2, \varepsilon^3)$	$-\varepsilon^3$	-0.015625	8
2	(0, 0, 1)	$(\varepsilon, \varepsilon^2, 1 - \varepsilon^3)$	$-1 + \varepsilon^3$	-0.984375	1
3	(0, 1, 0)	$(\varepsilon, 1 - \varepsilon^2, \varepsilon - \varepsilon^3)$	$-\varepsilon + \varepsilon^3$	-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. 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.