Kapitel 4

Hauptsatz und Optimalitätskriterium der Simplexmethode

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

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

eine zulässige Basislösung mit $\mathbf{x} > \mathbf{0}$.

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

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

□

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

Das geometrische Prinzip der Simplexmethode ist wie folgt:

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

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

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

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

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

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

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

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

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

$$z_j = c_1x_{1j} + c_2x_{2j} + \ldots + c_mx_{mj}, \quad j = m+1,\ldots,n, \tag{4.4}$$

eingeführt.

Satz 4.2 ***Hauptsatz der Simplexmethode.*** *Sei z_0 der Wert der Zielfunktion für die zulässige Basislösung $\mathbf{x} = (x_1,\ldots,x_m,0,\ldots,0)^T$, $x_i > 0$, $i = 1,\ldots,m$. Gilt für ein k mit $m+1 \le k \le n$, dass $z_k - c_k > 0$, so existiert wenigstens eine zulässige Basislösung mit einem Zielfunktionswert kleiner als z_0.*

Beweis: Sei $\theta > 0$ vorerst beliebig gewählt. Man multipliziere (4.3) und (4.4) für $j = k$ mit θ und bilde (4.1) - θ (4.3) und (4.2) - θ (4.4):

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

$$\begin{aligned} &c_1(x_1 - \theta x_{1k}) + c_2(x_2 - \theta x_{2k}) + \ldots + c_m(x_m - \theta x_{mk}) + \theta c_k \\ &\quad = \quad z_0 - \theta z_k + \theta c_k = z_0 + \theta(c_k - z_k). \end{aligned} \tag{4.6}$$

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

$$(x_1 - \theta x_{1k},\ldots,x_m - \theta x_{mk},0,\ldots,\theta,\ldots,0)^T. \tag{4.7}$$

Es wird in den Lemmata 4.3 und 4.6 gezeigt, dass man mit diesem Vektor eine Basislösung konstruieren kann. Man hätte eine zulässige Basislösung, wenn $x_i - \theta x_{ik} \ge 0$, $i = 1,\ldots,m$. Der zugehörige Zielfunktionswert ist durch die Gleichung (4.6) gegeben. Er ist kleiner als z_0 falls $\theta > 0$ und $z_k - c_k > 0$. ■

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

Lemma 4.3 *Sei*

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

Dann ist die Lösung (4.7) zulässig.

Beweis: Man hat

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

und die neue Lösung

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

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

Bemerkung 4.4 Damit die Wahl (4.8) von θ überhaupt funktioniert, brauchen wir ein $x_{ik} > 0$. Falls es kein solches x_{ik} gibt, dann folgt, dass die Zielfunktion nach unten unbeschränkt ist. Man kann nämlich in diesem Fall θ beliebig groß wählen, da stets $x_i - \theta x_{ik} \geq 0$. Aus (4.6) folgt dann, dass unter der Bedingung $c_k - z_k < 0$ die Zielfunktion unbeschränkt nach unten ist. Fazit: Falls für ein $z_k - c_k > 0$ alle $x_{ik} \leq 0$, dann ist die Zielfunktion nicht von unten beschränkt und man breche die Simplexmethode ab. □

Bemerkung 4.5 Wenn man Entartung ausschließt, dann ist θ eindeutig bestimmt, das heißt, das Minimum in (4.8) wird für genau einen Index l angenommen. Es gilt auch die Umkehrung, dass falls der Index l in (4.8) nicht eindeutig bestimmt ist, dann hat man Ausartung. Ausartung kann zur Folge haben, dass gilt $z(\mathbf{x}_i) = z(\mathbf{x}_{i+1}) = \ldots$. Das nennt man einen Basiszyklus. □

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

Lemma 4.6 *Sei* $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ *ein System linear unabhängiger Vektoren und sei*

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

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

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

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

In diese Gleichung wird (4.10) eingesetzt. Es folgt

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

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

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

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

Satz 4.7 ***Optimalitätskriterium.*** *Eine zulässige Basislösung* $\mathbf{x} \in \mathbb{R}^n$ *mit* $z_0 = \sum_{i=1}^{m} c_i x_i$ *ist optimale Basislösung, wenn für alle* $j = m+1, \ldots, n$ *gilt* $z_j - c_j \leq 0$.

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

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

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

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

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

wobei

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

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

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

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

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

Durch Umordnung nach den Basisvektoren folgt

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

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

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

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

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

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

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

Setzt man dies in (4.16), so erhält man

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

■

An dieser Stelle sollen die Ergebnisse und Beobachtungen dieses Abschnitts zusammengefasst werden:

- $\min_{\mathbf{x}\in\mathbb{R}^n}\{\mathbf{c}^T\mathbf{x} \;:\; A\mathbf{x}=\mathbf{b},\ \mathbf{x}\geq\mathbf{0}\}$ ist zu lösen.
- Man braucht eine erste Basis $A_B=(\mathbf{a}_1,\ldots,\mathbf{a}_m)$ mit der Basislösung $\mathbf{x}_B=(x_1,\ldots,x_m)^T$. Damit hat man einen Nichtbasisanteil $A_N=(\mathbf{a}_{m+1},\ldots,\mathbf{a}_n)$ und $\mathbf{x}_N=(x_{m+1},\ldots,x_n)^T$. Auch der Kostenvektor wird in dieser Form zerlegt $\mathbf{c}^T=\left(\mathbf{c}_B^T,\mathbf{c}_N^T\right)=(c_1,\ldots,c_m,c_{m+1},\ldots,c_n)^T$.
- Aus $A\mathbf{x}=\mathbf{b}$ folgt

  $$A_B\mathbf{x}_B+A_N\mathbf{x}_N=\mathbf{b} \implies \mathbf{x}_B=A_B^{-1}\mathbf{b}-A_B^{-1}A_N\mathbf{x}_N.$$

  Dieser Ausdruck wird in die zu minimierende Zielfunktion eingesetzt

  $$\mathbf{c}^T\mathbf{x}=\mathbf{c}_B^T\left(A_B^{-1}\mathbf{b}-A_B^{-1}A_N\mathbf{x}_N\right)+\mathbf{c}_N^T\mathbf{x}_N\to\min,$$

  oder

  $$\mathbf{c}^T\mathbf{x}=\mathbf{c}_B^TA_B^{-1}\mathbf{b}-\left(\mathbf{c}_B^TA_B^{-1}A_N-\mathbf{c}_N^T\right)\mathbf{x}_N\to\min.$$

  Für die gegenwärtige Basislösung ist der erste Term eine Konstante. Betrachte den zweiten Term:
  - 1. Fall: $\boldsymbol{\beta}=\mathbf{c}_B^TA_B^{-1}A_N-\mathbf{c}_N^T<\mathbf{0}$. Dann folgt

    $$\mathbf{c}^T\mathbf{x}=\mathbf{c}_B^TA_B^{-1}\mathbf{b}-\left(\underbrace{\beta_{m+1}}_{<0}x_{m+1}+\ldots+\underbrace{\beta_n}_{<0}x_n\right)\to\min.$$

    Das heißt, der Zielfunktionswert kann mit $x_i>0$ für $i=m+1,\ldots,n$ nicht verkleinert werden. Damit hat man Optimalität erreicht.
  - 2. Fall: $\mathbf{c}_B^TA_B^{-1}A_N-\mathbf{c}_N^T\leq\mathbf{0}$ und $\beta_j=0$ für mindestens einen Index j. Dann ist das Optimum nicht eindeutig bestimmt.
  - 3. Fall: $\mathbf{c}_B^TA_B^{-1}A_N-\mathbf{c}_N^T\not<\mathbf{0}$. Dann bewirkt die Aufnahme einer Nichtbasisvariablen in die Basis eine Verkleinerung der Zielfunktion. Nun ist noch die Frage zu klären, welche Nichtbasisvariable man in die Basis aufnehmen soll, falls es mehrere Indizes j mit $z_j-c_j>0$ gibt. In diesem Falle wähle man beispielsweise

    $$z_k-c_k=\max_{j\in\mathrm{NBV},z_j-c_j>0} z_j-c_j.$$

    Der Index der Basisvariablen, die aus der bisherigen Basis entfernt werden soll, ist durch (4.8) gegeben.