

## Das zweistufige Problem der stochastischen linearen Programmierung

PETER KALL

Eingegangen am 20. August 1966

### Zusammenfassung

Es wird das aus der Literatur ([1], [4], [7], [9]) bekannte zweistufige Problem behandelt, wobei allerdings nicht nur die „rechten Seiten“ bzw. die Koeffizienten der Zielfunktion stochastische Variable sind. Zunächst wird das Problem neu formuliert wie in [1], was sich für den Beweis der in [7] und [4] eher umständlich bewiesenen Ungleichungen als nützlich erweist. Schließlich wird ein Verfahren der zulässigen Richtungen zur Lösung des Problems angegeben und seine Konvergenz bewiesen.

### I. Einleitung

Gegeben sei ein lineares Programm der folgenden Art: Minimiere

$$(1) \quad \left\{ \begin{array}{l} \text{bezüglich} \\ (\alpha) \quad A_1 x = b_1, \quad x \geq 0, \\ (\beta) \quad Ax + My = b, \quad y \geq 0. \end{array} \right. \quad \begin{array}{l} c'x + q'y \\ \\ \end{array}$$

Die Entscheidungsvariablen sind die Komponenten von  $x$  und  $y$ . Dabei sind  $A_1$  eine  $(m_1 \times n_1)$ -Matrix,  $A$  eine  $(m \times n_1)$ -Matrix,  $M$  eine  $(m \times n)$ -Matrix,  $b_1$  ein  $m_1$ -Vektor,  $b$  ein  $m$ -Vektor;  $c$  und  $x$  sind  $n_1$ -Vektoren und  $q$  und  $y$  sind  $n$ -Vektoren. Solange die Elemente von  $A_1, A, M, c, q, b_1, b$  fest gegeben sind, ist dieses Problem bekanntlich numerisch lösbar. Nimmt man nun jedoch an — was wir in dieser Arbeit tun wollen —, daß die Elemente von  $A, b, c$  zufällige Größen mit fest gegebenen Verteilungen sind, so verliert das Problem (1) offenbar seinen Sinn, wenn man die Realisationen der Zufallsvariablen noch nicht kennt. Wir brauchen also für diesen Fall eine neue Problemstellung, die wir aus der folgenden Vorschrift gewinnen:

1. Bevor die Realisationen der Zufallsvariablen bekannt sind, bestimmt man ein  $x$  so, daß es (1 $\alpha$ ) genügt.

2. Wenn man nachher die Realisationen  $\hat{A}, \hat{b}$  kennt, bestimmt man die Lösung des sog. Notprogramms (Programm der zweiten Stufe)

$$(2) \quad \left\{ \begin{array}{l} \min q'y, \\ My = \hat{b} - \hat{A}\hat{x}, \\ y \geq 0. \end{array} \right.$$

3. Unter 1. hat man  $\hat{x}$  so zu wählen, daß die erwarteten Gesamtkosten minimiert werden.

Diese Vorschrift ergibt nun die folgende Problemstellung:

$$(3) \quad \left\{ \begin{array}{l} \text{Minimiere} \\ \text{bezüglich} \end{array} \right. \begin{array}{l} \psi(x) = E\{c'x + \varphi(x, A, b)\} \\ A_1x = b_1, \\ x \geq 0, \end{array}$$

wobei  $\varphi(x, A, b)$  der Optimalwert des Notprogramms ist:

$$(4) \quad \left\{ \begin{array}{l} \text{Minimiere} \\ \text{bezüglich} \end{array} \right. \begin{array}{l} q'y \\ My = b - Ax, \\ y \geq 0. \end{array}$$

In [4] wird das Problem (3) ausführlich diskutiert. Gestützt auf die dort erzielten Resultate machen wir einige Voraussetzungen, deren Bedeutung wir anschließend kurz diskutieren:

**V 1:** Die Matrix  $M$  gestattet zu jeder zulässigen Lösung  $x$  von (3) und zu jeder Realisation von  $A, b$  eine zulässige Lösung von (4).

**V 2:** Das System

$$M'z \leq q$$

hat eine zulässige Lösung  $z$ .

**V 3:** Der Bereich  $\mathfrak{B} = \{x \mid A_1x = b_1, x \geq 0\}$  ist beschränkt (und nicht leer).

**V 4:** Das Wahrscheinlichkeitsmaß  $P(A, b)$  ist absolut stetig in bezug auf das Lebesgue-Maß  $\mu(A, b)$  über dem  $R_{m(n_1+1)}$ , und der Erwartungswert  $E\{(A, b, c)\} = (\bar{A}, \bar{b}, \bar{c})$  existiert. [Der erste Teil dieser Voraussetzung ist erfüllt, wenn die gemeinsame Verteilungsfunktion  $F(A, b)$  absolut stetig ist.]

Bemerkungen zu diesen Voraussetzungen:

*Zu V 1:* In [4] werden hinreichende Bedingungen für  $M$  angegeben dafür, daß V 1 erfüllt ist. Diese Bedingungen lassen sich dadurch nachprüfen, daß man die Verträglichkeit eines linearen Restriktionensystems, in dem z. T. streng positive Variable verlangt werden, feststellt. Falls die Menge  $\{b - Ax \mid x \in \mathfrak{B}, (A, b) \text{ eine mögliche Realisation}\}$  eine Kugel um den Ursprung enthält, sind diese Bedingungen auch notwendig.

*Zu V 2:* Diese Voraussetzung garantiert (zusammen mit V 1), daß das Notprogramm (4) zu jedem  $x \in \mathfrak{B}$  und zu jeder (endlichen) Realisation von  $(A, b)$  eine endliche Lösung hat. Ist diese Voraussetzung verletzt, so folgt aus V 1 und dem Dualitätssatz, daß die Lösung des Notprogramms stets unendlich ist, womit dann die Problemstellung (3) sinnlos wird.

*Zu V 3:* Diese Bedingung garantiert die Existenz einer Lösung von (3). Vom theoretischen Standpunkt aus erscheint diese Voraussetzung sicher als sehr einschränkend; es ist jedoch zu beachten, daß es in den praktisch auftretenden Programmierungsproblemen im allgemeinen möglich ist, z. B. eine obere Schranke der Summe der Variablen anzugeben, womit V 3 bereits erfüllt ist.

Zu V 4: Diese Voraussetzung garantiert — zusammen mit V 1 und V 2 —, daß in jedem zulässigen  $x$  die Zielfunktion  $\varphi(x)$  endlich ist und einen stetigen Gradienten  $g(x)$  besitzt, wie in [4] gezeigt wird. Ferner ist  $\varphi(x)$  konvex.

## II. Eine äquivalente Formulierung des Problems

Die Problemstellung (3) erscheint etwas unhandlich, da die Funktion  $\varphi(x, A, b)$  nicht explizit gegeben ist. Für den Spezialfall, daß nur  $b$  zufällige Variable enthält und  $M = (I, -I)$  ist ( $I$  ist die Einheitsmatrix), hat WETS die Funktion  $\varphi(x, A, b)$  in [9] explizit angegeben, wobei er den Dualitätssatz benutzt. Für den weiteren Sonderfall, daß nur  $b$  stochastische Größen enthält,  $M$  aber nur der Bedingung V 1 unterworfen ist, haben CHARNES, COOPER und THOMPSON in [1] eine andere Formulierung gefunden, so daß  $\varphi(x, A, b)$  explizit gegeben ist. In dem in dieser Arbeit behandelten allgemeinen Fall kann man  $\varphi(x, A, b)$  mit Hilfe des Simplex-Kriteriums explizit darstellen, wie das in [4] geschehen ist. Für die numerische Auswertung scheint jedoch das Vorgehen von CHARNES, COOPER und THOMPSON geeigneter zu sein, das sich auf den allgemeinen Fall ohne weiteres übertragen läßt. Dazu benötigen wir den Begriff der verallgemeinerten Inversen einer Matrix.

**Definition.** Sei  $M$  eine  $(m \times n)$ -Matrix. Eine  $(n \times m)$ -Matrix  $M^+$  heißt die verallgemeinerte Inverse von  $M$ , wenn  $M^+$  den folgenden Bedingungen genügt:

$$(5) \quad \left\{ \begin{array}{l} (\alpha) \quad M M^+ M = M, \\ (\beta) \quad M^+ M M^+ = M^+, \\ (\gamma) \quad (M M^+)' = M M^+, \\ (\delta) \quad (M^+ M)' = M^+ M. \end{array} \right.$$

Aus der Literatur (s. [1]) ist bekannt, daß das System (5) zu jeder Matrix  $M$  genau eine Lösung besitzt. Falls  $m = n$  und  $M$  nicht singulär ist, gilt demzufolge  $M^+ = M^{-1}$ . Im weiteren ist der folgende leicht beweisbare Satz bekannt:

**Lemma 1.** *Das lineare Gleichungssystem  $Mx = d$  hat genau dann eine Lösung, wenn  $M M^+ d = d$  gilt. Dann lautet die Lösung  $x = M^+ d + (I - M^+ M)y$  mit beliebigem  $y \in R_n$ .*

Betrachten wir nun das Problem (4):

$$(4) \quad \left\{ \begin{array}{l} \varphi(x, A, b) = \min q' y \\ \text{bezüglich} \quad M y = b - A x, \\ y \geq 0. \end{array} \right.$$

Nach Lemma 1 gilt für jede zulässige Lösung  $y$  von (4):

$$y = M^+(b - Ax) + (I - M^+ M)z \geq 0, \quad z \in R_n,$$

sofern

$$M M^+(b - Ax) = b - Ax,$$

was durch V 1 sichergestellt ist.

Damit geht (4) über in

$$(6) \quad \left\{ \begin{array}{l} \varphi(x, A, b) = \min q' M^+(b - Ax) + q'(I - M^+ M)z \\ \text{bezüglich} \quad (I - M^+ M)z \geq M^+(Ax - b), \quad z \in R_n. \end{array} \right.$$

Nach dem Dualitätssatz gilt dann aber

$$(7) \quad \left\{ \begin{array}{l} \varphi(x, A, b) = \max w' M^+(Ax - b) + q' M^+(b - Ax) \\ \text{bezüglich } (I - M^+M)'w = (I - M^+M)'q, \\ w \geq 0, \quad w \in R_n. \end{array} \right.$$

Nach (5  $\delta$ ) ist das identisch mit

$$(8) \quad \left\{ \begin{array}{l} \varphi(x, A, b) = \max w' M^+(Ax - b) + q' M^+(b - Ax) \\ \text{bezüglich } (I - M^+M)w = (I - M^+M)q, \\ w \geq 0. \end{array} \right.$$

Nach unseren Voraussetzungen ist  $\varphi(x, A, b)$  immer endlich, d. h. (8) hat stets eine endliche Lösung, die bekanntlich immer durch einen Eckpunkt des zulässigen Bereiches angegeben werden kann. Seien daher  $w_1, \dots, w_k$  die Eckpunkte des zulässigen Bereiches von (8), die man bekanntlich berechnen kann. Dann ist

$$(9) \quad \varphi(x, A, b) = \max_{1 \leq i \leq k} w'_i M^+(Ax - b) + q' M^+(b - Ax).$$

Damit haben wir die gewünschte explizite Darstellung von  $\varphi(x, A, b)$ . Danach geht (3) über in:

$$(10) \quad \left\{ \begin{array}{l} \text{Minimiere} \\ \psi(x) = \bar{c}'x + q' M^+(b - Ax) + \int \max_{1 \leq i \leq k} w'_i M^+(Ax - b) dP(A, b) \\ \text{bezüglich } A_1 x = b_1, \\ x \geq 0. \end{array} \right.$$

Sei  $w_0$  die jeweils optimale Ecke von (8), die also von  $A, b$  und  $x$  abhängt. Wie in [4] gezeigt wird, hat  $\psi(x)$  unter unseren Voraussetzungen einen stetigen Gradienten, der nach (10) gegeben ist durch

$$(11) \quad g(x) = \bar{c} - \bar{A}' M^{+'} q + \int A' M^{+'} w_0 dP(A, b).$$

Hier ist natürlich das Integral über die einzelnen Komponenten der vektorwertigen Funktion  $A' M^{+'} w_0$  von  $(A, b)$  zu berechnen.

### III. Ungleichungen

Die Darstellung (10) hat, abgesehen davon, daß sie die Funktion  $\varphi(x, A, b)$  explizit angibt, noch den Vorteil, daß sich damit die in [4] und [7] mit ziemlich großem Aufwand hergeleiteten Ungleichungen wesentlich einfacher beweisen lassen. Dazu nehmen wir an, daß in (3) und (4) die Zufallsvariablen  $A, b, c$  durch ihre Erwartungswerte ersetzt werden. Dann geht (3) über in das lineare Programm

$$(12) \quad \left\{ \begin{array}{l} \text{Minimiere } \gamma = \bar{c}'x + q'y \\ \text{bezüglich } \quad A_1 x \quad = b_1, \\ \quad \quad \quad \bar{A}x + My = \bar{b}, \\ \quad \quad \quad x \geq 0, \quad y \geq 0. \end{array} \right.$$

Wir nehmen nun an, daß (12) einen endlichen Optimalwert  $\gamma_0$  besitzt. Folglich ist, wenn  $x_0 \in \mathfrak{B}$  diesen Optimalwert liefert, nach (10)

$$(13) \quad \gamma_0 = \bar{c}'x_0 + q' M^+(b - Ax_0) + w'_j M^+(\bar{A}x_0 - \bar{b}),$$

wobei

$$(14) \quad w'_j M^+(\bar{A}x - \bar{b}) = \max_{1 \leq i \leq k} w'_i M^+(\bar{A}x - \bar{b}).$$

Nun gilt der folgende

**Satz 1.** *Unter unseren Voraussetzungen gelten die Ungleichungen*

$$\gamma_0 \leq \min_{x \in \mathfrak{B}} \psi(x) \leq \psi(x_0).$$

*Beweis.* Offenbar gilt

$$(15) \quad w'_j M^+(Ax - b) \leq \max_{1 \leq i \leq k} w'_i M^+(Ax - b)$$

für alle  $x \in \mathfrak{B}$  und alle möglichen  $(A, b)$ .

Damit gilt nach (13) und (10)

$$(16) \quad \begin{cases} \gamma_0 = \bar{c}'x_0 + q' M^+(\bar{b} - \bar{A}x_0) + \int w'_j M^+(Ax_0 - b) dP(A, b) \\ \leq \bar{c}'x + q' M^+(\bar{b} - \bar{A}x) + \int w'_j M^+(Ax - b) dP(A, b) \\ \leq \bar{c}'x + q' M^+(\bar{b} - \bar{A}x) + \int \max_{1 \leq i \leq k} w'_i M^+(Ax - b) dP(A, b) \\ = \psi(x) \end{cases}$$

für alle  $x \in \mathfrak{B}$ .

Folglich gilt auch für einen Optimalpunkt  $\bar{x}$  von (10)

$$(17) \quad \gamma_0 \leq \psi(\bar{x}) = \min_{x \in \mathfrak{B}} \psi(x).$$

Damit ist der Satz bewiesen, da die zweite Ungleichung trivial ist.

#### IV. Lösungsverfahren

Wie wir gesehen haben, ist  $\psi(x)$  konvex und stetig differenzierbar nach allen Komponenten von  $x$ . Demzufolge bietet sich zur Lösung von (10) ein Gradientenverfahren geradezu an. Das hier beschriebene Verfahren gehört in die Klasse der von ZOUTENDIJK in [10] behandelten „Methoden der zulässigen Richtungen“ und arbeitet folgendermaßen:

1. *Schritt.* Bestimmung einer zulässigen Lösung  $x_0$  von

$$A_1 x = b_1, \quad x \geq 0,$$

d.h.  $x_0 \in \mathfrak{B}$ . Initialisierung  $k = 0$ .

2. *Schritt.* Bestimme

$$(18) \quad \begin{cases} \gamma_k = \text{Min } x'_k g(x_k) \\ \text{bezüglich } A_1 x = b_1, \quad x \geq 0. \end{cases}$$

3. *Schritt.* Ist  $\gamma_k = x'_k g(x_k)$ , so ist  $x_k$  eine Optimallösung von (10). Ist  $\gamma_k < x'_k g(x_k)$ , so sei  $\hat{x}_k$  eine Lösung von (18).

Dann folgt der

4. *Schritt.* Bestimme

$$(19) \quad \begin{cases} \text{Min } \psi(\lambda x_k + (1 - \lambda)\hat{x}_k) \\ \text{bezüglich } 0 \leq \lambda \leq 1. \end{cases}$$

Die Lösung sei  $\lambda_k$ . Dann setzt man

$$(20) \quad x_{k+1} = \lambda_k x_k + (1 - \lambda_k)\hat{x}_k \quad \text{und} \quad k = k + 1$$

und beginnt wieder mit dem 2. Schritt.

Wir haben zu zeigen, daß man mit diesem Verfahren eine Lösung von (10) bestimmen kann.

**Satz 2.** Die Zahlenfolge  $\psi(x_k)$  konvergiert gegen den Optimalwert der Zielfunktion von (10). Die Folge  $\{x_k\}$  enthält eine konvergente Teilfolge, und jede konvergente Teilfolge von  $\{x_k\}$  konvergiert gegen eine Lösung von (10).

*Beweis.* Es gilt  $x_0 \in \mathfrak{B}$ . Da  $\mathfrak{B}$  konvex ist, ist nach (18), (19) und (20) mit  $x_k \in \mathfrak{B}$  auch  $x_{k+1} \in \mathfrak{B}$ ,  $k = 0, 1, \dots$ . Da  $\mathfrak{B}$  kompakt und  $\psi(x)$  stetig ist, hat  $\psi(x)$  auf  $\mathfrak{B}$  ein Minimum  $\psi_0$ . Folglich gilt

$$(21) \quad \psi_0 \leq \psi(x_k), \quad k = 0, 1, \dots$$

Ist das Verfahren endlich, so gilt nach dem 3. Schritt für ein  $k$

$$x'_k g(x_k) \leq x' g(x_k) \quad \text{für alle } x \in \mathfrak{B}$$

und daher wegen der Konvexität von  $\psi(x)$

$$0 = (x'_k - x'_k) g(x_k) \leq (x' - x'_k) g(x_k) \leq \psi(x) - \psi(x_k) \quad \text{für alle } x \in \mathfrak{B}.$$

Also ist  $x_k$  optimal. Wir brauchen daher nur noch den unendlichen Fall zu betrachten. Wir nehmen nun an, daß für ein  $k$  die folgende Ungleichung gilt, wobei natürlich  $x_{k+1} \neq x_k$  ist:

$$(22) \quad \psi(x_{k+1}) \geq \psi(x_k).$$

Nach dem 4. Schritt gilt

$$(23) \quad \psi(x_{k+1}) \leq \psi(x) \quad \text{für alle } x \in [x_k, \hat{x}_k].$$

Aus (22) und (23) folgt

$$(24) \quad \psi(x) - \psi(x_k) \geq 0 \quad \text{für alle } x \in [x_k, \hat{x}_k].$$

Daher gilt für die Richtungsableitung von  $\psi(x)$  in  $x_k$  in der Richtung  $x - x_k = \gamma(\hat{x}_k - x_k)$ ,  $0 \leq \gamma \leq 1$ ,

$$(25) \quad \lim_{\substack{x \rightarrow x_k \\ x \in [x_k, \hat{x}_k]}} \frac{\psi(x) - \psi(x_k)}{|x - x_k|} = \frac{(\hat{x}_k - x_k)'}{|\hat{x}_k - x_k|} g(x_k) \geq 0.$$

Da wir angenommen haben, daß das Verfahren unendlich ist, muß aber im 3. Schritt

$$\hat{x}'_k g(x_k) < x'_k g(x_k)$$

und folglich

$$(26) \quad \frac{(\hat{x}_k - x_k)'}{|\hat{x}_k - x_k|} g(x_k) < 0$$

gelten. Aus dem Widerspruch zwischen (25) und (26) folgt, daß (22) falsch ist. Folglich gilt

$$(27) \quad \psi(x_{k+1}) < \psi(x_k) \quad \text{für } k = 0, 1, \dots$$

Aus (21) und (27) folgt die Konvergenz der Folge  $\{\psi(x_k)\}$ .

Wir nehmen nun an, daß

$$(28) \quad \lim_{k \rightarrow \infty} \psi(x_k) = \psi_1 > \psi_0$$

gilt und wollen diese Annahme zum Widerspruch führen. Sei  $\bar{x}$  eine optimale zulässige Lösung von (10), d. h.  $\psi(\bar{x}) = \psi_0$ . Da  $\mathfrak{B}$  kompakt ist, hat die Folge  $\{x_k\}$  mindestens einen Häufungspunkt  $\tilde{x} \in \mathfrak{B}$ . Folglich gibt es eine Teilfolge  $\{x_{k_i}\}$  derart, daß

$$\lim_{i \rightarrow \infty} (x_{k_i}) = \tilde{x} \quad \text{und} \quad \lim_{i \rightarrow \infty} \psi(x_{k_i}) = \psi(\tilde{x}) = \psi_1.$$

Zur Abkürzung setzen wir

$$(29) \quad q(x; y) = \frac{(y-x)'}{|y-x|} g(x); \quad x, y \in \mathfrak{B}; \quad x \neq y.$$

In  $\tilde{x}$  sei

$$(30) \quad (\hat{x}_i - \tilde{x})' g(\tilde{x}) = \text{Min}_{x \in \mathfrak{B}} (x - \tilde{x})' g(\tilde{x}) \quad \text{für} \quad i = 1, \dots, r,$$

wobei die  $\hat{x}_i$  alle die Eckpunkte von  $\mathfrak{B}$  sind, die in (30) optimal sind, d. h.  $r \geq 1$ . Dann gilt wegen (28) und der Konvexität von  $\psi(x)$  für  $i = 1, \dots, r$

$$(31) \quad (\hat{x}_i - \tilde{x})' g(\tilde{x}) \leq (\bar{x} - \tilde{x})' g(\tilde{x}) \leq \psi(\bar{x}) - \psi(\tilde{x}) = \psi_0 - \psi_1 < 0$$

und folglich

$$(32) \quad q(\tilde{x}; \hat{x}_i) \leq \beta < 0 \quad \text{für} \quad i = 1, \dots, r.$$

Ferner gilt für alle übrigen Eckpunkte  $\hat{x}_j$  des zulässigen Bereiches  $\mathfrak{B}$

$$(33) \quad (\hat{x}_j - \tilde{x})' g(\tilde{x}) \geq (\hat{x}_i - \tilde{x})' g(\tilde{x}) + \gamma; \quad i \leq r, \quad j > r, \quad \gamma > 0.$$

Nach (32) und (33) sind die folgenden Mengen  $U_1, U_2$  wegen der Stetigkeit von  $g(x)$  auf  $\mathfrak{B}$  Umgebungen von  $\tilde{x}$  bezüglich der relativen natürlichen Topologie auf  $\mathfrak{B}$ :

$$U_1 = \left\{ x \mid (\hat{x}_j - x)' g(x) \geq (\hat{x}_i - x)' g(x) + \frac{\gamma}{2}; \quad i \leq r, \quad j > r; \quad x \in \mathfrak{B} \right\},$$

$$U_2 = \left\{ x \mid q(x; \hat{x}_i) \leq \frac{\beta}{2}, \quad i \leq r; \quad x \in \mathfrak{B} \right\}.$$

Wegen (31) muß

$$(34) \quad |\hat{x}_i - \tilde{x}| \geq \delta > 0, \quad i \leq r$$

gelten. Dann ist auch

$$U_3 = \left\{ x \mid |\hat{x}_i - x| \geq \frac{\delta}{2}, \quad i \leq r; \quad x \in \mathfrak{B} \right\}$$

eine Umgebung von  $\tilde{x}$  in bezug auf die oben erwähnte Topologie.

Dann ist aber auch

$$U_4 = U_1 \cap U_2 \cap U_3$$

eine solche Umgebung von  $\tilde{x}$ , und es existiert daher eine Kugel  $K(\tilde{x}, \varepsilon)$  mit dem Zentrum  $\tilde{x}$  und dem Radius  $\varepsilon > 0$  derart, daß

$$K_{\mathfrak{B}} = \mathfrak{B} \cap K(\tilde{x}, \varepsilon) \subset U_4.$$

Nun gilt

$$(35) \quad \hat{x}_i \notin K_{\mathfrak{B}}, \quad i \leq r, \quad \text{da} \quad K_{\mathfrak{B}} \subset U_4 \subset U_3,$$

$$(36) \quad q(x; \hat{x}_i) \leq \frac{\beta}{2}, \quad i \leq r, \quad x \in K_{\mathfrak{B}}, \quad \text{da} \quad K_{\mathfrak{B}} \subset U_4 \subset U_2.$$

Weiter betrachten wir noch die Kugel  $K(\tilde{x}, \varepsilon/2)$  und definieren die Umgebung von  $\tilde{x}$  in  $\mathfrak{B}$

$$K_{\mathfrak{B}}^1 = \mathfrak{B} \cap K\left(\tilde{x}, \frac{\varepsilon}{2}\right) \subset K_{\mathfrak{B}}.$$

Offenbar gelten in  $K_{\mathfrak{B}}^1$  ebenfalls (35) und (36) und ferner

$$(37) \quad |x - x^0| \geq \frac{\varepsilon}{2}, \quad \text{falls } x^0 \in K_{\mathfrak{B}}^1 \quad \text{und} \quad x \in \overline{K_{\mathfrak{B}}^c},$$

wobei  $\overline{K_{\mathfrak{B}}^c}$  die abgeschlossene Hülle des Komplementes von  $K_{\mathfrak{B}}$  bezüglich  $\mathfrak{B}$  ist. Da  $K_{\mathfrak{B}}^1$  eine Umgebung von  $\tilde{x}$  bezüglich der relativen natürlichen Topologie auf  $\mathfrak{B}$  ist, existiert ein  $j_0$  derart, daß  $x_{k_j} \in K_{\mathfrak{B}}^1$  für alle  $j > j_0$ , d. h.  $K_{\mathfrak{B}}^1$  enthält unendlich viele Punkte  $x_{k_j}$ . Da  $K_{\mathfrak{B}}^1 \subset U_1$ , wird das im 2. Schritt zu bestimmende Minimum für jedes  $x_{k_j}$ ,  $j > j_0$ , in einem der Eckpunkte  $\hat{x}_i$ ,  $i \leq r$ , angenommen. Sei zu einem derartigen  $x_{k_j}$  ein optimaler Eckpunkt  $\hat{x}_i$  bestimmt. Dann haben wir im 4. Schritt einen Punkt  $x_{k_{j+1}} \in S = \{x \mid x = x_{k_j} + \lambda(\hat{x}_i - x_{k_j}), 0 \leq \lambda \leq 1\}$  so zu bestimmen, daß  $\psi(x)$  auf  $S$  minimal wird. Nach (35) und (36) gilt  $x_{k_{j+1}} \notin K_{\mathfrak{B}}$ . Sei  $x_{k_j, R} \in K_{\mathfrak{B}}$  der Schnittpunkt von  $S$  mit dem Rand von  $K(\tilde{x}, \varepsilon)$ . Dann gilt nach (37)

$$(38) \quad |x_{k_{j+1}} - x_{k_j}| \geq |x_{k_j, R} - x_{k_j}| \geq \frac{\varepsilon}{2}.$$

Damit folgt wegen (36) nach dem Mittelwertsatz mit  $0 \leq \Theta \leq 1$

$$(39) \quad \begin{aligned} \psi(x_{k_{j+1}}) &\leq \psi(x_{k_j, R}) = \psi(x_{k_j}) + |x_{k_j, R} - x_{k_j}| q(\Theta x_{k_j} + (1 - \Theta)x_{k_j, R}; \hat{x}_i) \leq \\ &\leq \psi(x_{k_j}) + \frac{\varepsilon}{2} \frac{\beta}{2}, \end{aligned}$$

d. h. die Zielfunktion nimmt in jedem von einem  $x_{k_j} \in K_{\mathfrak{B}}^1$  ausgehenden Iterationsschritt mindestens um  $\frac{1}{4}|\varepsilon\beta|$  ab. Da  $K_{\mathfrak{B}}^1$  unendlich viele  $x_{k_j}$  enthält, würde deshalb  $\psi(x)$  auf  $\mathfrak{B}$  gegen  $-\infty$  streben. Aus diesem Widerspruch zu (21) folgt, daß die Annahme (28) falsch war. Also gilt

$$(40) \quad \lim_{k \rightarrow \infty} \psi(x_k) = \psi_0.$$

Wegen der Kompaktheit von  $\mathfrak{B}$  und der Stetigkeit von  $\psi(x)$  ist dann der Rest des Satzes trivial.

Das oben angegebene Verfahren, das WETS in [9] zur Lösung seines bereits erwähnten speziellen Problems benutzt hat, führt theoretisch ganz allgemein bei der Minimierung einer konvexen, stetig differenzierbaren Funktion über einem beschränkten konvexen Polyeder zur Lösung. Praktisch können jedoch im 4. Schritt bei der Minimierung von  $\psi(\lambda x_k + (1 - \lambda)\hat{x}_k)$ ,  $0 \leq \lambda \leq 1$ , erhebliche Schwierigkeiten auftreten, obwohl es sich ja hier nur noch um die Minimierung einer Funktion von einer Variablen über einem beschränkten abgeschlossenen Intervall handelt. Bei WETS erscheint das Verfahren jedoch deshalb anwendbar, weil sich in seinem Fall aufgrund der speziellen Struktur der Matrix  $M$  — und da nur  $b$  stochastisch ist —  $\psi(x)$  und  $g(x)$  mit Hilfe einfacher Integrale von linearen Funktionen von  $(A, b)$  darstellen lassen. Betrachtet man jedoch das allgemeine Problem (10), so sieht man sofort, daß  $\psi(x)$  und  $g(x)$  hier komplizierter dargestellt werden müssen. Nun ist es aber ohne weiteres möglich, das obige allgemeine Lö-



sungsverfahren auf verschiedene Arten derart abzuändern, daß die strenge Lösung der im 4. Schritt gestellten Minimierungsaufgabe umgangen wird. Eine Möglichkeit, den 4. Schritt des Verfahrens zu ändern, wollen wir hier angeben:

4. Schritt. Setze  $x = \hat{x}_k$  und

$$(41) \quad y = x_k + \frac{1}{2}(x - x_k).$$

Ist  $\psi(y) \geq \psi(x)$ , so setze  $x_{k+1} = x$  und beginne wieder mit dem 2. Schritt.

Ist  $\psi(y) < \psi(x)$ , so setze  $x = y$  und kehre zurück nach (41).

Wir haben zu zeigen, daß auch dieses abgeänderte Verfahren zum Ziel führt. Dazu beweisen wir zunächst das

**Lemma 2.** *Bei dem abgeänderten Verfahren erhält man in jedem Schritt eine Verbesserung des Wertes der Zielfunktion.*

*Beweis.* Nach (27) existiert ein  $\lambda$ ,  $0 \leq \lambda < 1$ , derart, daß für  $y = \lambda x_k + (1 - \lambda)\hat{x}_k = (1 - \lambda)(\hat{x}_k - x_k) + x_k$

$$\psi(y) < \psi(x_k)$$

gilt. Dann wählen wir eine natürliche Zahl  $n$  so, daß

$$\frac{1}{2^n} \leq 1 - \lambda$$

gilt und setzen

$$x = x_k + \frac{1}{2^n}(\hat{x}_k - x_k) = \mu x_k + (1 - \mu)y$$

mit

$$\mu = 1 - \frac{1}{2^n(1 - \lambda)} < 1.$$

Dann gilt nach unserer Vorschrift im 4. Schritt und wegen der Konvexität von  $\psi(x)$

$$\psi(x_{k+1}) \leq \psi(x) \leq \mu \psi(x_k) + (1 - \mu)\psi(y) < \psi(x_k);$$

und das war zu zeigen.

Schließlich beweisen wir noch den

**Satz 3.** *Im abgeänderten Verfahren konvergiert die Folge  $\{\psi(x_k)\}$  gegen den Optimalwert der Zielfunktion von (10). Eine Teilfolge der  $x_k$  konvergiert gegen eine Optimallösung.*

*Beweis.* Lemma 2 zeigt, daß die Folge  $\{\psi(x_k)\}$  monoton abnimmt. Da  $\psi(x_k) \geq \psi_0$  gilt, folgt die Konvergenz der Folge. Wir haben wieder zu zeigen, daß  $\lim_{k \rightarrow \infty} \psi(x_k) = \psi_0$  gilt, und gehen dabei analog wie im Beweis von Satz 2 vor, d. h. wir nehmen  $\lim_{k \rightarrow \infty} \psi(x_k) > \psi_0$  an. Die im Beweis von Satz 2 definierten Mengen und Größen werden auch hier benutzt. Ferner ist

$$(42) \quad \alpha = \max \{ |\hat{x}_i - x| \mid x \in K_{\mathfrak{B}}^1, i \leq r \} \geq \frac{\delta}{2} > 0,$$

da  $K_{\mathfrak{B}}^1 \subset U_3$ . Wir wählen eine natürliche Zahl  $n_0$  so groß, daß

$$(43) \quad 2^{n_0-1} \geq \frac{\alpha}{\varepsilon}.$$

Dann gilt

$$(44) \quad \frac{1}{2^{n_0}} |\hat{x}_i - x| \leq \frac{1}{2^{n_0}} \alpha \leq \frac{\varepsilon}{2}, \quad i \leq r, \quad x \in K_{\mathfrak{B}}^1.$$

Andererseits gilt, da  $K_{\mathfrak{B}}^1 \subset U_3$ ,

$$(45) \quad \frac{\varepsilon}{2} \geq \frac{1}{2^{n_0}} |\hat{x}_i - x| \geq \frac{\delta}{2^{n_0+1}} = 3\rho > 0 \quad \text{für } x \in K_{\mathfrak{B}}^1.$$

Nun betrachten wir die Kugeln  $K(\tilde{x}, \rho)$  und  $K(\tilde{x}, 2\rho)$  sowie die Menge  $K_{\mathfrak{B}}^0 = \mathfrak{B} \cap K(\tilde{x}, \rho)$ , die offenbar wieder eine Umgebung von  $\tilde{x}$  bezüglich der relativen natürlichen Topologie auf  $\mathfrak{B}$  ist und daher, da  $\tilde{x}$  Häufungspunkt der  $x_k$  ist, unendlich viele  $x_k$  enthält. Für jedes  $x \in K_{\mathfrak{B}}^0$  gilt mit  $i \leq r$ :

$$(46) \quad \left\{ \begin{aligned} 2\rho &\leq \frac{1}{2^{n_0}} |\hat{x}_i - x| - \rho \leq \frac{1}{2^{n_0}} |\hat{x}_i - x| - |x - \tilde{x}| \\ &\leq \left| \frac{1}{2^{n_0}} (\hat{x}_i - x) - (\tilde{x} - x) \right| \\ &= \left| x + \frac{1}{2^{n_0}} (\hat{x}_i - x) - \tilde{x} \right| \\ &\leq |x - \tilde{x}| + \frac{1}{2^{n_0}} |\hat{x}_i - x| \\ &\leq \rho + \frac{\varepsilon}{2} \leq \varepsilon. \end{aligned} \right.$$

Diese Abschätzung zeigt, daß wir im 4. Schritt des Verfahrens, ausgehend von  $x_k \in K_{\mathfrak{B}}^0$ , jedenfalls aus der Kugel  $K(\tilde{x}, 2\rho)$  hinauskommen. Folglich gilt für  $x_k \in K_{\mathfrak{B}}^0$ , wenn  $y = x_k + (1/2^{n_0})(\hat{x}_i - x_k)$  gesetzt wird,  $i \leq r$  und  $0 \leq \Theta \leq 1$ , unter Beachtung von (45)

$$(47) \quad \begin{aligned} \psi(x_{k+1}) &\leq \psi(y) = \psi(x_k) + \frac{1}{2^{n_0}} |\hat{x}_i - x_k| q(\Theta x_k + (1 - \Theta)y; \hat{x}_i) \leq \\ &\leq \psi(x_k) + 3\rho \frac{\beta}{2}, \end{aligned}$$

da nach (46)

$$\begin{aligned} |\Theta x_k + (1 - \Theta)y - \tilde{x}| &= |\Theta(x_k - \tilde{x}) + (1 - \Theta)(y - \tilde{x})| \\ &\leq \Theta |x_k - \tilde{x}| + (1 - \Theta) |y - \tilde{x}| \\ &\leq \Theta \rho + (1 - \Theta) \varepsilon \leq \varepsilon \end{aligned}$$

und daher  $(\Theta x_k + (1 - \Theta)y) \in K_{\mathfrak{B}} \subset U_2$ .

Aus (47) würde wieder folgen, daß  $\psi(x_k)$  gegen  $-\infty$  strebt, womit unsere Annahme, daß  $\lim_{k \rightarrow \infty} \psi(x_k) > \psi_0$  sei, sich wieder als falsch erweist. Der Rest des Satzes ist nun wieder trivial.

Numerisch kann man das Verfahren noch weiter vereinfachen, indem man z. B. im 4. Schritt nur so lange die Strecke  $[x_k, \hat{x}_k]$  halbiert, bis man einen kleineren Wert als  $\psi(x_k)$  gefunden hat, und dann diese Stelle als  $x_{k+1}$  bezeichnet. Man muß aber bei allen derartigen Veränderungen des Verfahrens selbstverständlich damit rechnen, daß die Konvergenzrate gegenüber dem allgemeinen Verfahren abnimmt.

## V. Bemerkungen zum 2. Schritt

Wie wir gesehen haben, hat man in dem hier besprochenen Verfahren jeweils im zweiten Schritt ein lineares Programm zu lösen. Dabei wird sich im allgemeinen

der Gradient der Zielfunktion  $g(x_k)$  von Mal zu Mal ändern. Eine zulässige Lösung  $x_k$  ist bekannt. Ferner ist auch eine zulässige Basislösung  $\hat{x}_{k-1}$  bekannt. Damit kann man natürlich ohne weiteres mit dem Simplexverfahren beginnen. Allerdings wird man dann u. U. relativ viele Transformationen durchführen müssen, da  $g(x_k)$  in vielen Fällen nach dem 4. Schritt orthogonal ist zu  $[x_{k-1}, \hat{x}_{k-1}]$ . Deshalb wollen wir uns noch einmal kurz mit der im II. Abschnitt eingeführten verallgemeinerten Inversen einer Matrix befassen und eine aus der Literatur bekannte Eigenschaft besonders erwähnen.

Nach (5δ) ist die Matrix  $A_1^+ A_1$  symmetrisch und, wie man leicht zeigen kann, vom Rang  $r = \text{Rang}(A_1)$ . Ferner gilt nach (5)

$$\begin{aligned} A_1^+ A_1 A_1^+ A_1 &= A_1^+ A_1, \\ (I - A_1^+ A_1)(I - A_1^+ A_1) &= I - A_1^+ A_1 - A_1^+ A_1 + A_1^+ A_1 A_1^+ A_1 = I - A_1^+ A_1, \\ A_1^+ A_1 (I - A_1^+ A_1) &= A_1^+ A_1 - A_1^+ A_1 A_1^+ A_1 = A_1^+ A_1 - A_1^+ A_1 = 0, \\ A_1^+ A_1 + (I - A_1^+ A_1) &= I, \\ A_1 (I - A_1^+ A_1) &= A_1 - A_1 A_1^+ A_1 = A_1 - A_1 = 0. \end{aligned}$$

Aus diesen einfachen Beziehungen folgt, daß  $A_1^+ A_1$  und  $I - A_1^+ A_1$  Projektionsmatrizen sind. Dabei projiziert die Matrix  $I - A_1^+ A_1$  alle Vektoren des  $R_n$  auf den Nullraum der Matrix  $A_1$ . Diese Tatsache nutzen wir nun aus, um — ähnlich dem Vorgehen von H. P. KÜNZLI in [6] — längs der Projektion des negativen Gradienten in den linearen Raum  $\{x \mid A_1 x = b_1\}$  so weit wie möglich durch den zulässigen Bereich vorzudringen. Genau gesagt gehen wir folgendermaßen vor:

Wir kennen im 2. Schritt ein  $x_k \in \mathfrak{B}$  und daher auch den Gradienten der Zielfunktion von (18),  $g(x_k)$ . Dann setzen wir

$$\begin{aligned} x &= x_k, \\ s &= -(I - A_1^+ A_1)g(x_k). \end{aligned}$$

$s$  ist nach dem oben Gesagten die Projektion des negativen Gradienten in den Nullraum  $\{z \mid A_1 z = 0\}$ . Ist  $s = 0$ , so hat man offenbar das Optimum von  $\psi(x)$  erreicht. Sei daher  $s \neq 0$ . Nun bestimmt man ein  $\lambda \geq 0$  derart, daß

$$y = x + \lambda s$$

gerade noch in  $\mathfrak{B}$  liegt, d. h.

$$\lambda = \text{Min}_{1 \leq i \leq n_1} \left( \left| \frac{x_i}{s_i} \right| \mid s_i < 0 \right).$$

Danach beginnt man in  $y$ , das Problem (18) mit dem von D. ONTKEIT in [8] angegebenen linearen Programmierverfahren zu lösen.

### Literatur

1. CHARNES, A., W. W. COOPER and G. L. THOMPSON: Constrained generalized medians and hypermedians as deterministic equivalents for two-stage programs under uncertainty. *Management Science* **12**, 83ff. (1965).
2. DANTZIG, G. B.: *Linear programming and extensions*. Princeton: Princeton University Press 1963.
3. HADLEY, G.: *Nonlinear and dynamic programming*. Reading, Mass.: Addison-Wesley Publishing Company, Inc., 1964.

4. KALL, P.: Qualitative Aussagen zu einigen Problemen der stochastischen Programmierung. Z. Wahrscheinlichkeitstheorie verw. Geb. **6**, 246—272 (1966).
5. KÜNZI, H. P., u. W. KRELLE: Nichtlineare Programmierung. Berlin-Göttingen-Heidelberg: Springer 1962.
6. — Die Duoplex-Methode. Unternehmensforsch. **7**, 103—116 (1963).
7. MADANSKY, A.: Inequalities for stochastic linear programming problems. Management Science **6**, 197—204 (1960).
8. ONIGKEIT, D.: Ein lineares Programmierungsverfahren. Z. Wahrscheinlichkeitstheorie verw. Geb. **1**, 82—108 (1962).
9. WETS, R.: Programming under uncertainty: The complete problem. Z. Wahrscheinlichkeitstheorie verw. Geb. **4**, 316—339 (1966).
10. ZOUTENDIJK, G.: Methods of feasible directions. London-Amsterdam: Elsevier Publishing Company 1960.

Seminar für Angewandte Mathematik und  
Statistik der Universität Zürich