

Lineare Optimierung mit dem Simplexverfahren

Fritz Philipp

Inhaltsverzeichnis

1	Problemstellung	3
2	Basisvektoren	4
3	Das Simplexverfahren	6
3.1	Allgemeine Vorbereitungen	6
3.2	Idee des Simplexverfahrens	8
3.3	Formulierung des Simplexverfahrens	9
3.4	Wohldefiniiertheit des Verfahrens	10
3.5	Ablauf des Simplexverfahrens	11
3.6	Bestimmung eines Basisvektors	14
3.7	Implementierung des Simplexverfahrens	15

1 Problemstellung

Ein lineares Optimierungsproblem besteht darin, zu gegebenem $\mathbf{c} \in \mathbb{R}^n$ ein $\mathbf{x} \in \mathbb{R}^n$ zu finden, so daß $\mathbf{c}^T \mathbf{x}$ unter bestimmten Nebenbedingungen minimal oder maximal wird. Da die Komponenten von \mathbf{x} oftmals Variablen für Anzahl, Währung oder Ähnliches darstellen, wird meistens $\mathbf{x} \geq \mathbf{0}$ vorausgesetzt. Das bedeutet, daß alle Komponenten von \mathbf{x} größer oder gleich Null sind. Wir wollen hier ein kleines Beispiel angeben, um den Leser mit der Gestalt einer solchen Aufgabe vertraut zu machen.

Beispiel: (*Lineare Optimierungsaufgabe und deren mathematische Formulierung*)

Aufgabe: Ein Bauer hat 50 ha Land zu bebauen, wahlweise mit Weizen, Rüben oder Mais, wobei die Anbaufläche für Rüben 20 ha nicht überschreiten darf. Es stehen Arbeitskräfte für insgesamt maximal 1300 Stunden zur Verfügung, wobei 20 Stunden je Hektar Weizen, 40 Stunden je Hektar Rüben und 30 Stunden je Hektar Mais benötigt werden. Der erzielbare Gewinn beträgt je Hektar Weizen 5000 DM, je Hektar Rüben 8000 DM und je Hektar Mais 6000 DM. Wie erreicht der Bauer den größten Gewinn?

Mathem. Formulierung: Es sei x_1 die Anbaufläche für Weizen, x_2 die für Rüben und x_3 die für Mais. Der Gewinn $\mathbf{c}^T \mathbf{x}$ ergibt sich dann mit $\mathbf{c} = (5000, 8000, 6000)^T$. Die Nebenbedingungen sind $x_1 + x_2 + x_3 \leq 50$, $x_2 \leq 20$ und $20x_1 + 40x_2 + 30x_3 \leq 1300$. Wir haben also das lineare Optimierungsproblem

$$\begin{aligned} \mathbf{c}^T \mathbf{x} &= \max \\ x_1 + x_2 + x_3 &\leq 50 \\ x_2 &\leq 20 \\ 20x_1 + 40x_2 + 30x_3 &\leq 1300 \\ \mathbf{x} &\geq \mathbf{0} \end{aligned} \quad (*)$$

zu lösen.

Bei der Form einer Optimierungsaufgabe wie im Beispiel stören uns natürlich die Ungleichungen und die Unübersichtlichkeit. Diese lassen sich aber durch folgende äquivalente Umformungen des Problems ausmerzen:

1. Ein Maximumsproblem kann in ein Minimierungsproblem übergeführt werden (und umgekehrt), indem \mathbf{c} in $-\mathbf{c}$ übergeht:

$$\mathbf{c}^T \mathbf{x} = \max \iff -\mathbf{c}^T \mathbf{x} = \min .$$

2. Eine Ungleichung $a_1x_1 + \dots + a_nx_n \leq b$ kann durch Einführung einer sogenannten *Schlupfvariablen* $y \geq 0$ zu einer Gleichung $a_1x_1 + \dots + a_nx_n + y = b$ übergehen.

3. Gibt es für eine Komponente x_j von \mathbf{x} keine Vorzeichenbedingung, so kann man $x_j = y - z$ mit $y, z \geq 0$ schreiben.

Mit solchen Umformungen können wir nämlich das Optimierungsproblem in die übersichtliche und einheitliche Gestalt der folgenden *Normalform* bringen, die mit $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, und $\mathbf{A} \in \mathbb{R}^{m \times n}$ folgendermaßen aussieht:

$$\begin{aligned} \mathbf{c}^T \mathbf{x} &= \min \\ \mathbf{A} \mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0} \end{aligned} \quad (1)$$

Beispiel: (*Fortsetzung des obigen Beispiels*) Die Optimierungsaufgabe war durch (*) gegeben. Wir führen nun die Schlupfvariablen x_4 , x_5 und x_6 ein, um die Ungleichungen zu Gleichungen zu machen. Natürlich ist dieses Problem nun kein dreidimensionales Problem mehr, sondern ein sechsdimensionales. Mit den Größen $\mathbf{c} = (-5000, -8000, -6000, 0, 0, 0)^T$, $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6)^T$,

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 20 & 40 & 30 & 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad \mathbf{b} = \begin{pmatrix} 50 \\ 20 \\ 1300 \end{pmatrix} \quad (2)$$

haben wir das obige Maximierungsproblem in ein äquivalentes Minimierungsproblem in Normalform übergeführt.

Wir wollen uns jetzt nur noch mit Optimierungsproblemen in Normalform befassen, da ja jede lineare Optimierungsaufgabe in Normalform zu fassen ist.

Weiter setzen wir nun noch für alle weiteren Untersuchungen die Bedingung $\text{Rang} \mathbf{A} = m$ voraus, denn ist $\text{Rang} \mathbf{A} < m$, also die Zeilenvektoren linear abhängig, so ist das System $\mathbf{A} \mathbf{x} = \mathbf{b}$ entweder unlösbar, oder es können Zeilen von \mathbf{A} und entsprechend Komponenten in \mathbf{b} gestrichen werden, so daß die obige Voraussetzung wieder erfüllt ist. Daraus folgt weiterhin $m \leq n$. Mit diesen Voraussetzungen ist das lineare Gleichungssystem $\mathbf{A} \mathbf{x} = \mathbf{b}$ stets lösbar.

Definition: *Es sei ein lineares Optimierungsproblem in Normalform (1) gegeben. Die zu minimierende Funktion $\mathbf{c}^T \mathbf{x}$ nennen wir Zielfunktion. Weiter heißt jeder Vektor $\mathbf{x} \in \mathbb{R}^n$ mit $\mathbf{A} \mathbf{x} = \mathbf{b}$ und $\mathbf{x} \geq \mathbf{0}$ zulässig. Die Menge der zulässigen Vektoren $M = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ wird dementsprechend zulässiger Bereich des Optimierungsproblems genannt. Ist $\hat{\mathbf{x}} \in M$ und $\mathbf{c}^T \hat{\mathbf{x}} \leq \mathbf{c}^T \mathbf{x}$ für alle $\mathbf{x} \in M$, so heißt $\hat{\mathbf{x}}$ Lösung des Problems (1).*

2 Basisvektoren

Im Folgenden werden wir eine für ein lineares Optimierungsproblem sehr wichtige Menge von zulässigen Vektoren diskutieren—die Menge der *Basisvektoren*. Diese ist

endlich und enthält im Falle der Lösbarkeit des Problems eine Optimallösung, wie wir noch sehen werden.

Definition: Ein Basisvektor ist ein zulässiger Vektor \mathbf{x} , zu dem es eine m Elemente enthaltende Indexmenge $I \subset \{1, \dots, n\}$ gibt, so daß

- (i) die Spaltenvektoren \mathbf{a}_j von \mathbf{A} für $j \in I$ linear unabhängig sind und
- (ii) $x_j = 0$ für alle $j \notin I$.

Ist ein Basisvektor $\hat{\mathbf{x}}$ eine Lösung des Optimierungsproblems, so heißt $\hat{\mathbf{x}}$ auch Basislösung.

Es sei $\hat{\mathbf{x}}$ ein zulässiger Vektor. Die Menge $M' = \{\mathbf{x} | \mathbf{A}\mathbf{x} = \mathbf{b}\}$ ist dann $\hat{\mathbf{x}} + \text{Ker}\mathbf{A}$, denn ist $\mathbf{x} \in \hat{\mathbf{x}} + \text{Ker}\mathbf{A}$, so gibt es $\mathbf{v} \in \text{Ker}\mathbf{A}$, so daß $\mathbf{x} = \hat{\mathbf{x}} + \mathbf{v}$. Damit folgt $\mathbf{A}\mathbf{x} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{A}\mathbf{v} = \mathbf{b} + \mathbf{0} = \mathbf{b}$, also $\mathbf{x} \in M'$. Ist andersherum $\mathbf{x} \in M'$, so ist $\mathbf{v} := \mathbf{x} - \hat{\mathbf{x}} \in \text{Ker}\mathbf{A}$, also $\mathbf{x} = \hat{\mathbf{x}} + \mathbf{v}$ und somit $\mathbf{x} \in \hat{\mathbf{x}} + \text{Ker}\mathbf{A}$.

Der zulässige Bereich läßt sich in Anbetracht dieser Feststellung assoziieren mit einer Geraden oder Ebene im \mathbb{R}^3 . Die Basisvektoren als Punkte gesehen sind dann im Falle einer Geraden in der Menge der Durchstoßpunkte der Geraden mit den Hauptebenen und im Falle einer Ebene in der Menge der Schnittpunkte eben dieser mit den Hauptachsen enthalten, in beiden Fällen beschränkt auf den ersten Oktanten (siehe Abbildung 1). Im Allgemeinen ist der zulässige Bereich ein abgeschlossenes Polyeder und die Basisvektoren dessen Ecken.

Satz 2.1 *Ist der zulässige Bereich nicht leer, so existiert ein Basisvektor.*

Wir geben nun einen Beweis dieses Satzes an, der auf folgendem Prinzip beruht: Wir suchen uns einen Vektor $\mathbf{v} \in \text{Ker}\mathbf{A}$ und laufen mit \mathbf{v} von einem zulässigen $\hat{\mathbf{x}}$ ausgehend linear auf M' entlang bis eine neue Komponente Null ist. Dabei gewährleisten wir, daß die Nullkomponenten von $\hat{\mathbf{x}}$ nicht verändert werden, und daß in dem neuen zulässigen Vektor eine Nullkomponente mehr enthalten ist (siehe Abbildung 1).

Beweis: Es sei $\hat{\mathbf{x}}$ ein zulässiger Vektor und $I(\mathbf{x}) = \{i | x_i > 0\}$ für $\mathbf{x} \in M$.

a) Sind die Spaltenvektoren \mathbf{a}_i von \mathbf{A} linear unabhängig für $i \in I := I(\hat{\mathbf{x}})$, so ist $|I| \leq m$ wegen $\text{Rang}\mathbf{A} = m$. Daher ist $\hat{\mathbf{x}}$ bereits ein Basisvektor, denn für $|I| = m$ ist das klar, und ist $|I| < m$, so können wir die $\mathbf{a}_i (i \in I)$ durch andere $\mathbf{a}_j (j \notin I)$ zu einer Basis des \mathbb{R}^m ergänzen und I entsprechend erweitern.

b) Es seien die $\mathbf{a}_i (i \in I)$ linear abhängig. Dann gibt es einen Vektor $\mathbf{v} \in \text{Ker}\mathbf{A}$ mit $v_j = 0$ für $j \notin I$ und $v_i > 0$ für *mindestens* ein $i \in I$. (Sind alle $v_i \leq 0$ für $i \in I$, so wähle man $-\mathbf{v}$) Mit diesem \mathbf{v} sei $\mathbf{x}(\lambda) = \hat{\mathbf{x}} - \lambda\mathbf{v}$. Für hinreichend kleine $\lambda \in \mathbb{R}$ ist $\mathbf{x}(\lambda)$ zulässig, $x_j(\lambda) = 0$ für alle $j \notin I$ und $\lambda \in \mathbb{R}$, und $x_i(\lambda) = 0$ ist gleichwertig mit $\lambda = \hat{x}_i/v_i$ für $i \in I$. Wir definieren also

$$\rho = \min_{i \in I} \left\{ \frac{\hat{x}_i}{v_i} \mid v_i > 0 \right\} > 0$$

und zeigen, daß $\tilde{\mathbf{x}} = \mathbf{x}(\rho)$ zulässig ist. Zunächst ist doch $\tilde{x}_j = 0$ für alle $j \notin I$. Ist $v_i < 0$ für ein i , so gilt weiter $\tilde{x}_i = \hat{x}_i - \rho v_i = \hat{x}_i + \rho|v_i| > 0$, und im Fall $v_i > 0$ gilt

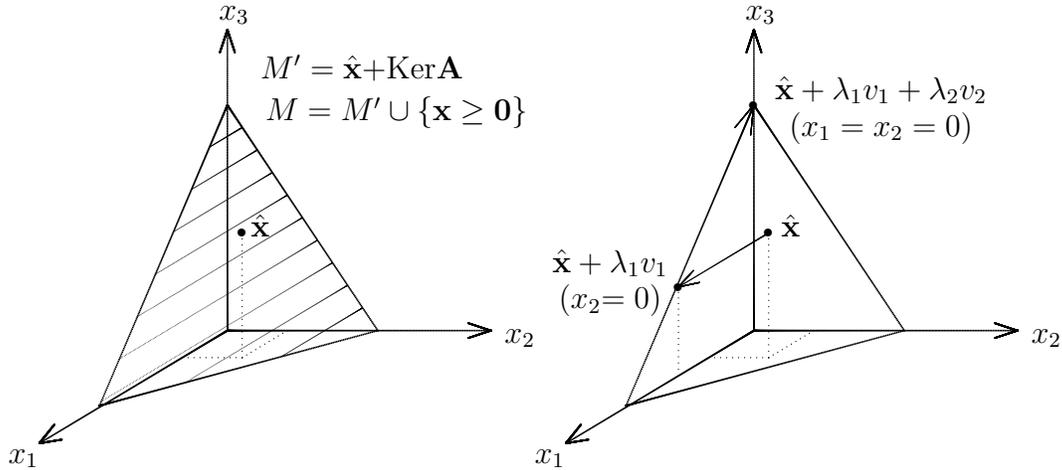


Abbildung 1: Zulässiger Bereich und Basisvektoren

$\tilde{x}_i = \hat{x}_i - \rho v_i \geq \hat{x}_i - (\hat{x}_i/v_i)v_i = 0$. Damit ist $\tilde{\mathbf{x}} \geq \mathbf{0}$ und $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$. Außerdem ist in $\tilde{\mathbf{x}}$ mindestens eine Nullkomponente mehr als in $\hat{\mathbf{x}}$, also $|I(\tilde{\mathbf{x}})| < |I(\hat{\mathbf{x}})|$. Dieses Vorgehen wiederholen wir immer wieder bis die \mathbf{a}_i , $i \in I(\mathbf{x})$, linear unabhängig sind, bis wir also nach Teil a) einen Basisvektor gefunden haben. \square

Dieser Satz ist von fundamentaler Bedeutung für das Verfahren, das wir im nächsten Abschnitt diskutieren wollen. In diesem starten wir nämlich mit einem Basisvektor, welcher natürlich existieren muß – sonst hätte alle Theorie keinen Nutzen.

3 Das Simplexverfahren

Das Simplexverfahren bestimmt zu einem Basisvektor $\hat{\mathbf{x}}$ einen neuen Basisvektor $\tilde{\mathbf{x}}$, so daß $\mathbf{c}^T \tilde{\mathbf{x}} \leq \mathbf{c}^T \hat{\mathbf{x}}$. In der Folge unserer Untersuchungen werden wir sehen, daß wir so der Lösung, falls sie existiert, nicht nur näher kommen, sondern sie sogar mit einem Basisvektor bestimmen können, d.h. wir finden auf diese Weise einen Basisvektor $\tilde{\mathbf{x}}$ mit $\mathbf{c}^T \tilde{\mathbf{x}} \leq \mathbf{c}^T \mathbf{x}$ für alle $\mathbf{x} \in M$. Falls es also eine Lösung des Problems gibt, können wir sicher sein, eine solche in einer endlichen Menge zu finden.

3.1 Allgemeine Vorbereitungen

Es sei $I \subset \{1, \dots, n\}$ eine m -elementige Indexmenge, so daß die Spaltenvektoren \mathbf{a}_i , $i \in I$, von \mathbf{A} linear unabhängig sind. Es sei dann $\mathbf{A}\mathbf{x} = \mathbf{b}$, zunächst für einen

beliebigen Vektor \mathbf{x} . Wir setzen nun $k = n - m$ und damit

$$I := \{i_1, i_2, \dots, i_m\} \quad \text{und} \quad J := \{1, \dots, n\} \setminus I := \{j_1, j_2, \dots, j_k\}. \quad (3)$$

Mit diesen Indexmengen definieren wir (B für Basis und NB für nicht-Basis)

$$\begin{aligned} \mathbf{A}_B &= (\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \dots, \mathbf{a}_{i_m}), \quad \mathbf{c}_B = (c_{i_1}, c_{i_2}, \dots, c_{i_m})^\top, \quad \mathbf{x}_B = (x_{i_1}, x_{i_2}, \dots, x_{i_m})^\top \\ \mathbf{A}_{NB} &= (\mathbf{a}_{j_1}, \mathbf{a}_{j_2}, \dots, \mathbf{a}_{j_k}), \quad \mathbf{c}_{NB} = (c_{j_1}, c_{j_2}, \dots, c_{j_k})^\top, \quad \mathbf{x}_{NB} = (x_{j_1}, x_{j_2}, \dots, x_{j_k})^\top. \end{aligned}$$

Es gilt $\mathbf{A}\mathbf{x} = \mathbf{b}$ und damit $\mathbf{A}_B\mathbf{x}_B + \mathbf{A}_{NB}\mathbf{x}_{NB} = \mathbf{b}$. \mathbf{A}_B ist eine reguläre $m \times m$ -Matrix, also $\mathbf{x}_B = \mathbf{A}_B^{-1}\mathbf{b} - \mathbf{A}_B^{-1}\mathbf{A}_{NB}\mathbf{x}_{NB}$. Mit $\mathbf{y} = \mathbf{A}_B^{-1}\mathbf{b}$ und $\mathbf{D} = \mathbf{A}_B^{-1}\mathbf{A}_{NB}$ ist dann

$$\mathbf{x}_B = \mathbf{y} - \mathbf{D}\mathbf{x}_{NB}. \quad (\text{I})$$

Für die Zielfunktion gilt $\mathbf{c}^\top\mathbf{x} = \mathbf{c}_B^\top\mathbf{x}_B + \mathbf{c}_{NB}^\top\mathbf{x}_{NB} = \mathbf{c}_B^\top\mathbf{y} + (\mathbf{c}_{NB}^\top - \mathbf{c}_B^\top\mathbf{D})\mathbf{x}_{NB}$, mit $\Phi = \mathbf{c}_{NB}^\top - \mathbf{D}^\top\mathbf{c}_B$ also

$$\mathbf{c}^\top\mathbf{x} = \mathbf{c}_B^\top\mathbf{y} + \Phi^\top\mathbf{x}_{NB}. \quad (\text{II})$$

Die Zielfunktion ist somit nur von dem Vektor \mathbf{x}_{NB} abhängig, da Φ für jede gewählte Indexmenge I konstant ist. Ist nun $\Phi \geq \mathbf{0}$, so sehen wir an (II), daß die Zielfunktion um so kleiner ist, je kleiner die Komponenten in \mathbf{x}_{NB} sind. Der kleinste Wert wird also bei $\mathbf{x}_{NB} = \mathbf{0}$ angenommen. Dieses Ergebnis formulieren wir in dem nächsten Satz.

Satz 3.1 *Es sei $\hat{\mathbf{x}}$ ein Basisvektor zu der Indexmenge I und $\Phi \geq \mathbf{0}$. Dann ist $\hat{\mathbf{x}}$ eine Lösung des linearen Optimierungsproblems.*

Beweis: Nach Definition eines Basisvektors ist $\hat{\mathbf{x}}_{NB} = \mathbf{0}$, also $\hat{\mathbf{x}}_B = \mathbf{y}$ nach (I). Ist nun $\Phi \geq \mathbf{0}$, so gilt für jedes zulässige $\mathbf{x} \in \mathbb{R}^n$ mit (II)

$$\mathbf{c}^\top\mathbf{x} = \mathbf{c}_B^\top\mathbf{y} + \Phi^\top\mathbf{x}_{NB} \geq \mathbf{c}_B^\top\mathbf{y} = \mathbf{c}^\top\hat{\mathbf{x}}.$$

Demnach ist $\hat{\mathbf{x}}$ eine Lösung des Problems. □

Im nächsten Satz nehmen wir an, daß nicht $\Phi \geq \mathbf{0}$ gilt, und finden ein gutes Kriterium für die Unlösbarkeit des Problems.

Satz 3.2 *Es sei $\hat{\mathbf{x}}$ ein Basisvektor zu der Indexmenge I und $\Phi_r < 0$ für ein $r \leq k$.*

Weiter sei \mathbf{d}_r der r -te Spaltenvektor der Matrix \mathbf{D} . Dann gilt

a) $\hat{\mathbf{x}}$ ist keine Lösung des Problems.

b) Ist $\mathbf{d}_r \leq \mathbf{0}$, so ist das Problem unlösbar.

Beweis: Es sei $\tilde{\mathbf{x}} \in \mathbb{R}^n$ mit $\tilde{x}_j = 0$ für alle $j \in (J \setminus \{j_r\})$, $\tilde{x}_{j_r} = \varepsilon > 0$ und $\tilde{\mathbf{x}}_B = \hat{\mathbf{x}}_B - \mathbf{D}\tilde{\mathbf{x}}_{NB}$. Für genügend kleines $\varepsilon > 0$ ist $\tilde{\mathbf{x}} \geq \mathbf{0}$, denn $\tilde{\mathbf{x}}_{NB} \geq \mathbf{0}$ und $\tilde{\mathbf{x}}_B = \hat{\mathbf{x}}_B - \tilde{x}_{j_r}\mathbf{d}_r = \hat{\mathbf{x}}_B - \varepsilon\mathbf{d}_r$. Es gilt ferner

$$\mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}_B\tilde{\mathbf{x}}_B + \mathbf{A}_{NB}\tilde{\mathbf{x}}_{NB} = \mathbf{A}_B\hat{\mathbf{x}}_B + (\mathbf{A}_{NB} - \mathbf{A}_B\mathbf{D})\tilde{\mathbf{x}}_{NB} = \mathbf{A}_B\hat{\mathbf{x}}_B = \mathbf{b} .$$

Damit ist $\tilde{\mathbf{x}}$ für genügend kleine $\varepsilon > 0$ zulässig, und es gilt mit der Formel (II) $\mathbf{c}^T\tilde{\mathbf{x}} = \mathbf{c}_B^T\mathbf{y} + \Phi^T\tilde{\mathbf{x}}_{NB} = \mathbf{c}_B^T\mathbf{y} + \Phi_r\varepsilon$.

a) Die Behauptung ergibt sich aus $\mathbf{c}^T\tilde{\mathbf{x}} = \mathbf{c}_B^T\mathbf{y} + \Phi_r\varepsilon < \mathbf{c}_B^T\mathbf{y} = \mathbf{c}^T\hat{\mathbf{x}}$.

b) Ist $\mathbf{d}_r \leq \mathbf{0}$, so ist $\tilde{\mathbf{x}}$ mit $\tilde{\mathbf{x}}_B = \hat{\mathbf{x}}_B - \varepsilon\mathbf{d}_r$ für *alle* $\varepsilon > 0$ zulässig. Wegen $\mathbf{c}^T\tilde{\mathbf{x}} = \mathbf{c}_B^T\mathbf{y} + \Phi_r\varepsilon$ können wir die Zielfunktion somit im zulässigen Bereich beliebig klein machen. Sie ist dort unbeschränkt. Es existiert demnach kein Minimum. \square

3.2 Idee des Simplexverfahrens

Es sei wieder $\hat{\mathbf{x}}$ ein Basisvektor zur Indexmenge I . Für jeden Vektor \mathbf{x} mit $\mathbf{A}\mathbf{x} = \mathbf{b}$ gelten die Gleichungen (I) und (II). Es sei nun $\Phi_r < 0$ für ein $r \leq k$. *Wir suchen einen neuen Basisvektor $\tilde{\mathbf{x}}$ mit $\mathbf{c}^T\tilde{\mathbf{x}} \leq \mathbf{c}^T\hat{\mathbf{x}}$.* Der Clou besteht nun darin, $\tilde{x}_j = 0$ für $j \in (J \setminus \{j_r\})$ und $\tilde{x}_{j_r} \geq 0$ so zu wählen, daß $\tilde{\mathbf{x}}_B$ in (I) eine Nullkomponente mehr besitzt als $\hat{\mathbf{x}}_B$ und die übrigen $\tilde{\mathbf{x}}_B$ -Komponenten ≥ 0 zu lassen. Damit ist dann $\mathbf{c}^T\tilde{\mathbf{x}} = \mathbf{c}_B^T\mathbf{y} + \Phi_r\tilde{x}_{j_r} \leq \mathbf{c}_B^T\mathbf{y} = \mathbf{c}^T\hat{\mathbf{x}}$.

Beispiel: (*Anwendung des Simplexverfahrens in einem Schritt*)

Wir wollen die lineare Optimierungsaufgabe $x + y + 3z = 6$, $2x + y + 5z = 8$, ZF = $4x - 6y - 10z = \min$ lösen. Wir starten mit dem Basisvektor $(2, 4, 0)$ mit x und y als sogenannten *Basisvariablen*. So nennt man die Komponenten im Basisvektor \mathbf{x}_B . Wir stellen die Gleichungen gemäß (I) und (II) um: $x = 2 - 2z$, $y = 4 - z$ und ZF = $-16 - 12z$. Da $k = n - m = 3 - 2 = 1$, ist Φ eine reelle Zahl des Wertes -12 . Unser Basisvektor ist nach Satz 3.2a) also keine Lösung. Um ZF kleiner zu machen, können wir den Wert von z erhöhen. Dabei müssen aber $x, y \geq 0$ bleiben. Daher wählen wir für z das Minimum von 1 und 4, was natürlich 1 ist, nehmen also x als neue Nullkomponente. Der neue Basisvektor ist dann $(0, 3, 1)$, und $z = 1 - x/2$, $y = 3 + x/2$. Weiterhin ist ZF = $-16 - 12 + 6x = -28 + 6x$. D.h. $\Phi = 6 > 0$, und wir haben mit der Lösung $(0, 3, 1)$ das Minimum -28 erreicht.

Wir verfahren nun allgemein wie im Beispiel und setzen zunächst wie oben angedeutet $\tilde{x}_j = 0$ für alle $j \in (J \setminus \{j_r\})$. Nach (I) ist dann $\tilde{\mathbf{x}}_B = \mathbf{y} - \tilde{x}_{j_r}\mathbf{d}_r = \hat{\mathbf{x}}_B - \tilde{x}_{j_r}\mathbf{d}_r$. In jegliche Zeilen aufgedröselst ist diese Gleichung äquivalent mit

$$\tilde{x}_{i_s} = \hat{x}_{i_s} - \tilde{x}_{j_r}d_{sr} \quad \text{für } s = 1, \dots, m . \quad (4)$$

Dabei sei $\mathbf{D} = (d_{ij})$. Im Beispiel sind die \tilde{x}_{i_s} die Variablen x und y . Nun wollen wir doch, daß ein \tilde{x}_{i_s} Null ist. Dann ist $\tilde{x}_{j_r} = \hat{x}_{i_s}/d_{sr}$. Diese Werte vergleichen wir wie im Beispiel für $s = 1, \dots, m$ und wählen davon den kleinsten, also

$$\tilde{x}_{j_r} = \min_{s=1, \dots, m} \left\{ \hat{x}_{i_s}/d_{sr} \mid d_{sr} > 0 \right\} \geq 0 . \quad (5)$$

Damit ist dann $\tilde{x}_{i_s} = \hat{x}_{i_s} - \tilde{x}_{j_r} d_{sr} \geq 0$ für alle $s = 1, \dots, m$ und somit $\tilde{\mathbf{x}}$ zulässig ($\mathbf{A}\tilde{\mathbf{x}} = \mathbf{b}$, da $\tilde{\mathbf{x}}$ (I) erfüllt, und (I) $\iff \mathbf{A}\mathbf{x} = \mathbf{b}$). Wir fassen unsere Ergebnisse nun zusammen und zeigen, daß der neue Vektor $\tilde{\mathbf{x}}$ ein Basisvektor ist.

3.3 Formulierung des Simplexverfahrens

Bevor wir mit der Beschreibung der Funktionsweise des Simplexverfahrens durch den nächsten Satz beginnen, wollen wir noch ein kleines Lemma anbringen.

Lemma 3.3 *Die Spaltenvektoren \mathbf{a}_i , $i \in I$, von \mathbf{A} seien linear unabhängig. Erfüllt ein $\mathbf{x} \in \mathbb{R}^n$ die Gleichung (I), so auch $\mathbf{A}\mathbf{x} = \mathbf{b}$.*

Beweis: Mit $\mathbf{y} = \mathbf{A}_B^{-1}\mathbf{b}$ und $\mathbf{D} = \mathbf{A}_B^{-1}\mathbf{A}_{NB}$ gilt $\mathbf{x}_B = \mathbf{y} - \mathbf{D}\mathbf{x}_{NB}$. Daraus folgt

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{A}_B\mathbf{x}_B + \mathbf{A}_{NB}\mathbf{x}_{NB} \\ &= \mathbf{A}_B(\mathbf{y} - \mathbf{D}\mathbf{x}_{NB}) + \mathbf{A}_{NB}\mathbf{x}_{NB} \\ &= \mathbf{A}_B\mathbf{y} - \mathbf{A}_{NB}\mathbf{x}_{NB} + \mathbf{A}_{NB}\mathbf{x}_{NB} \\ &= \mathbf{A}_B(\mathbf{A}_B^{-1}\mathbf{b}) \\ &= \mathbf{b} . \end{aligned}$$

□

Satz 3.4 (Formulierung des Simplexverfahrens) *Es sei $\hat{\mathbf{x}}$ ein Basisvektor zu der Indexmenge I . Dann gilt*

- Ist $\Phi \geq 0$, so ist $\hat{\mathbf{x}}$ eine Lösung des Problems (1).
- Ist $\Phi_r < 0$ für ein $r \leq k$, und ist $\mathbf{d}_r \leq \mathbf{0}$, so hat das Problem keine Lösung.
- Ist $\Phi_r < 0$ für ein $r \leq k$, und ist b) nicht der Fall, so ist mit

$$\mu = \min_{s=1, \dots, m} \left\{ \frac{\hat{x}_{i_s}}{d_{sr}} \mid d_{sr} > 0 \right\} = \frac{\hat{x}_{i_p}}{d_{pr}} \quad (6)$$

der Vektor $\tilde{\mathbf{x}}$, definiert durch $\tilde{x}_j = 0$ für $j \in (J \setminus \{j_r\})$, $\tilde{x}_{j_r} = \mu$ und $\tilde{\mathbf{x}}_B = \hat{\mathbf{x}}_B - \mu\mathbf{d}_r$ ein Basisvektor zur Indexmenge $\tilde{I} = (I \cup \{j_r\}) \setminus \{i_p\}$.

d) Für $\tilde{\mathbf{x}}$ aus c) gilt $\mathbf{c}^T\tilde{\mathbf{x}} \leq \mathbf{c}^T\hat{\mathbf{x}}$.

Beweis: a) und b) haben wir bereits in den Sätzen 3.1 und 3.2 bewiesen.

c) Zunächst ist $\tilde{x}_j \geq 0$ für $j \in J$. Sei nun $s \in \{1, \dots, m\}$. Ist $d_{sr} > 0$, so gilt $\tilde{x}_{i_s} = \hat{x}_{i_s} - \mu d_{sr} \geq \hat{x}_{i_s} - (\hat{x}_{i_s}/d_{sr})d_{sr} = 0$. Aber auch im Falle von $d_{sr} \leq 0$ gilt $\tilde{x}_{i_s} = \hat{x}_{i_s} - \mu d_{sr} = \hat{x}_{i_s} + \mu |d_{sr}| \geq 0$. Damit ist $\tilde{\mathbf{x}} \geq \mathbf{0}$. Da weiterhin $\tilde{\mathbf{x}}$ die Gleichung (I) erfüllt, genügt $\tilde{\mathbf{x}}$ nach Lemma 3.3 dem Gleichungssystem $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{b}$ und ist somit zulässig. Analog zu den Vorbereitungen sei $\tilde{J} = \{1, \dots, n\} \setminus \tilde{I} = \{j_1, \dots, j_{r-1}, i_p, j_{r+1}, \dots, j_k\}$. Da $\tilde{x}_{i_p} = \hat{x}_{i_p} - \mu d_{pr} = \hat{x}_{i_p} - (\hat{x}_{i_p}/d_{pr})d_{pr} = 0$, ist $\hat{x}_j = 0$ für alle $j \in \tilde{J}$. Es bleibt zu zeigen, daß die \mathbf{a}_i , $i \in \tilde{I}$, linear unabhängig sind. Dazu setzen wir wie üblich $\sum_{i \in \tilde{I}} \lambda_i \mathbf{a}_i = \mathbf{0}$ mit $\lambda_i \in \mathbb{R}$ für $i \in \tilde{I}$. Dann folgt mit $\mathbf{A}_{NB} = \mathbf{A}_B\mathbf{D} \implies \mathbf{a}_{j_r} = \mathbf{A}_B\mathbf{d}_r = \sum_{s=1}^m d_{sr}\mathbf{a}_{i_s}$:

$$\begin{aligned}
\mathbf{0} &= \sum_{i \in \tilde{I}} \lambda_i \mathbf{a}_i = \sum_{\substack{s=1 \\ s \neq p}}^m \lambda_{i_s} \mathbf{a}_{i_s} + \lambda_{j_r} \mathbf{a}_{j_r} \\
&= \sum_{\substack{s=1 \\ s \neq p}}^m \lambda_{i_s} \mathbf{a}_{i_s} + \sum_{s=1}^m \lambda_{j_r} d_{sr} \mathbf{a}_{i_s} \\
&= \sum_{\substack{s=1 \\ s \neq p}}^m (\lambda_{i_s} + \lambda_{j_r} d_{sr}) \mathbf{a}_{i_s} + \lambda_{j_r} d_{pr} \mathbf{a}_{i_p} .
\end{aligned}$$

Wegen der linearen Unabhängigkeit der \mathbf{a}_{i_s} , $s = 1, \dots, m$, folgt zunächst, daß $\lambda_{j_r} d_{pr} = 0$, was $\lambda_{j_r} = 0$ bedeutet, da ja aus der Definition von μ hervorgeht, daß $d_{pr} > 0$. Somit sind alle $\lambda_i = 0$ für $i \in \tilde{I}$.

d) Nach der Gleichung (II) und $\Phi^T \tilde{\mathbf{x}}_{NB} = \Phi_r \mu$ gilt

$$\mathbf{c}^T \tilde{\mathbf{x}} = \mathbf{c}_B^T \mathbf{y} + \Phi_r \mu \leq \mathbf{c}_B^T \mathbf{y} = \mathbf{c}^T \hat{\mathbf{x}} ,$$

da ja $\Phi_r < 0$ und $\mu \geq 0$. □

Man kann also, ausgehend von einem Basisvektor $\hat{\mathbf{x}}^1$, mit Satz 3.4c) eine Folge $(\hat{\mathbf{x}}^k)$ von Basisvektoren bestimmen, mit denen die Folge $(\mathbf{c}^T \hat{\mathbf{x}}^k)$ monoton fallend ist. Tritt für ein Folgenglied einer der Fälle a) oder b) aus Satz 3.4 ein, so haben wir ein Ergebnis für die Optimierungsaufgabe gefunden. Dieses Vorgehen nennt man das **Simplexverfahren**.

Man kann in diesem Verfahren aber auch auf gewisse Hürden stoßen, nämlich auf Basisvektoren, in denen eine Basiskomponente verschwindet (d.h. $\hat{x}_i = 0$ für ein $i \in I$). Solche Basisvektoren nennen wir *entartet*. Wir wollen nun überprüfen, wann solche entarteten Basisvektoren auftreten, und auf welche Weise sie das Simplexverfahren behindern können.

3.4 Wohldefiniertheit des Verfahrens

Auf den ersten Blick sieht das Simplexverfahren mit Satz 3.4 so aus, als ob keine Probleme auftauchen könnten. Es wird in jedem Schritt ein Basisvektor erzeugt, und die Zielfunktion wird immer kleiner... Und genau dort hakt's. Ein kleiner Blick auf den Teil d) des Satzes zeigt, daß auch der Fall $\mathbf{c}^T \tilde{\mathbf{x}} = \mathbf{c}^T \hat{\mathbf{x}}$ auftreten kann. Im Beweis dieses Teils wird deutlich, daß das genau dann auftritt, wenn $\Phi_r \mu = 0$. Da ja $\Phi_r < 0$, muß also $\mu = 0$ sein. Und $\mu = 0$ genau dann, wenn $\hat{x}_{i_p} = 0$ und $d_{pr} > 0$ für ein $p \in \{1, \dots, m\}$. $\hat{\mathbf{x}}$ ist dann notwendig entartet.

Ist $\hat{\mathbf{x}}$ entartet mit $\hat{x}_{i_p} = 0$ und $p \in \{1, \dots, m\}$, dann ist $\mathbf{c}^T \tilde{\mathbf{x}} = \mathbf{c}^T \hat{\mathbf{x}}$ genau dann, wenn $d_{pr} > 0$. In diesem Fall ist $\tilde{\mathbf{x}} = \hat{\mathbf{x}}$, aber $\tilde{I} \neq I$. Tritt in den nächsten Schritten des Simplexverfahrens immer wieder dieser Fall auf, entsteht eine Folge

von Basisvektoren $\hat{\mathbf{x}}^k$ zu Indexmengen I_k , wobei $\hat{\mathbf{x}}^{k+1} = \hat{\mathbf{x}}^k$ und $I_{k+1} \neq I_k$. So kann dann der Fall eintreten, daß es $n_1, n_2 \in \mathbb{N}$ gibt, so daß $I_{n_1} = I_{n_2}$, wobei $n_1 \neq n_2$. Dann befinden wir uns in einem Schlamassel, denn wir kommen nicht weiter. Ein solches Schlamassel nennt man *Zyklus*. Man kann aber solche Zyklen durch bestimmte Modifizierungen des Simplexverfahrens vermeiden, auf die wir hier jedoch nicht eingehen wollen.

Man kann schon in der obigen Erklärung, wie es zu solchen Zyklen kommen kann, sehen, daß diese sehr unwahrscheinlich sind. Tatsächlich kommen sie in praktischen Anwendungen so gut wie nie vor. Da man sie vermeiden kann, schließen wir diese Möglichkeit im Folgenden aus. Wir wollen uns aber doch nochmal mit der Entstehung von entarteten Basisvektoren befassen.

Angenommen, der Basisvektor $\hat{\mathbf{x}}$ zur Indexmenge I aus Satz 3.4 sei *nicht* entartet. Wir wollen nun untersuchen, was für $\hat{\mathbf{x}}$ vorausgesetzt werden muß, damit $\tilde{\mathbf{x}}$ entartet ist, damit also $\tilde{x}_i = 0$ für mindestens ein $i \in \tilde{I} = \{i_1, \dots, i_{p-1}, j_r, i_{p+1}, \dots, i_m\}$. Das ist genau dann der Fall, wenn $\mu = 0$ oder $\hat{x}_{i_q} = \mu d_{qr}$ für ein $q \in \{1, \dots, m\} \setminus \{p\}$. Wäre $\mu = 0$, dann ist $\hat{x}_{i_p} = 0$, also $\hat{\mathbf{x}}$ entartet, was unserer Annahme widerspricht. Es sei also $\mu \neq 0$. Wegen $d_{qr} > 0$ (denn wäre $d_{qr} \leq 0$, so wäre $\hat{x}_{i_q} \leq 0$, also $\hat{\mathbf{x}}$ entartet oder nicht zulässig) ist dann $\mu = \hat{x}_{i_q}/d_{qr}$. D.h., das Minimum, mit dem μ definiert wurde, wird zweimal angenommen, nämlich für $s = q$ und $s = p$.

Satz 3.5 *Das Simplexverfahren bricht nach endlich vielen Schritten ab.*

Beweis: Das Verfahren erzeugt eine Folge (I_j) von Indexmengen. Da wir die Möglichkeit von Zyklen ausgeschlossen haben, gilt $I_j \neq I_i$ für alle $i \neq j$. Da es nur endlich viele solcher m -elementigen Indexmengen gibt, ist (I_j) endlich, und damit die Behauptung gezeigt. \square

Satz 3.6 *Wenn (1) lösbar ist, bricht das Simplexverfahren (nach endlich vielen Schritten) mit einer Basislösung von (1) ab.*

Beweis: Da (1) lösbar ist, kann der Fall b) im Satz 3.4 im Laufe des Verfahrens nicht eintreten. Das Simplexverfahren erzeugt wieder eine Folge (I_j) von m -elementigen Indexmengen. Nach dem Satz 3.5 ist diese endlich. Es sei I_R die Indexmenge, mit der das Verfahren endet. Der Fall c) aus Satz 3.4 kann nun nicht eintreten, denn sonst gäbe es eine weitere Indexmenge in der Folge. Es bleibt der Fall a) übrig, nach dem der Basisvektor $\hat{\mathbf{x}}_R$ zur Indexmenge I_R eine Basislösung von (1) ist. \square

3.5 Ablauf des Simplexverfahrens

In diesem Kapitel wollen wir darauf eingehen, wie man das Simplexverfahren mit wenigen Operationen durchführen kann. Mit diesen Ergebnissen stellen wir dann einen Algorithmus auf und führen dann mit dessen Hilfe das Simplexverfahren an unserem Beispiel aus Kapitel 1 durch.

Wir nehmen an, wir hätten einen Basisvektor $\hat{\mathbf{x}}$ zur Indexmenge I . Ein Schritt im Simplexverfahren besteht nun darin, die Fälle a), b) und c) des Satzes 3.4 durchzugehen und bei Eintreffen von a) oder b) das Verfahren zu stoppen. Ist c) der Fall, so bestimmen wir μ , also p , und damit \tilde{I} .

Um einen der Fälle a), b) oder c) auf Richtigkeit zu prüfen, müssen wir $\Phi = \mathbf{c}_{NB} - \mathbf{D}^T \mathbf{c}_B$ berechnen. Wir wollen uns dabei aber die Berechnung der Matrix \mathbf{D} sparen, indem wir uns die Beziehung $\mathbf{D} = \mathbf{A}_B^{-1} \mathbf{A}_{NB}$ zunutze machen. Ist nämlich $\mathbf{z} \in \mathbb{R}^m$ die Lösung des LGS $\mathbf{A}_B^T \mathbf{z} = \mathbf{c}_B$, dann gilt $\mathbf{D}^T \mathbf{c}_B = \mathbf{A}_{NB}^T \mathbf{z}$, also $\Phi = \mathbf{c}_{NB} - \mathbf{A}_{NB}^T \mathbf{z}$.

Fall a) ist somit leicht zu prüfen. Im Fall b) braucht man jedoch den Vektor \mathbf{d}_r . Die Beziehung, die uns im vorigen Absatz bereits geholfen hat, ist auch hier angebracht. Die r -te Zeile lautet nämlich $\mathbf{d}_r = \mathbf{A}_B^{-1} \mathbf{a}_{j_r}$. Der Vektor \mathbf{d}_r läßt sich somit aus dem LGS $\mathbf{A}_B \mathbf{d}_r = \mathbf{a}_{j_r}$ gewinnen. Damit ist dann auch der Fall b) leicht zu prüfen.

Es bleibt der Fall c). Hier brauchen wir den Vektor $\hat{\mathbf{x}}$, der noch nicht unbedingt gegeben zu sein braucht. Den berechnet man natürlich mit dem LGS $\mathbf{A}_B \hat{\mathbf{x}}_B = \mathbf{b}$. Mit den Vektoren $\hat{\mathbf{x}}_B$ und \mathbf{d}_r kann nun auch μ berechnet werden. Wir fassen diese Ergebnisse in einen

Algorithmus: (Das Simplexverfahren)

(S.1) Berechne das LGS

$$\mathbf{A}_B^T \mathbf{z} = \mathbf{c}_B \quad . \quad (\text{S1})$$

Dann ist $\Phi = \mathbf{c}_{NB} - \mathbf{A}_{NB}^T \mathbf{z}$.

Berechne damit $\Phi_{\min} = \min\{\Phi_s \mid s = 1, \dots, k\} = \Phi_r$.

Ist $\Phi_{\min} \geq 0$: STOP

(S.2) Der Vektor $\mathbf{d} := \mathbf{d}_r$ ergibt sich aus dem LGS

$$\mathbf{A}_B \mathbf{d} = \mathbf{a}_{j_r} \quad . \quad (\text{S2})$$

Berechne $d_{\min} = \max\{d_s \mid s = 1, \dots, k\}$.

Ist $d_{\min} \leq 0$: STOP

(S.3) Der Vektor \mathbf{x}_B ergibt sich aus dem LGS

$$\mathbf{A}_B \mathbf{x}_B = \mathbf{b} \quad . \quad (\text{S3})$$

Damit ermittle den Wert von p aus

$$\frac{x_{i_p}}{d_p} = \min_{s=1, \dots, m} \left\{ \frac{x_{i_s}}{d_s} \mid d_s > 0 \right\} \quad .$$

(S.4) Setze $I_{\text{neu}} = (I \cup \{j_r\}) \setminus \{i_p\}$.

Gehe zu (S.1).

Beispiel: (Fortsetzung des Beispiels aus Kapitel 1) Wir hatten das lineare Bauern-Optimierungsproblem bereits auf Normalform gebracht mit

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 20 & 40 & 30 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 50 \\ 20 \\ 1300 \end{pmatrix} \quad \text{und} \quad \mathbf{c} = \begin{pmatrix} -5000 \\ -8000 \\ -6000 \\ \mathbf{0} \end{pmatrix},$$

wobei $\mathbf{0} \in \mathbb{R}^3$. Wir wenden nun das Simplexverfahren an:

1. *Schritt:* $I = \{4, 5, 6\}; J = \{1, 2, 3\}; \mathbf{A}_B = \mathbf{E}$

Das LGS (S1) hat die Lösung $\mathbf{z} = \mathbf{0}$, also $\Phi = \mathbf{c}_{NB} = (-5000, -8000, -6000)^T$. Daraus folgt dann $r = 2$. Nun lösen wir die Gleichungssysteme (S2) und (S3). Diese haben die Lösungen $\mathbf{d} = (1, 1, 40)^T$ und $\mathbf{x}_B = (50, 20, 1300)^T$. Damit ist $\mu = \min_{s=1,2,3} \{x_{i_s}/d_s \mid d_s > 0\} = \min\{50, 20, 32.5\}$, also $p = 2$. Es werden die Indizes i_2 und j_2 getauscht, d.h. 5 und 2: $I_{neu} = \{4, 2, 6\}$.

2. *Schritt:* $I = \{4, 2, 6\}; J = \{1, 5, 3\}; \mathbf{A}_B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 40 & 1 \end{pmatrix}$

(S1) hat die Lösung $\mathbf{z} = (0, -8000, 0)^T$. Damit ist $\Phi = (-5000, 8000, -6000)^T$. Daraus folgt $r = 3$. (S2) und (S3) besitzen die Lösungen $\mathbf{d} = (0, 0, 1)^T$ und $\mathbf{x}_B = (30, 20, 500)^T$. Hieraus sehen wir sofort, daß $p = 3$. Es werden die Indizes i_3 und j_3 getauscht, d.h. 6 und 3: $I_{neu} = \{4, 2, 3\}$.

3. *Schritt:* $I = \{4, 2, 3\}; J = \{1, 5, 6\}; \mathbf{A}_B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 40 & 30 \end{pmatrix}$

Die Lösung von (S1) ist $\mathbf{z} = (0, 0, -200)^T$. Damit ist $\Phi = (-1000, 0, 200)^T$, also ist $r = 1$. Wir lösen wieder die Gleichungssysteme (S2) und (S3): $\mathbf{d} = (-0.33, 0, 0.33)^T$ und $\mathbf{x}_B = (13.33, 20, 16.66)^T$. Damit ist $p = 3$. Es werden die Indizes i_3 und j_1 getauscht, d.h. 3 und 1: $I_{neu} = \{4, 2, 1\}$.

4. *Schritt:* $I = \{4, 2, 1\}; J = \{3, 5, 6\}; \mathbf{A}_B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 40 & 20 \end{pmatrix}$

$\mathbf{z} = (0, 2000, -250)^T$ ist die Lösung von (S1). Daraus folgt $\Phi = (1500, -2000, 250)^T$ und somit $r = 2$. $\mathbf{d} = (1, 1, -2)^T$ und $\mathbf{x}_B = (5, 20, 25)^T$ lösen die Systeme (S2) und (S3). Damit ist $\mu = \min\{5, 20\} = 5$, also $p = 1$. Es werden die Indizes i_1 und j_2 vertauscht, d.h. 4 und 5: $I_{neu} = \{5, 2, 1\}$.

5. *Schritt:* $I = \{5, 2, 1\}; J = \{3, 4, 6\}; \mathbf{A}_B = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 40 & 20 \end{pmatrix}$

Die Lösung von (S1) ist $\mathbf{z} = (-2000, 0, -150)^T$. Es folgt $\Phi = (500, 2000, 150)^T > \mathbf{0}$. Wir sind also bei der Lösung angekommen. Die Lösung des Systems (S3) ist der Vektor $\mathbf{x}_B = (5, 15, 35)^T$. Der Lösungsvektor für das Optimierungsproblem (1) ist dann $\mathbf{x} = (35, 15, 0, 0, 5, 0)^T$. Der Zielfunktionswert ist $\mathbf{c}^T \mathbf{x} = -295000$.

Der Landwirt kann also 35 ha Weizen und 15 ha Rüben anbauen, um auf den maximal zu erreichenden Gewinn von 295000,00 DM zu kommen.

3.6 Bestimmung eines Basisvektors

Um das Simplexverfahren starten zu können, müssen wir einen Basisvektor kennen. Bei kleineren Dimensionen können wir, ausgehend von einem zulässigen Vektor, wie im Beweis von Satz 2.1 vorgehen. Das erfordert für größere n bzw. m jedoch viel zuviel Rechenaufwand. Wir können uns in diesem Falle eines Hilfsproblems bedienen. Da das Lösen von Problem (1) dann aus zwei Phasen besteht, nennt man diese Methode logischerweise *Zweiphasenmethode*.

Zunächst bemerken wir, dass wir im Ausgangsproblem (1) $\mathbf{b} \geq \mathbf{0}$ annehmen dürfen. Ist dies nämlich nicht der Fall, so können wir doch die entsprechenden Zeilen von \mathbf{A} mit -1 multiplizieren. Mit $\mathbf{v} \in \mathbb{R}^m$, $\mathbf{z}^T = [\mathbf{x}^T, \mathbf{v}^T]$, $\tilde{\mathbf{c}}^T = [\mathbf{0}, \mathbf{e}]$, $\mathbf{e} = (1, 1, \dots, 1) \in \mathbb{R}^m$, $\tilde{\mathbf{A}} = [\mathbf{A}, \mathbf{E}]$, wobei \mathbf{E} die $m \times m$ -Einheitsmatrix sei, und $\mathbf{b} \geq \mathbf{0}$ formulieren wir nun das Hilfsproblem

$$\begin{aligned} \tilde{\mathbf{c}}^T \mathbf{z} &= \min \\ \tilde{\mathbf{A}} \mathbf{z} &= \mathbf{b} \\ \mathbf{z} &\geq \mathbf{0} \end{aligned} \quad (7)$$

Satz 3.7 *Das Problem (7) ist lösbar, und der Vektor $\bar{\mathbf{z}}^T = [\mathbf{0}, \mathbf{b}^T]$ ist ein Basisvektor zur Indexmenge $I = \{n+1, n+2, \dots, n+m\}$. Weiter gilt mit einer Basislösung $\mathbf{z}^* = \begin{bmatrix} \mathbf{x}^* \\ \mathbf{v}^* \end{bmatrix}$ von (7) zur Indexmenge I^* :*

- a) *Ist $\tilde{\mathbf{c}}^T \mathbf{z}^* > 0$, so ist der zulässige Bereich von (1) leer.*
- b) *Ist $\tilde{\mathbf{c}}^T \mathbf{z}^* = 0$, so ist \mathbf{x}^* ein Basisvektor von (1).*

Beweis: Da $\bar{\mathbf{z}}$ zulässig ist, ist der zulässige Bereich von (7) nicht leer, und die Zielfunktion ist durch 0 nach unten beschränkt. Daher ist (7) lösbar. Daß $\bar{\mathbf{z}}$ ein Basisvektor ist, muß wohl nicht gezeigt werden.

- a) In diesem Fall gibt es keinen zulässigen Vektor der Form $\mathbf{z}^T = [\mathbf{x}^T, \mathbf{0}]$, denn sonst wäre $\tilde{\mathbf{c}}^T \mathbf{z} = 0 < \tilde{\mathbf{c}}^T \mathbf{z}^*$, was einen Widerspruch darstellt. Es gibt also keinen Vektor mit $\mathbf{A}\mathbf{x} = \mathbf{b}$ und $\mathbf{x} \geq \mathbf{0}$. Der zulässige Bereich von (1) ist somit leer.
- b) Es sei $\tilde{\mathbf{c}}^T \mathbf{z}^* = 0$, also $\mathbf{v}^* = \mathbf{0}$. Dann ist $\mathbf{A}\mathbf{x}^* = \mathbf{b}$ und $\mathbf{x}^* \geq \mathbf{0}$. \mathbf{x}^* ist also für (1) zulässig. In \mathbf{z}^* verschwinden mindestens n Komponenten. Darunter befinden sich die \mathbf{v}^* -Komponenten; in \mathbf{x}^* sind somit mindestens $k = n - m$ Komponenten Null. Sind genau k Nullkomponenten in \mathbf{x}^* , so enthält I^* nur Indizes aus $\{1, \dots, n\}$, und \mathbf{x}^* ist ein Basisvektor von (1) zur Indexmenge I^* . Besitzt \mathbf{x}^* mehr als k Nullkomponenten, so ist \mathbf{z}^* entartet. I^* kann dann auch Elemente aus $\{n+1, n+2, \dots, n+m\}$ enthalten. Ist dies nicht der Fall, so ist \mathbf{x}^* entarteter Basisvektor von (1) zur Indexmenge I^* . Sei nun aber $I^* \not\subset \{1, \dots, n\}$. Dann sind die Vektoren $\mathbf{a}_i, i \in I_{(1)}^* := I^* \cap \{1, \dots, n\}$, linear unabhängig. Wegen $\text{Rang } \mathbf{A} = m$ kann man aber nun eine $(m - |I_{(1)}^*|)$ -elementige Menge $I \subset \{1, \dots, n\}$ finden, so daß die Vektoren

$$\mathbf{a}_i, i \in I_{(1)}^* \quad \text{und} \quad \mathbf{a}_j, j \in I$$

eine Basis des \mathbb{R}^m darstellen. \mathbf{x}^* ist dann ein entarteter Basisvektor von (1) zur Indexmenge $I_{(1)}^* \cup I$. In jedem Fall ist \mathbf{x}^* ein Basisvektor von (1). \square

Wenn man also noch keinen Basisvektor für (1) gefunden hat, dann kann man das Simplexverfahren auf das Hilfsproblem (7) anwenden. Starten läßt es sich mit dem Basisvektor $\mathbf{z} = \begin{pmatrix} \mathbf{0} \\ \mathbf{b} \end{pmatrix}$, bzw. mit der Indexmenge $I = \{n+1, n+2, \dots, n+m\}$. Da das Problem lösbar ist, bricht das Simplexverfahren nach Satz 3.6 mit einer Basislösung \mathbf{z}^* ab. Diese hilft uns dann, gemäß Satz 3.7, beim Finden einer Lösung von (1) weiter.

3.7 Implementierung des Simplexverfahrens

Der nun folgende Code ist für MATLAB geschrieben. Er beinhaltet eine Funktion, die das Simplexverfahren durchführt mit den Parametern \mathbf{A} , \mathbf{b} , \mathbf{c} und I . I ist dabei eine Indexmenge, zu der ein Basisvektor existiert.

Programm 3.8 (*Das Simplexverfahren mit vorgegebener Indexmenge*)

```
function [xL,fL,IL,SL]=Simplex(A,b,c,I)

% Dimensionsbestimmungen und Anderes
n=length(c); m=length(b); k=n-m; SL=1;
error1='Dimensions do not fit. Check your inputs!';
error2='No solution because of unbounded function';
if ((any(size(A)=[m,n])|(length(I)~=m))
    error(error1), end;

STOP=0; J=1:n; J(I)=[];
AB=A(:,I); ANB=A(:,J); cB=c(I); cNB=c(J);
z=AB\'cB; Phi=cNB-ANB'*z;
[PhiMin,r]=min(Phi);
if (PhiMin>=0), STOP=1; end;

while (STOP==0)
    d=AB\ANB(:,r); dmax=max(d);
    if (dmax<=0), error(error2), end;
    xB=AB\b; s=1;
    for i=1:m
        if (d(i)>0)
            q=xB(i)/d(i);
            if (s==1)
                qmin=q; s=2; end;
            if (q<=qmin)
                p=i; qmin=q; end;
        end;
    end;
end;
```

```

    end;
end;
dummy=I(p); I(p)=J(r); J(r)=dummy;
AB(:,p)=A(:,I(p)); ANB(:,r)=A(:,J(r));
cB(p)=c(I(p)); cNB(r)=c(J(r)); SL=SL+1;
z=AB'\cB; Phi=cNB-ANB'*z;
[PhiMin,r]=min(Phi);
if (PhiMin>=0), STOP=1; end;
end;

%Ausgabeparameter vorbereiten
IL=I; xB=AB\b;
xL=(1:n)'; xL(I)=xB; xL(J)=zeros(k,1);
fL=c'*xL;

```

Das nun folgende Programm ist ebenfalls in MATLAB Code. Die Funktion löst das Hilfsproblem (7) und startet dann mit einem Basisvektor das Simplexverfahren nach Programm 3.8.