

# Eine Alternativ-Methode für die numerische Lösung linearer Optimierungsaufgaben

Horst Hollatz  
Fakultät für Mathematik  
Otto-von-Guericke-Universität  
39016 Magdeburg  
Postfach 4120

9. September 1997

## Zusammenfassung

Auf der Grundlage eines modifizierten Farkas-Lemmas über die Lösbarkeit eines linearen Ungleichungssystems wird eine endliche, allgemeine Methode zur numerischen Lösung linearer Optimierungsaufgaben entwickelt.

## Abstract

Using a modified Farkas-lemma on the solvability of a linear inequality system it is given a finite, general method for solving linear optimization problems.

Für die Untersuchungen ist das folgende Lemma wesentlich.

**Lemma 1** *Es seien  $A$  eine beliebige  $(m, n)$ -Matrix und  $c \in \mathbb{R}^n$ . Das System*

$$A^T u = c, u \geq 0$$

*ist genau dann unlösbar, wenn das System  $Ax \leq 0, c^T x > 0$  eine Lösung der Form  $x = \alpha c + A^T u$  besitzt.*

**Beweis.** Nehmen wir zunächst an, daß der Durchschnitt beider Mengen nichtleer ist. Insbesondere ist dann das System  $Ax \leq 0, c^T x > 0$  lösbar, woraus mit dem bekannten Lemma von Farkas (vgl. z. B. [3]) die Unlösbarkeit des ersten Systems folgt. Um die Notwendigkeit der Bedingung zu beweisen, haben wir vorauszusetzen, daß das System  $A^T u = c, u \geq 0$  unlösbar ist. Nach dem Farkas-Lemma gilt dann

$$\{ x \mid Ax \leq 0, c^T x > 0 \} \neq \emptyset.$$

Angenommen, die beiden Mengen in der Behauptung des Lemmas haben einen leeren Durchschnitt. Dann gibt es nach dem Trennungssatz (vgl. z. B. [2]) eine Hyperebene, die beide Mengen trennt; es existiert somit ein Vektor  $d \in \mathbb{R}^n, d \neq 0$  mit

$$d^T x \leq 0 \quad \forall x : x = \alpha c + A^T u$$

und

$$d^T x \geq 0 \quad \forall x : Ax \leq o, c^T x > 0.$$

Wiederum mit dem Farkas-Lemma, jetzt auf die zweite Bedingung angewendet, folgt, daß das System

$$d = A^T u + \alpha c$$

lösbar ist. Mit der ersten Bedingung ergibt sich daraus, daß  $d^T d \leq 0$ , also  $d = o$  sein muß. Dieser Widerspruch beweist das Lemma. \*  
Als Folgerung ergibt sich:

Wenn das System  $Ax \leq o, c^T x > 0$  lösbar ist, so hat es auch eine Lösung der Form  $x = \alpha c + A^T u$ .

Betrachtet wird das lineare Optimierungsproblem

$$P : \quad \max\{c^T x \mid Ax \leq b\}$$

mit  $c \in \mathbb{R}^n, b \in \mathbb{R}^m, A = (a_{ij})_{m,n}$ . Es seien  $A_1^T, \dots, A_m^T$  die Zeilenvektoren der Matrix  $A$ . Ohne Beschränkung der Allgemeinheit wird angenommen, daß  $\|c\| = \|A_i\| = 1, i = 1, \dots, m$  gilt. Zu einer beliebigen zulässigen Lösung  $x$  betrachten wir die  $\varepsilon$ -aktive Indexmenge

$$I_\varepsilon(x) = \{i \mid A_i^T x \geq b_i - \varepsilon\}, \quad (\varepsilon \geq 0).$$

Zu einer Indexmenge  $I \subseteq \{1, \dots, m\}$  und einem beliebigen nichtnegativen Vektor  $z \in \mathbb{R}_+^m$  bilden wir die Menge

$$\Upsilon(I, z) = \{y \mid A_i^T y = -z_i, i \in I, c^T y = 1, y = \alpha_0 c + \sum_{i \in I} \alpha_i A_i\},$$

deren Elemente wir Achsvektoren nennen. Achsvektoren definieren zulässige Richtungen, auf denen die Zielfunktion wächst, falls der Vektor  $z$  so gewählt wurde, daß das System lösbar ist. Wir betrachten daher auch noch die Menge aller Lösbarkeitsvektoren

$$\mathcal{Z}(I) = \{z \in \mathbb{R}_+^m \mid \Upsilon(I, z) \neq \emptyset\}.$$

Als Indexmengen  $I$  werden wir hier die zu einer zulässigen Lösung  $x$  gehörige  $\varepsilon$ -aktive Indexmenge  $I_\varepsilon(x)$  verwenden und entsprechend von Achsvektoren in einer zulässigen Lösung sprechen.

Für eine optimale Lösung  $x^*$  gilt bekanntlich

$$c = \sum_{i \in I_0(x^*)} u_i A_i, \quad u_i \geq 0, i \in I_0(x^*).$$

Folglich existiert in einer optimalen Lösung kein Achsvektor. Daher haben wir mit dem Lemma den folgenden Optimalitätstest:

Eine zulässige Lösung  $x^*$  von  $P$  ist genau dann optimal, wenn  $\mathcal{Z}(I_0(x^*)) = \emptyset$  gilt.

Es sei nun  $x$  eine nichtoptimale, zulässige Lösung von  $P$ ,  $z \in \mathcal{Z}(I_\varepsilon(x))$  und  $y \in \Upsilon(I_\varepsilon(x), z)$  ein Achsvektor. Der Strahl

$$x(\lambda) = x + \lambda y, \quad \lambda > 0$$

weist in den zulässigen Bereich. Von einer  $\varepsilon$ -aktiven Hyperebene

$$H_i = \{ \mathbf{x} \mid \mathbf{A}_i^T \mathbf{x} = b_i \}, (i \in I_\varepsilon(\mathbf{x}))$$

hat ein fixierter Punkt  $\mathbf{x} + \lambda \mathbf{y}$  des Strahls höchstens den Abstand  $\varepsilon + \lambda z_i$ :

$$\varrho(\mathbf{x}(\lambda), H_i) \leq \varepsilon + \lambda z_i.$$

Es sei

$$\mu = \min\{z_i, i \in I_\varepsilon(\mathbf{x})\}.$$

Die Kugel

$$K_\lambda = \{ \mathbf{x} \mid \|\mathbf{x}(\lambda) - \mathbf{x}\| \leq \lambda \mu \}$$

mit dem Mittelpunkt  $\mathbf{x}(\lambda)$  und dem Radius  $\lambda \mu$  liegt für kleine  $\lambda$  im zulässigen Bereich. Man kann die Zielfunktion von P über einer solchen Kugel  $K_\lambda$  maximieren:

$$\max\{ \mathbf{c}^T \mathbf{x} \mid \|\mathbf{x}(\lambda) - \mathbf{x}\| \leq \lambda \mu \}.$$

Diese Aufgabe hat den Optimalpunkt

$$\mathbf{x}(\lambda) + \lambda \mu \cdot \mathbf{c}.$$

Somit ist

$$\mathbf{x}(\lambda) + \lambda \mu \cdot \mathbf{c} = \mathbf{x} + \lambda \cdot (\mathbf{y} + \mu \mathbf{c}), \quad \lambda > 0$$

ein vom zulässigen Punkt  $\mathbf{x}$  ausgehender Strahl, der ein Stück durch den zulässigen Bereich verläuft und auf dem der Wert der Zielfunktion um  $\lambda(1 + \mu)$  wächst:

$$\mathbf{c}^T (\mathbf{x} + \lambda(\mathbf{y} + \mu \mathbf{c})) = \mathbf{c}^T \mathbf{x} + \lambda(1 + \mu).$$

Ist nun  $\mathbf{x}$  ein innerer Punkt des zulässigen Bereiches, also  $I_0(\mathbf{x}) = \emptyset$ , so gilt offenbar  $\mathbf{y} = \mathbf{c}$ . Wir haben daher den maximalen Wert für  $\lambda$  zu wählen, um einen maximalen Zuwachs im Wert der Zielfunktion zu erreichen, d.h. für  $\lambda$  ist jener Wert zu nehmen, bei dem der Strahl  $\mathbf{x} + \lambda(\mathbf{y} + \mu \mathbf{c})$ ,  $\lambda > 0$  den zulässigen Bereich verläßt. Dieses geschieht für

$$\lambda = \Lambda_\varepsilon(\mathbf{x}, \mathbf{y}) = \min \left\{ \frac{b_i - \mathbf{A}_i^T \mathbf{x}}{\mathbf{A}_i^T (\mathbf{y} + \mu \mathbf{c})} \mid \mathbf{A}_i^T (\mathbf{y} + \mu \mathbf{c}) > 0, i \notin I_\varepsilon(\mathbf{x}) \right\}.$$

Aus diesen Überlegungen leitet sich das folgende Algorithmenmodell ab, das mit einer beliebigen zulässigen Lösung  $\mathbf{x}^0$  und  $\varepsilon > 0$  startet:

$$\mathcal{A}(\mathbf{x}^0, \varepsilon) : \begin{cases} \mathbf{z}^r & = \varphi(\mathcal{Z}(I_\varepsilon(\mathbf{x}^r))), \\ \mu^r & = \min\{z_i^r, i \in I_\varepsilon(\mathbf{x}^r)\}, \\ \mathbf{y}^r & = \psi(\Upsilon(I_\varepsilon(\mathbf{x}^r), \mathbf{z}^r)), \\ \lambda^r & = \Lambda_\varepsilon(\mathbf{x}^r, \mathbf{y}^r), \\ \mathbf{x}^{r+1} & = \mathbf{x}^r + \lambda^r (\mathbf{y}^r + \mu^r \mathbf{c}). \end{cases}$$

Hierin seien  $\varphi$  eine Auswahlfunktion für die Menge  $\mathcal{Z}(I)$  und  $\psi$  eine Auswahlfunktion für die Menge  $\Upsilon(I, \mathbf{z})$ . Der Algorithmus soll stoppen, wenn eine Operation nicht ausführbar ist. Hier gibt es nur eine Operation, die gegebenenfalls nicht ausgeführt werden kann: Die Wahl eines Lösbarkeitsvektors  $\mathbf{z}^r \in \mathcal{Z}(I_\varepsilon(\mathbf{x}^r))$ . Folglich ergibt sich als Stopp-Regel: Wenn  $\mathcal{Z}(I_\varepsilon(\mathbf{x}^r)) = \emptyset$ , so stoppe.

**Satz 1** Wenn der zulässige Bereich von  $P$  beschränkt ist, dann ist für jedes  $\varepsilon > 0$  und jeden zulässigen Startpunkt  $\mathbf{x}^0$  das Algorithmenmodell  $\mathcal{A}(\mathbf{x}^0, \varepsilon)$  endlich.

**Beweis.** Angenommen, es gibt ein  $\varepsilon > 0$  und einen zulässigen Startpunkt  $\mathbf{x}^0$ , so daß  $\mathcal{A}(\mathbf{x}^0, \varepsilon)$  eine unendliche Punktfolge  $\{\mathbf{x}^r\}$  produziert. Da es nur endlich viele  $\varepsilon$ -aktive Indexmengen gibt, wird im Algorithmus auch nur endlich oft ein neuer Lösbarkeitsvektor  $\mathbf{z}$  gewählt. Folglich wird im Algorithmus zu jeder Indexmenge  $I \subseteq \{1, \dots, m\}$  höchstens ein Achsvektor berechnet. Somit existiert eine positive Zahl  $L$ , so daß

$$\mathbf{A}_i(\mathbf{y}^r + \mu^r \mathbf{c}) \leq L, \quad i = 1, \dots, m, \forall r.$$

Wegen der Beschränktheit des zulässigen Bereiches und

$$\mathbf{c}^T(\mathbf{x}^{r+1} - \mathbf{x}^r) \geq \lambda^r$$

muß die Schrittweitenfolge  $\{\lambda^r\}$  eine Nullfolge sein. Damit gibt es ein  $r_0$ , so daß für alle  $r \geq r_0$

$$\lambda^r L < \varepsilon \quad \forall r \geq r_0$$

ausfällt. Zu  $\lambda^r, r \geq r_0$  existiert ein  $i_* \notin I_\varepsilon(\mathbf{x}^r)$  mit

$$\lambda^r = \frac{b_{i_*} - \mathbf{A}_{i_*}^T \mathbf{x}^r}{\mathbf{A}_{i_*}^T (\mathbf{y}^r + \mu^r \mathbf{c})}.$$

Der Punkt  $\mathbf{x}^r$  hat von der Hyperebene  $H_{i_*}$  einen größeren Abstand als  $\varepsilon$ . Damit folgt

$$0 = \varrho(\mathbf{x}^{r+1}, H_{i_*}) = b_{i_*} - \mathbf{A}_{i_*}^T \mathbf{x}^r - \lambda^r \mathbf{A}_{i_*}^T (\mathbf{y}^r + \mu^r \mathbf{c}) > \varepsilon - \lambda^r L > 0.$$

Dieser Widerspruch beweist den Satz.  $\clubsuit$

Das Algorithmenmodell endet also stets nach endlich vielen Schritten - etwa  $r_*$  - mit einer Indexmenge  $I_* = I_\varepsilon(\mathbf{x}^{r_*})$ . Im Falle  $I_* = I_0(\mathbf{x}^{r_*})$  ist  $\mathbf{x}^{r_*}$  optimale Lösung der Aufgabe  $P$ . Dies kann durch eine hinreichend kleine Wahl von  $\varepsilon$  erreicht werden.

Wählt man als Startpunkt eine Ecke des Polyeders und als Vektor  $\mathbf{z}$  stets einen geeigneten Einheitsvektor, so erhält man als Iterationspunkte nur Ecken und man kann  $\varepsilon = 0$  wählen. Daher ist die Simplexmethode ein Spezialfall des Algorithmenmodells.

In gewissen Fällen kann man einen Achsvektor  $\mathbf{y}^r$  in  $\mathbf{x}^r$  besonders einfach angeben. Dazu nehmen wir an, daß der zulässige Bereich als konvexes Polyeder regulär ist (also jede  $s$ -dimensionale Randseite Durchschnitt von genau  $n - s$  linear unabhängigen Hyperebenen  $H_i$  ist) und die Niveauebenen der Zielfunktion nicht parallel zu einer Randseite des Polyeders verlaufen. Dann gibt es genau eine optimale Lösung (die natürlich Ecke des Polyeders ist) und für jede zulässige, nicht optimale Lösung  $\mathbf{x}$  von  $P$ , die nicht Eckpunkt des zulässigen Bereiches ist, sind die Vektoren  $\mathbf{c}, \mathbf{A}_i, i \in I_0(\mathbf{x})$  linear unabhängig. Für hinreichend kleines  $\varepsilon$  sind die Vektoren  $\mathbf{c}, \mathbf{A}_i, i \in I_\varepsilon(\mathbf{x})$  immer noch linear unabhängig. Damit können wir im Algorithmus Achsvektoren mit  $\mathbf{z}^T = (1, 1, \dots, 1)$  wählen, also den Achsvektor  $\mathbf{y}^r$  in  $\mathbf{x}^r$  als Lösung des Systems

$$\mathbf{A}_i^T \mathbf{y} = -1, i \in I_\varepsilon(\mathbf{x}^r), \mathbf{c}^T \mathbf{y} = 1, \mathbf{y} = \alpha_0 \mathbf{c} + \sum_{i \in I_\varepsilon(\mathbf{x}^r)} \alpha_i \mathbf{A}_i$$

bestimmen. Dieses System hat unter den gemachten Voraussetzungen genau eine Lösung. Die Koordinaten des gesuchten Achsvektors  $y^r$  bezüglich der Vektoren  $c, A_i, i \in I_\varepsilon(x)$  erhält man als Lösung von

$$\alpha_0 + \sum_{l \in I_\varepsilon(x^r)} \alpha_l c^T A_l = 1,$$

$$\alpha_0 A_i^T c + \sum_{l \in I_\varepsilon(x^r)} \alpha_l A_i^T A_l = -1, \quad i \in I_\varepsilon(x^r).$$

Das System hat eine symmetrische, positiv definite Koeffizientenmatrix, sofern  $\varepsilon$  hinreichend klein gewählt wurde. Ist die Koeffizientenmatrix nicht positiv definit, so hat man den Parameter  $\varepsilon$  zu reduzieren. Bei dieser algorithmischen Variante ist  $\mu^r = 1$  und der Punkt  $x^r + \lambda y^r$  hat von allen Hyperebenen  $H_i, i \in I_0(x^r)$  den Abstand  $\lambda$ . Die obige Kugel  $K_\lambda$  berührt daher alle diese Hyperebenen. Achsvektoren, die keine Wachstumsrichtungen der Zielfunktion sind, wurden in [1] betrachtet, jedoch ohne ein Lösungsverfahren aufzustellen.

Für Optimierungsaufgaben in Normalform

$$P_n : \quad \max \{ c^T x \mid Ax \leq b, x \geq 0 \}$$

reduziert sich das System zur Bestimmung eines Achsvektors für den betrachteten Fall eines regulären zulässigen Bereiches in folgender Weise:

Es sei

$$I_\varepsilon(x) = \{ i \mid A_i^T x \geq b_i - \varepsilon \}, \quad J_\varepsilon(x) = \{ j \mid x_j \leq \varepsilon \}, \quad (\varepsilon \geq 0).$$

Eine kleine Rechnung zeigt, daß der Achsvektor  $y$  die Komponenten

$$y_j = 1, \quad j \in J_\varepsilon(x), \quad y_j = \alpha_0 c_j + \sum_{l \in I_\varepsilon(x)} \alpha_l a_{lj}, \quad j \notin J_\varepsilon(x)$$

hat, wobei  $\alpha_0, \alpha_l, l \in I_\varepsilon(x)$  Lösung von

$$\alpha_0 \sum_{k \notin J_\varepsilon(x)} c_k c_k + \sum_{l \in I_\varepsilon(x)} \left( \sum_{k \notin J_\varepsilon(x)} c_k a_{lk} \right) \alpha_l = 1 - \sum_{k \in J_\varepsilon(x)} c_k$$

$$\alpha_0 \sum_{k \notin J_\varepsilon(x)} a_{ik} c_k + \sum_{l \in I_\varepsilon(x)} \left( \sum_{k \notin J_\varepsilon(x)} a_{ik} a_{lk} \right) \alpha_l = -(1 + \sum_{k \in J_\varepsilon(x)} a_{ik}), \quad i \in I_\varepsilon(x)$$

sein muß.

Falls es in der Optimierungsaufgabe Restriktionen in Gleichungsform gibt, ist  $z_i = 0$  in der Menge  $\Upsilon(I, z)$  für jede Gleichung  $A_i^T x = b_i$  zu wählen. Für die Effektivität der Methode ist es jedoch zweckmäßig, die Gleichungen über eine Variablenreduktion zu eliminieren.

## Literatur

- [1] Nožička, F.: Über eine Berührungshyperkugel-Methode in der Theorie der Programmierung. Mathematik und Kybernetik in der Ökonomie, Tagungsprotokoll, Teil II, S. 26-53, Berlin 1964

- [2] Stoer, J., and Ch. Witzgall: Convexity and Optimization in Finite Dimensions I. Springer-Verlag, Berlin, Heidelberg, New York 1970
- [3] Tschernikow, S. N.: Lineare Ungleichungen. Deutscher Verlag der Wissenschaften, Berlin 1971