Z. Wahrscheinlichkeitstheorie verw. Geb. 6, 246—272 (1966)

Qualitative Aussagen zu einigen Problemen der stochastischen Programmierung*

P. KALL

Eingegangen am 15. Oktober 1965

§ 1. Einführung

In den letzten zwei Jahrzehnten ist eine ziemlich umfangreiche Literatur entstanden über lineare Programme, die sich stets in der Form

$$(1.1)\quad \begin{cases} & \min c'x \\ \text{unter den Restriktionen} & \\ & Ax = b \\ & x \geqq 0 \end{cases}$$

oder auch in der Form

$$(1.2)\quad \begin{cases} & \min c'x \\ \text{bezüglich} & \\ & Ax \geqq b \\ & x \geqq 0 \end{cases}$$

darstellen lassen, wobei A eine $(m \times n)$-Matrix, c und x n-Vektoren und b ein m-Vektor sind[1]. In der linearen Programmierung werden nun die Daten A, b, c stets als fest gegeben betrachtet. Dann ist es mit Hilfe bekannter Verfahren ohne weiteres möglich, die Probleme (1.1) und (1.2) zu lösen. Praktisch kann man diese Probleme interpretieren als Produktionspläne, bei denen aus den Faktoren x_j $(j = 1, \ldots, n)$ Endprodukte entweder genau in der Menge der Nachfragen b_i $(i = 1, \ldots, m)$ oder mindestens in der Menge der Nachfragen b_i $(i = 1, \ldots, m)$ hergestellt werden sollen derart, daß die Gesamtkosten $c'x$ minimiert werden. Natürlich arbeitet man in der Praxis mit nichtnegativen Faktormengen x_j $(j = 1, \ldots, n)$. Die Elemente der Matrix A bezeichnet man dann als technische Koeffizienten.

Nun kann man sich mit Recht fragen, ob die Daten A, b, c in der Praxis wirklich deterministischer Natur sind, wie in der linearen Programmierung angenommen wird. In vielen Fällen wird man dann feststellen, daß A, b, c oder wenigstens ein Teil dieser Daten in Wirklichkeit stochastische Variable sind. In diesem Fall verlieren die Problemstellungen (1.1) und (1.2) weitgehend ihren Sinn. Wir wollen daher in dieser Arbeit einige vernünftig erscheinende Problemstellungen für den Fall, daß A, b, c Zufallsvariable sind, untersuchen.

---

* Ich danke Herrn Prof. Dr. B. L. VAN DER WAERDEN und den Herren Dr. D. ONIGKEIT und Dr. K. KLEIBOHM recht herzlich für eingehende Diskussionen und wertvolle Ratschläge. Mein besonderer Dank gilt Herrn Prof. Dr. H. P. KÜNZI für seine mir stets gewährte Unterstützung, die diese Arbeit erst ermöglicht hat.

[1] Im Falle von Restriktionen in Gleichungsform nehmen wir stets an, daß $m \leqq n$ ist.

Zunächst kann man sich fragen, wie hoch die erwarteten Kosten sind, wenn man zu jeder Realisation der Zufallsvariablen A, b, c die optimale Lösung des Problems (1.1) bestimmt, sofern diese existiert.

Das Problem lautet also:

$$\text{(1.3)} \qquad \text{Bestimme } E\,(\min c'x \mid A\,x = b,\ x \geqq 0),$$

wobei der Erwartungswert bezüglich der gemeinsamen Verteilung der stochastischen Variablen in A, b, c zu bilden ist, sofern er existiert. Offenbar ist (1.3) kein eigentliches Entscheidungsproblem in dem Sinne, daß von vornherein ein optimales x zu bestimmen ist.

Ein echtes Entscheidungsproblem liegt hingegen vor, wenn man verlangt, daß ein $x \geqq 0$ so bestimmt wird, daß die Wahrscheinlichkeit dafür, daß x in (1.2) zulässig ist, maximiert wird. Wir haben also folgende Aufgabe:

$$\text{(1.4)} \quad \begin{cases} & \max P(\{A, b \mid A\,x \geqq b\}) \\ \text{bezüglich} & \\ & x \geqq 0\,. \end{cases}$$

P steht hier für das Wahrscheinlichkeitsmaß.

Mit diesem Problem hängt ein weiteres eng zusammen, das verlangt, ein $x \geqq 0$ zu bestimmen derart, daß die erwarteten Kosten minimiert werden unter der Bedingung, daß x mindestens mit einer vorgegebenen Wahrscheinlichkeit zulässig ist, also:

$$\text{(1.5)} \quad \begin{cases} & \min E\,(c'x) \\ \text{bezüglich} & \\ & P(\{A, b \mid A\,x \geqq b\}) \geqq 1 - \alpha \\ & x \geqq 0 \end{cases}$$

mit $0 \leqq \alpha < 1$.

Ein weiteres Problem entsteht, wenn man folgende Möglichkeit in Betracht zieht. Es wird zunächst eine Entscheidung über x getroffen, also $x = \hat{x} \geqq 0$ gesetzt. Nachdem dann die Realisationen von A, b, c bekannt sind, kann man feststellen, in welchem Maße die Restriktionen z. B. in (1.1) verletzt sind, d. h. man bestimmt den Vektor $b - A\,\hat{x}$, der die Fehlbeträge der Produktion angibt. In der Praxis führen solche Fehlbeträge im allgemeinen zu Strafkosten, die sich beispielsweise durch den Wert des Nachfrageüberschusses an Endprodukten bzw. durch die Lagerhaltungskosten für das die Nachfrage übersteigende Angebot an Endprodukten ergeben.

Man kann etwas allgemeiner annehmen, daß eine feste Matrix M und ein fester Vektor q existieren, so daß das lineare Programm

$$\text{(1.6)} \quad \begin{cases} & \min q'y \\ \text{bezüglich} & \\ & M\,y = b - A\,\hat{x} \\ & y \geqq 0 \end{cases}$$

für jede Realisation von A, b und für jede Entscheidung $\hat{x}$ eine zulässige Lösung y hat. M enthält die technischen Koeffizienten der zur Deckung des Nachfrageüberschusses nötigen Produktion und der zur Beseitigung des Überangebotes notwendigen Maßnahmen. Der Vektor y enthält die dazu nötigen Faktoren, und q

besteht aus den Preisen der Faktoren. Wegen seiner Bedeutung bezeichnen wir (1.6) als Notprogramm. Es erscheint dann vernünftig, $\hat{x}$ so zu bestimmen, daß der Erwartungswert der Gesamtkosten, die sich aus den Produktionskosten $c'x$ und den minimalen Strafkosten gemäß (1.6) zusammensetzen, minimiert wird. Das Problem lautet also:

$$(1.7) \qquad \left\{ \text{bezüglich} \quad \begin{array}{l} \min E[c'x + (\min q'y \mid My = b - Ax, y \geqq 0)] \\ x \geqq 0 . \end{array} \right.$$

Problem (1.7) wird in der Literatur als zweistufig bezeichnet, während die vorher genannten einstufig heißen.

Wir wollen nun untersuchen, welche Aussagen sich über die genannten Probleme machen lassen. Dabei wird sich zeigen, daß die Probleme (1.4) und (1.5) nach dem heutigen Stand der Optimierungstheorie im allgemeinen nicht lösbar sind, während (1.7) unter gewissen Voraussetzungen recht angenehme Eigenschaften hat.

§ 2. Betrachtungen zur Meßbarkeit

Offenbar ist die Aufgabenstellung (1.3) sinnlos, wenn nicht sichergestellt ist, daß $(\min c'x \mid Ax = b, x \geqq 0)$ eine meßbare Funktion von A, b, c ist. Andererseits ist diese Funktion sicher dann meßbar, wenn man zu jeder Realisation von A, b, c eine Optimallösung $x_{\text{opt}}(A, b, c)$ so angeben kann, daß $x_{\text{opt}}(A, b, c)$ eine meßbare Transformation von A, b, c in den R_n ist. In der Tat läßt sich nun folgendes beweisen:

Satz 1. *$x_{\text{opt}}(A, b, c)$ läßt sich als meßbare Transformation von A, b, c konstruieren.*

Beweis. Den durch die Elemente von A, b, c aufgespannten Raum nennen wir $R(A, b, c)$. Zunächst betrachten wir die Menge $\mathfrak{M} \subset R(A, b, c)$ der Punkte, die keine zulässige Lösung gestatten. Nach der Theorie der linearen Programmierung zerfällt $\mathfrak{M}$ in zwei Teile:

$\mathfrak{M}_1$, die Menge aller Punkte (A, b, c), für die jede Determinante aus Elementen von A in einer bestimmten Zeilenkombination verschwindet, während mindestens eine dieser Determinanten von Null verschieden ist, wenn man eine geeignete Spalte durch die entsprechenden Elemente von b ersetzt. Das Gleichungssystem $Ax = b$ ist dann offenbar unverträglich.

$\mathfrak{M}_2$, die Menge der Punkte (A, b, c), die nicht zu $\mathfrak{M}_1$ gehören, d.h. $(A, b, c) \in \overline{\mathfrak{M}}_1$, für die in jeder Basislösung von $Ax = b$ mindestens eine Komponente streng negativ ist. Das Gleichungssystem $Ax = b$ ist also lösbar, hat jedoch keine nichtnegativen Lösungen.

$\mathfrak{M}_1$ und $\mathfrak{M}_2$ sind dadurch bestimmt, daß gewisse Determinanten und Vektorprodukte in vorgegebenen Borelmengen liegen. Da Determinanten und Vektorprodukte stetige und damit meßbare Funktionen sind, sind $\mathfrak{M}_1$ und $\mathfrak{M}_2$ und daher $\mathfrak{M} = \mathfrak{M}_1 \cup \mathfrak{M}_2$ meßbare Mengen. Da auf $\mathfrak{M}$ ohnehin kein zulässiges x existiert, können wir auf $\mathfrak{M}$ $x_{\text{opt}}(A, b, c)$ folgendermaßen definieren:

$$(2.1) \qquad x_{\text{opt}\,i} = \left\{ \begin{array}{ll} +\infty & \text{falls} \quad c_i \geqq 0 \\ -\infty & \text{falls} \quad c_i < 0 \end{array} \right. \quad i = 1, \ldots, n; \quad (A, b, c) \in \mathfrak{M} .$$

Die Definition (2.1) drückt lediglich aus, daß im Falle der Unerfüllbarkeit der Restriktionen die Kosten als unendlich hoch angenommen werden können.

Bevor wir nun die Punkte $(A, b, c) \in R(A, b, c)$ betrachten, die zulässige Lösungen gestatten, wollen wir noch ausschließen, daß die Matrix A einen kleineren Rang als m hat (m ist die Zeilenzahl und n die Spaltenzahl von A, und es gilt nach der Fußnote in § 1 $m \leqq n$). Die Matrix A enthält $r = \binom{n}{m}$ m-reihige quadratische Untermatrizen, die mit B_i bezeichnet werden. Ist $|B_i|$ die Determinante von B_i, so sondern wir also die Menge

$$\mathfrak{N} = \bigcap_{i=1}^{r} \{A, b, c \,|\, |B_i| = 0\} \tag{2.2}$$

aus. Bis auf weiteres betrachten wir nur noch Punkte (A, b, c), die in

$$R'(A, b, c) = R(A, b, c) - \mathfrak{M} - \mathfrak{N} \tag{2.3}$$

liegen. Auf $R'(A, b, c)$ definieren wir folgende Mengen:

$$\mathfrak{B}_i = \{A, b, c \,|\, |B_i| \neq 0;\ B_i^{-1} b \geqq 0\}, \quad i = 1, \ldots, r. \tag{2.4}$$

Die $\mathfrak{B}_i$ sind offensichtlich meßbar, und nach Konstruktion gilt

$$\bigcup_{i=1}^{r} \mathfrak{B}_i = R'(A, b, c). \tag{2.5}$$

Die Indexmenge $\{1, \ldots, r\}$ hat, wenn man die leere Menge mitzählt, 2^r Teilmengen $I(\nu)$, $\nu = 1, \ldots, 2^r$. Um x_{opt} definieren zu können, definieren wir folgende Mengen:

$$\begin{aligned} R_{i\infty} = \bigcup_{j=1}^{n} \{A, b, c \,|\, (A, b, c) \in \overline{\mathfrak{B}}_i;\ B_i^{-1} A_j \leqq 0;\ c_j < \tilde{c}' B_j^{-1} A_j\} - \\ - \bigcup_{j=1}^{i-1} R_{j\infty}; \quad i = 1, \ldots, r. \end{aligned} \tag{2.6}$$

Darin ist A_j die j-te Spalte von A, c_j die j-te Komponente von c, und $\tilde{c}$ ist der Vektor aus den Komponenten von c, die zu der Basis B_i gehören. Ferner definieren wir

$$\begin{aligned} R_i = \bigcup_{\nu=1}^{2^r} [\mathfrak{B}_i \cap \bigcap_{j \in I(\nu)} \mathfrak{B}_j \cap \bigcap_{k \in \overline{I(\nu)}} \overline{\mathfrak{B}}_k \cap \bigcap_{\substack{j \in I(\nu) \\ j \geqq i}} \{A, b, c \,|\, \tilde{c}' B_i^{-1} b \leqq \tilde{c}' B_j^{-1} b\} \cap \\ \cap \bigcap_{\substack{j \in I(\nu) \\ j < i}} \{A, b, c \,|\, \tilde{c}' B_i^{-1} b < \tilde{c}' B_j^{-1} b\}] \cap \overline{\bigcup_{j=1}^{r} R_{j\infty}}; \quad i = 1, \ldots, r. \end{aligned} \tag{2.7}$$

$\overline{I(\nu)}$, $\overline{\mathfrak{B}}_k$ und $\overline{\bigcup_{j=1}^{r} R_{j\infty}}$ sollen angeben, daß die Komplemente der betreffenden Mengen zu bilden sind. In (2.7) ist zu beachten, daß der Durchschnitt über die leere Indexmenge nach Definition gleich dem ganzen betrachteten Raum, also $R'(A, b, c)$ ist. Auf Grund der Konstruktion sind die Mengen $R_{i\infty}$ und R_i alle disjunkt. Ferner sind sie meßbar. Auf $R_{i\infty}$ hat das lineare Programm offenbar eine unendliche Lösung, während auf R_i eine endliche Lösung vorliegt. Offenbar gilt

$$\bigcup_{i=1}^{r} (R_i \cup R_{i\infty}) = R'(A, b, c). \tag{2.8}$$

Nun definieren wir

$$\tilde{x}_{\text{opt}} = B_i^{-1} b \quad \text{falls} \quad (A, b, c) \in R_i; \quad i = 1, \ldots, r. \tag{2.9}$$

$\tilde{x}_{\text{opt}}$ ist hier der Vektor der zur Basis B_i gehörenden Komponenten von x_{opt}. Alle anderen Komponenten sind gleich Null. Ferner definieren wir x_{opt} auf $R_{i\infty}$ folgendermaßen:

(2.10) Die Komponenten von $\tilde{x}_{\text{opt}}$, die zu den verschwindenden Komponenten von $B_i^{-1} A_j$ gehören, bleiben gegenüber (2.9) unverändert. Die übrigen Komponenten von $\tilde{x}_{\text{opt}}$ setzen wir gleich ∞. Außerdem wird natürlich die Komponente $x_{\text{opt}\,j} = \infty$. Um die Eindeutigkeit dieser Definition zu sichern, beziehen wir uns hier auf den kleinsten Index j, für den gemäß (2.6) $B_i^{-1} A_j \leqq 0$ und $c_j < \tilde{c}' B_i^{-1} A_j$ gilt.

Durch (2.9) und (2.10) haben wir x_{opt} auf $R'(A, b, c)$ als meßbare Transformation definiert. Auf $\mathfrak{M}$ war x_{opt} bereits in (2.1) festgelegt worden. Lediglich auf $\mathfrak{N} - \mathfrak{M}$ fehlt noch eine Definition, die man jedoch ganz analog zu der in (2.9) und (2.10) geben kann, indem man mit Basen von geringerer Dimension als m arbeitet. Damit ist der Satz bewiesen.

Zu diesem Satz sind zwei Bemerkungen zu machen. Zunächst wird im Beweis $m \leqq n$ angenommen. Diese Voraussetzung ist jedoch nicht wesentlich. Ist $m > n$, so ist das System $A x = b$ nur dann lösbar, wenn jede $(n + 1)$-reihige Determinante der erweiterten Matrix (A, b) verschwindet. Daraus ergibt sich die erforderliche Abänderung des Beweises: $\mathfrak{M}$ ist um die Menge der Punkte (A, b, c) zu erweitern, für die mindestens eine der erwähnten $(n + 1)$-reihigen Determinanten von Null verschieden ist. Auf der Restmenge $R(A, b, c) - \mathfrak{M}$ läßt sich dann die Beweisidee analog durchführen, d.h. man konstruiert eine disjunkte Zerlegung meßbarer Mengen dieser Restmenge, auf denen sich x_{opt} eindeutig und meßbar definieren läßt.

Zum anderen ist darauf hinzuweisen, daß die Meßbarkeit für die Existenz des in (1.3) zu bildenden Erwartungswertes zwar notwendig, aber keineswegs hinreichend ist. Für die Existenz sind gewisse Bedingungen an die gemeinsame Wahrscheinlichkeitsverteilung $P(A, b, c)$ zu stellen. Zum Beispiel ist die Bedingung

$$P(\mathfrak{M}) + \sum_{i=1}^{r} P(R_{i\infty}) = 0 \quad \text{sicher notwendig}.$$

Damit wollen wir dieses Problem verlassen und uns den nächsten beiden zuwenden.

§ 3. Aussagen über $\mathbf{P}(\{\mathbf{A}, \mathbf{b} \mid \mathbf{A}\mathbf{x} \geqq \mathbf{b}\})$

In (1.4) und (1.5) haben wir Programmierungsaufgaben formuliert, deren Beschaffenheit wir uns jetzt genauer ansehen wollen. Dabei wird sich herausstellen, daß wir es hier im allgemeinen nicht mit konvexen Programmierungsproblemen zu tun haben, wenn man unter einem konvexen Problem die Minimierung einer konvexen oder die Maximierung einer konkaven Funktion über einem konvexen Bereich versteht. Da nach dem gegenwärtigen Stand der Optimierungstheorie für nicht-konvexe Probleme praktisch keine brauchbaren Lösungsverfahren existieren, bedeutet das, daß man vorerst im allgemeinen nicht damit rechnen kann, die Probleme (1.4) und (1.5) zu lösen.

Dafür, daß die genannten Probleme im allgemeinen nicht konvex sind, geben wir das folgende kleine Beispiel: Gegeben sei die zweidimensionale Zufallsvariable $\begin{pmatrix} a \\ b \end{pmatrix}$ mit der diskreten Verteilung

$$P\left[\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -3 \\ -1 \end{pmatrix}\right] = \frac{1}{3}\,; \quad P\left[\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}\right] = \frac{2}{3}\,. \tag{3.1}$$

Sei

1. $$x \leqq +\frac{1}{3}\,.$$

Daraus folgt

$$-3x + 1 \geqq 0$$

und

$$3x - 2 \leqq +1 - 2 < 0\,.$$

2. $$+\frac{1}{3} < x < \frac{2}{3}\,.$$

Dann gilt

$$-3x + 1 < -1 + 1 = 0$$

und

$$3x - 2 < 2 - 2 = 0\,.$$

3. $$x \geqq \frac{2}{3}\,.$$

Dann ist

$$\begin{aligned} -3x + 1 &\leqq -2 + 1 < 0 \\ 3x - 2 &\geqq 0\,. \end{aligned}$$

Folglich gilt

$$P(\{a, b \,|\, a x \geqq b\}) = \begin{cases} \frac{1}{3} & \text{falls} \quad x \leqq +\frac{1}{3} \\ 0 & \text{falls} \quad +\frac{1}{3} < x < \frac{2}{3} \\ \frac{2}{3} & \text{falls} \quad x \geqq \frac{2}{3}\,. \end{cases} \tag{3.2}$$

Offensichtlich ist $P(\{a, b \,|\, ax \geqq b\})$ über $x \geqq 0$ nicht konkav. Also ist (1.4) in diesem Fall kein konvexes Problem. Bei diesem Beispiel ist aber auch (1.5) nicht immer konvex, nämlich dann nicht, wenn wir in (1.5) α so wählen, daß $\frac{2}{3} \leqq \alpha < 1$.

Nachdem wir gezeigt haben, daß es für $\alpha > 0$ vorkommen kann, daß der zulässige Bereich von (1.5) nicht konvex ist, wollen wir einen wichtigen Spezialfall erwähnen.

Lemma 1. *Ist $\alpha = 0$, dann ist der zulässige Bereich von* (1.5) *konvex.*

Beweis. Seien x^1 und x^2 zulässige Lösungen von (1.5). Sei

$$\begin{aligned} \Omega_1 &= \{A, b \,|\, A\,x^1 \geqq b\}\,, \\ \Omega_2 &= \{A, b \,|\, A\,x^2 \geqq b\}\,. \end{aligned}$$

$R(A,b)$ sei der durch alle Punkte (A,b) aufgespannte Raum. Dann gilt, da $\alpha = 0$ und x^1 zulässig ist,

$$0 \leqq P(\Omega_2 - \Omega_1) \leqq P(R(A,b) - \Omega_1) = 0 .$$

Folglich ist

$$P(\Omega_2 \cap \Omega_1) = P(\Omega_2) - P(\Omega_2 - \Omega_1) = 1 ,$$

da x^2 zulässig ist. Für $(A, b) \in \Omega_1 \cap \Omega_2$ gilt aber für jedes λ mit $0 < \lambda < 1$

$$A(\lambda x^1 + (1-\lambda) x^2) \geqq b .$$

Also ist $\lambda x^1 + (1-\lambda) x^2$ zulässig. Damit ist der Beweis des Lemmas erbracht.

§ 4. Einführung des zweistufigen Modells

Bevor wir zu der in (1.7) formulierten Aufgabe kommen, betrachten wir zunächst ein praktisches Problem. Wir nehmen an, ein Unternehmer müsse mit seiner Produktion, die zunächst mit einer festen technischen Matrix A und festen Faktorpreisen c arbeitet, eine feste Nachfrage b mindestens decken. Verlangt man nun die Minimierung der Kosten, so entsteht bekanntlich das lineare Programm

$$\text{(4.1)} \quad \left\{ \begin{array}{rl} & \min c' x \\ \text{bezüglich} & A x \geqq b \\ & x \geqq 0 . \end{array} \right.$$

Nehmen wir nun an, daß einige oder alle Daten in A, b, c stochastische Variable sind, so müssen wir natürlich eine neue Entscheidungsregel suchen. Da man im Zeitpunkt der Entscheidung die Realisationen der Daten nicht kennt, muß man damit rechnen, daß die getroffene Entscheidung die Restriktionen $Ax \geqq b$ verletzt. Man kann nun versuchen, durch Einführung von Strafkosten zu erreichen, daß diese Bereichsüberschreitungen nicht zu häufig vorkommen. Da eine Bereichsüberschreitung hier einen Angebotsmangel bei einzelnen Produkten bedeutet, erscheint es vernünftig, die Strafkosten gleich dem Wert der fehlenden Produktmengen zu setzen, wobei wir die Produktpreise als fest annehmen. Bezeichnen wir die Produktpreise mit p_i und die i-te Zeile von A mit A_i, und wollen wir unsere Entscheidung über x so treffen, daß die erwarteten Gesamtkosten (gleich Produktions- plus Strafkosten) minimiert werden, dann haben wir folgende Problemstellung:

$$\text{(4.2)} \quad \left\{ \begin{array}{rl} & \min \left[\sum\limits_{j=1}^{n} \int c_j x_j \, dP + \sum\limits_{i=1}^{m} p_i \int\limits_{A_i x < b_i} (b_i - A_i x) \, dP \right] \\ \text{bezüglich} & x \geqq 0 . \end{array} \right.$$

P ist darin die gemeinsame Wahrscheinlichkeitsverteilung von A, b, c. Wir setzen für diesen und alle folgenden Paragraphen voraus, daß die Erwartungswerte $\bar{A} = E(A)$, $\bar{b} = E(b)$ und $\bar{c} = E(c)$ existieren.

Im Gegensatz zu (4.2) entsteht das übliche lineare Programm aus (4.1) dadurch — oder man kann es jedenfalls so interpretieren —, daß alle zufälligen

Daten durch ihre Erwartungswerte ersetzt werden. Wir haben also

$$(4.3)\qquad \left\{\begin{array}{ll} & \min \bar{c}'x \\ \text{bezüglich} & \bar{A}x \geqq \bar{b} \\ & x \geqq 0 . \end{array}\right.$$

Hat (4.3) eine endliche Lösung, so hat das dazu duale Problem

$$(4.4)\qquad \left\{\begin{array}{ll} & \max \bar{b}'p \\ \text{bezüglich} & \bar{A}'p \leqq \bar{c} \\ & p \geqq 0 \end{array}\right.$$

eine zulässige Lösung. Geht man von den festen Daten $\bar{A}, \bar{b}, \bar{c}$ aus und interpretiert man p als den Vektor der Produktpreise, so bedeutet (4.4) bekanntlich die Annahme vollständiger Konkurrenz. Man beweist nun leicht den folgenden

Satz 1. *Unter der Annahme, daß die p_i eine zulässige Lösung von* (4.4) *darstellen, ist $\sum_{i=1}^{m} p_i \bar{b}_i$ eine untere Schranke der Zielfunktion von* (4.2).

Beweis. Sei $\hat{x} \geqq 0$ beliebig gewählt. Dann gilt, da p zulässig ist in (4.4),

$$\sum_{j=1}^{n} \int c_j \hat{x}_j \, dP + \sum_{i=1}^{m} p_i \int\limits_{A_i\hat{x}<b_i} (b_i - A_i\hat{x})\, dP = \bar{c}'\hat{x} + \sum_{i=1}^{m} p_i \int\limits_{A_i\hat{x}<b_i} (b_i - A_i\hat{x})\, dP \geqq$$
$$\geqq \bar{c}'\hat{x} + \sum_{i=1}^{m} p_i(\bar{b}_i - \bar{A}_i x) = (\bar{c}' - p'\bar{A})\hat{x} + p'\bar{b} \geqq p'\bar{b},$$

und das war zu beweisen.

Für die Lösbarkeit des Problems (4.2) ist nach dem heutigen Stand der Optimierungstheorie der nächste Satz besonders wichtig.

Satz 2. *Gilt $p_i \geqq 0$, $i = 1, \ldots, m$, was für Strafkosten in der Praxis ohnehin gegeben ist, dann ist die Zielfunktion von* (4.2) *konvex in x.*

Beweis. Seien x^1 und x^2 beliebige Vektoren, $0 < \lambda < 1$, $\hat{x} = \lambda x^1 + (1-\lambda)x^2$. Da der erste Teil der Zielfunktion von (4.2), $\bar{c}'x$, ohnehin konvex in x ist, genügt es, die Konvexität von $p_i \int\limits_{A_i x<b_i} (b_i - A_i x)\, dP$ in x zu beweisen.

Seien dazu

$$\begin{aligned} \mathfrak{B} &= \{A, b \mid A_i\hat{x} < b_i\}, \\ \mathfrak{B}_1 &= \{A, b \mid A_i\hat{x}^1 < b_i\} \cap \{A, b \mid A_i x^2 < b_i\}, \\ \mathfrak{B}_2 &= \{A, b \mid A_i\hat{x} < b_i\} \cap \{A, b \mid A_i x^2 \geqq b_i\}, \\ \mathfrak{B}_3 &= \{A, b \mid A_i\hat{x} < b_i\} \cap \{A, b \mid A_i x^1 \geqq b_i\}. \end{aligned}$$

Offenbar gilt

$$\mathfrak{B} = \mathfrak{B}_1 \cup \mathfrak{B}_2 \cup \mathfrak{B}_3$$

und

$$\mathfrak{B}_i \cap \mathfrak{B}_j = \emptyset, \quad \text{falls} \quad i \neq j .$$

Dabei ist $\emptyset$ die leere Menge. Seien ferner

$$\begin{aligned} \mathfrak{A}_1 &= \{A, b \mid A_i x^1 < b_i\} = \mathfrak{B}_1 \cup \mathfrak{B}_2 \cup \mathfrak{D}_1, \\ \mathfrak{A}_2 &= \{A, b \mid A_i x^2 < b_i\} = \mathfrak{B}_1 \cup \mathfrak{B}_3 \cup \mathfrak{D}_2 . \end{aligned}$$

$\mathfrak{D}_1$ und $\mathfrak{D}_2$ sind eindeutig bestimmt gemäß

$$\mathfrak{D}_j = \mathfrak{A}_j - (\mathfrak{B}_1 \cup \mathfrak{B}_{j+1}), \quad j = 1, 2.$$

Nun ist

$$\begin{aligned}
p_i \int\limits_{A_i\hat{x}<b_i} (b_i - A_i\hat{x})\, dP &= p_i \int\limits_{\mathfrak{B}} (b_i - A_i\hat{x})\, dP \\
&= \lambda p_i \int\limits_{\mathfrak{B}} (b_i - A_i x^1)\, dP + (1-\lambda) p_i \int\limits_{\mathfrak{B}} (b_i - A_i x^2)\, dP \\
\text{1.} \qquad &= \lambda p_i \int\limits_{\mathfrak{B}_1\cup\mathfrak{B}_2} (b_i - A_i x^1)\, dP + \lambda p_i \int\limits_{\mathfrak{B}_3} (b_i - A_i x^1)\, dP + \\
&\quad + (1-\lambda) p_i \int\limits_{\mathfrak{B}_1\cup\mathfrak{B}_3} (b_i - A_i x^2)\, dP + (1-\lambda) p_i \int\limits_{\mathfrak{B}_2} (b_i - A_i x^2)\, dP \\
\text{2.} \qquad &\leqq \lambda p_i \int\limits_{\mathfrak{A}_1} (b_i - A_i x^1)\, dP + (1-\lambda) p_i \int\limits_{\mathfrak{A}_2} (b_i - A_i x^2)\, dP \\
&= \lambda p_i \int\limits_{A_i x^1<b_i} (b_i - A_i x^1)\, dP + (1-\lambda) p_i \int\limits_{A_i x^2<b_i} (b_i - A_i x^2)\, dP.
\end{aligned}$$

Die Relationen 1. und 2. folgen aus der Tatsache, daß

1. die $\mathfrak{B}_j$ disjunkt sind und ihre Vereinigung gleich $\mathfrak{B}$ ist, und
2. $b_i - A_i x_0^j > 0$ auf $\mathfrak{D}_j$, $j = 1, 2$ und
   $b_i - A_i x^1 \leqq 0$ auf $\mathfrak{B}_3$ und
   $b_i - A_i x^2 \leqq 0$ auf $\mathfrak{B}_2$

sowie aus der Voraussetzung $p_i \geqq 0$. Damit ist die Behauptung bewiesen.

Nun wollen wir die Verbindung zwischen den Problemen (4.2) und (1.7) herstellen. Dazu beweisen wir den

Satz 3. *Sei $p_i \geqq 0$, $i = 1, \ldots, m$. Dann ist Problem* (4.2) *äquivalent zu*

$$(4.5) \quad \left\{ \begin{array}{ll} & \min E[c'x + \min(p'w \mid w - v = b - Ax, v \geqq 0, w \geqq 0)] \\ \textit{bezüglich} & x \geqq 0. \end{array} \right.$$

Beweis. Zum Beweis betrachten wir für alle festen A, b und x das Problem

$$(4.6) \quad \left\{ \begin{array}{ll} & \min p'w \\ \text{bezüglich} & w - v = b - Ax, \\ & v \geqq 0, \quad w \geqq 0 \end{array} \right.$$

und definieren (v^*, w^*) wie folgt:

$$\begin{aligned}
v_i^* &= 0, \quad w_i^* = b_i - A_i x, \quad \text{falls} \quad b_i - A_i x > 0, \\
v_i^* &= A_i x - b_i, \quad w_i^* = 0, \quad \text{falls} \quad b_i - A_i x \leqq 0.
\end{aligned}$$

Offenbar ist (v^*, w^*) eine optimale zulässige Lösung von (4.6). Wäre nämlich $b_i - A_i x > 0$ und $v_i = \alpha > 0$, so wäre $w_i = b_i - A_i x + v_i = b_i - A_i x + \alpha$, und die entsprechende Kostenkomponente wäre $p_i w_i = p_i w_i^* + p_i \alpha \geqq p_i w_i^*$, da $p_i \geqq 0$ ist nach Voraussetzung. Analog beweist man die Optimalität von (v^*, w^*), falls $b_i - A_i x \leqq 0$ ist. Mit dieser Optimallösung bilden wir den Erwartungswert $E(p'w^*)$ und erhalten

$$\sum_{i=1}^{m} p_i \int w_i^*\, dP = \sum_{i=1}^{m} p_i \int\limits_{A_i x<b_i} (b_i - A_i x)\, dP.$$

Damit geht (4.5) über in

$$(4.7)\quad \begin{cases} \min\left[\bar{c}'x + \sum_{i=1}^{m} p_i \int\limits_{A_i x < b_i} (b_i - A_i x)\, dP\right] \\ \text{bezüglich} \quad x \geqq 0\,, \end{cases}$$

was mit (4.2) identisch ist.

Damit haben wir gezeigt, daß (4.2) unter der Voraussetzung $p_i \geqq 0$, $i = 1, \ldots, m$, ein spezielles zweistufiges Problem ist. In (4.2) konnten wir aus praktischen Gründen stets annehmen, daß $p \geqq 0$ war. Es wird sich nun zeigen, daß in (4.5) diese Annahme auch notwendig ist, wenn dieses Problem sinnvoll sein soll. Dazu verallgemeinern wir (4.5) folgendermaßen:

$$(4.8)\quad \begin{cases} \min E[c'x + \min(q^{+\prime} y^{+} + q^{-\prime} y^{-} \,|\, y^{+} - y^{-} = b - Ax; \\ \qquad y^{+} \geqq 0, y^{-} \geqq 0)] \\ \text{bezüglich} \quad x \geqq 0\,. \end{cases}$$

Für dieses Problem gilt nun

Satz 4. *Das lineare Programm der zweiten Stufe*

$$(4.9)\quad \begin{cases} \min(q^{+\prime} y^{+} + q^{-\prime} y^{-}) \\ \textit{bezüglich} \quad y^{+} - y^{-} = b - Ax, \\ \qquad y^{+} \geqq 0\,, \quad y^{-} \geqq 0 \end{cases}$$

hat dann und nur dann eine endliche Lösung, wenn $\tilde{q} = q^{+} + q^{-} \geqq 0$ *ist.*

Beweis. (4.9) hat offenbar stets zulässige Lösungen. Bekanntlich ist für feste A, b, x jedoch die optimale zulässige Lösung von (4.9) dann und nur dann endlich, wenn das zugehörige duale Programm mindestens eine zulässige Lösung besitzt. Dieses Programm lautet

$$(4.10)\quad \begin{cases} \max(b - Ax)'z \\ \text{bezüglich} \quad z \leqq q^{+}, \\ \qquad -z \leqq q^{-}. \end{cases}$$

Diese Restriktionen lassen aber offenbar dann und nur dann Lösungen zu, wenn $-q^{-} \leqq q^{+}$ ist, und das ist die Behauptung.

Damit ist die Bedingung für die Endlichkeit der Lösung von (4.9) unabhängig von A, b und x. Umgekehrt bedeutet das, daß in (4.9) für jedes A, b und x der Optimalwert der Zielfunktion gleich $-\infty$ wird, falls die Bedingung $\tilde{q} \geqq 0$ verletzt ist. Damit verliert aber offenbar (4.8) seinen Sinn, da dann die Zielfunktion für jedes x ebenfalls gleich $-\infty$ wird. Da in (4.5) $q^{-} = 0$ war, ist dort die Bedingung $\tilde{q} \geqq 0$ identisch mit $p = q^{+} \geqq 0$.

Offenbar ist das in (1.7) angegebene Programm eine Verallgemeinerung von (4.8). Damit wollen wir uns im nächsten Paragraphen befassen.

§ 5. Das allgemeine zweistufige Problem

Wir wollen nun also das Problem

$$(1.7)\quad \begin{cases} \min E[c'x + (\min q'y \,|\, My = b - Ax, y \geqq 0)] \\ \text{bezüglich} \quad x \geqq 0 \end{cases}$$

untersuchen. Da es im allgemeinen zu jedem $x \geqq 0$ und zu jedem A, b möglich sein soll, die Bereichsüberschreitung $b - Ax$ durch My mit $y \geqq 0$ zu kompensieren, ist die folgende Voraussetzung gerechtfertigt:

V 1. Zu jedem $z \in R_m$ existiert ein $y \geqq 0$, so daß $My = z$ ist.

Wir wollen uns überlegen, welche Konsequenzen die Voraussetzung für die Struktur der Matrix M hat. Zunächst ist klar, daß M den Rang m haben muß, wobei m die Zeilenzahl von A und M ist. Ferner beweist man leicht das

Lemma 1. *Unter der Voraussetzung* **V 1** *hat M mindestens $m+1$ Spalten.*

Beweis. Mit M_i bezeichnen wir die Spalten von M. Da M den Rang m hat, können wir annehmen, daß $M_1, \ldots, M_m$ linear unabhängig sind. Nehmen wir nun an, M habe genau m Spalten. Dann ist mit $\alpha_i > 0$, $i = 1, \ldots, m$,

$$z_0 = \sum_{i=1}^{m} \alpha_i M_i \in R_m .$$

Der Vektor $-z_0 \in R_m$ hat dann wegen der linearen Unabhängigkeit der M_i, $i = 1, \ldots, m$, die eindeutige Darstellung

$$-z_0 = \sum_{i=1}^{m} -\alpha_i M_i$$

im Widerspruch zur Voraussetzung **V 1**. Also hat M mindestens $m+1$ Spalten.

Im weiteren ist es zweckmäßig, zwei Fälle zu unterscheiden: M hat genau $m+1$ Spalten, oder M hat mehr als $m+1$ Spalten. Ohne Einschränkung der Allgemeinheit können wir in diesem Paragraphen annehmen, daß die ersten m Spalten von M linear unabhängig sind.

Satz 1. *M habe $m+1$ Spalten. Unter der Voraussetzung* **V 1** *sind je m Spalten linear unabhängig.*

Beweis. Nehmen wir an, die Spalten $M_1, \ldots, M_{m-1}, M_{m+1}$ seien linear abhängig. Dann ist, da $M_1, \ldots, M_m$ linear unabhängig sind,

$$M_{m+1} = \sum_{i=1}^{m-1} \alpha_i M_i . \tag{5.1}$$

Nach **V 1** kann man $-M_m$ als nichtnegative Linearkombination von $M_1, \ldots, M_{m+1}$ darstellen, d. h.

$$\begin{aligned} -M_m &= \sum_{i=1}^{m+1} \beta_i M_i \quad \text{mit} \quad \beta_i \geqq 0 \\ &= \sum_{i=1}^{m} \beta_i M_i + \beta_{m+1} \sum_{i=1}^{m-1} \alpha_i M_i \quad \text{nach (5.1)} \\ &= \sum_{i=1}^{m-1} (\beta_i + \beta_{m+1} \alpha_i) M_i + \beta_m M_m . \end{aligned}$$

Folglich gilt

$$\sum_{i=1}^{m-1} (\beta_i + \beta_{m+1} \alpha_i) M_i + (1 + \beta_m) M_m = 0 .$$

Also sind, da $1 + \beta_m > 0$ ist, die Spalten $M_1, \ldots, M_m$ linear abhängig im Widerspruch zur Voraussetzung. Daraus folgt die Behauptung.

Bis jetzt haben wir nur Bedingungen abgeleitet, die für das Erfülltsein der Voraussetzung **V 1** notwendig waren. Man kann jedoch auch Bedingungen an die Struktur von M stellen, die für die Gültigkeit von **V 1** notwendig und hinreichend sind. Seien v, v_i Vektoren und λ_i Skalare; dann führen wir folgende Bezeichnungen ein:

$$K^+(v_1, \ldots, v_n) = \{v \mid v = \sum_{i=1}^{n} \lambda_i v_i, \lambda_i \geqq 0\},$$

$$K^-(v_1, \ldots, v_n) = \{v \mid v = \sum_{i=1}^{n} \lambda_i v_i, \lambda_i < 0\}.$$

Damit können wir folgende Aussage beweisen:

Satz 2. *Wenn M genau $m + 1$ Spalten hat, dann ist für* **V1** *notwendig und hinreichend, daß*

$$M_{m+1} \in K^-(M_1, \ldots, M_m).$$

Beweis. Die Bedingung ist notwendig, denn sei

$$M_{m+1} = \sum_{i=1}^{m} \alpha_i M_i. \tag{5.2}$$

Diese Darstellung ist wegen der linearen Unabhängigkeit von $M_1, \ldots, M_m$ eindeutig. Wir nehmen nun an, es existiere mindestens ein $i \leqq m$, so daß $\alpha_i \geqq 0$ ist. Ohne Einschränkung der Allgemeinheit können wir $\alpha_m \geqq 0$ annehmen. Nach **V1** gibt es Koeffizienten $\beta_i \geqq 0$ derart, daß

$$-M_m = \sum_{i=1}^{m+1} \beta_i M_i.$$

Nach (5.2) ist dann

$$-M_m = \sum_{i=1}^{m} \beta_i M_i + \beta_{m+1} \sum_{i=1}^{m} \alpha_i M_i = \sum_{i=1}^{m} (\beta_i + \beta_{m+1} \alpha_i) M_i$$

oder

$$\sum_{i=1}^{m-1} (\beta_i + \beta_{m+1} \alpha_i) M_i + (1 + \beta_m + \beta_{m+1} \alpha_m) M_m = 0.$$

Da $1 + \beta_m + \beta_{m+1} \alpha_m > 0$ ist, sind $M_1, \ldots, M_m$ linear abhängig im Widerspruch zur Voraussetzung. Folglich ist unsere Annahme falsch. Damit ist die Notwendigkeit bewiesen.

Die Bedingung ist aber auch hinreichend. Sei nämlich

$$M_{m+1} = \sum_{i=1}^{m} \alpha_i M_i \text{ mit } \alpha_i < 0 \text{ für } 1 \leqq i \leqq m, \tag{5.3}$$

und sei $z \in R_m$ beliebig gewählt. Wegen der linearen Unabhängigkeit von $M_1, \ldots, M_m$ existiert eine eindeutige Darstellung

$$z = \sum_{i=1}^{m} \beta_i M_i. \tag{5.4}$$

In (5.4) sind zwei Fälle möglich:

1. $\beta_i \geqq 0$, $1 \leqq i \leqq m$; dann sind wir bereits fertig, da $z \in K^+(M_1, \ldots, M_m)$.

2. Es ist mindestens ein $\beta_i < 0$.

Wir müssen nun zeigen, daß außer der Darstellung (5.4) noch eine Darstellung von z möglich ist, in der alle Koeffizienten nichtnegativ sind. Dazu suchen wir das

$$\max_{1 \leqq i \leqq m} \frac{\beta_i}{\alpha_i},$$

das sicher positiv ist, da alle $\alpha_i < 0$ und mindestens ein $\beta_i < 0$ sind. Wir können ohne Einschränkung der Allgemeinheit annehmen, daß dieses Maximum von $\frac{\beta_m}{\alpha_m}$ erreicht wird, daß also

$$\frac{\beta_m}{\alpha_m} = \max_{1 \leqq i \leqq m} \frac{\beta_i}{\alpha_i} > 0 \quad \text{ist.} \tag{5.5}$$

Aus (5.3) folgt, da $M_1, \ldots, M_m$ linear unabhängig sind, auch die lineare Unabhängigkeit von $M_1, \ldots, M_{m-1}, M_{m+1}$. Folglich gibt es für z die weitere eindeutige Darstellung

$$z = \sum_{i=1}^{m-1} \gamma_i M_i + \gamma_{m+1} M_{m+1}. \tag{5.6}$$

Wir zeigen nun, daß $\gamma_i \geqq 0$ gilt für $i = 1, \ldots, m-1, m+1$. Setzt man (5.3) in (5.6) ein, so ergibt sich

$$z = \sum_{i=1}^{m-1} \gamma_i M_i + \gamma_{m+1} \sum_{i=1}^{m} \alpha_i M_i = \sum_{i=1}^{m-1} (\gamma_i + \gamma_{m+1} \alpha_i) M_i + \gamma_{m+1} \alpha_m M_m. \tag{5.7}$$

Vergleicht man (5.7) mit (5.4), so folgt

$$\sum_{i=1}^{m} \beta_i M_i = \sum_{i=1}^{m-1} (\gamma_i + \gamma_{m+1} \alpha_i) M_i + \gamma_{m+1} \alpha_m M_m. \tag{5.8}$$

Da $M_1, \ldots, M_m$ linear unabhängig sind, folgt aus (5.8)

$$\beta_i = \gamma_i + \gamma_{m+1} \alpha_i, \qquad 1 \leqq i \leqq m-1, \tag{5.9}$$

$$\beta_m = \gamma_{m+1} \alpha_m. \tag{5.10}$$

Folglich ist nach (5.5) und (5.10)

$$\gamma_{m+1} = \frac{\beta_m}{\alpha_m} > 0 \tag{5.11}$$

und nach (5.9) und (5.5)

$$\gamma_i = \beta_i - \alpha_i \frac{\beta_m}{\alpha_m} \geqq 0; \tag{5.12}$$

also ist $z \in K^+(M_1, \ldots, M_{m-1}, M_{m+1})$. Damit ist der Satz vollständig bewiesen.

Hat M mehr als $m+1$ Spalten, dann gilt eine ähnliche Aussage.

Satz 3. *Hat M $m+n$ Spalten $(n > 1)$, dann ist für* **V1** *notwendig und hinreichend, daß es Koeffizienten $\delta_i \geqq 0$ gibt, $i = m+1, \ldots, m+n$, so daß*

$$M_{m+n+1} = \sum_{i=m+1}^{m+n} \delta_i M_i \in K^-(M_1, \ldots, M_m). \tag{5.13}$$

Beweis. Die Bedingung ist hinreichend. Denn wenn wir statt der Matrix $M = (M_1, \ldots, M_{m+n})$ die Matrix $M^* = (M_1, \ldots, M_m, M_{m+n+1})$ betrachten, dann gibt es nach Satz 2 zu jedem $z \in R_m$ nichtnegative Koeffizienten $\beta_1, \ldots, \beta_m, \beta_{m+n+1}$ derart, daß

$$\begin{aligned} z &= \sum_{i=1}^{m} \beta_i M_i + \beta_{m+n+1} M_{m+n+1} \\ &= \sum_{i=1}^{m} \beta_i M_i + \beta_{m+n+1} \sum_{i=m+1}^{m+n} \delta_i M_i \qquad \text{nach (5.13).} \end{aligned} \tag{5.14}$$

Folglich ist $z \in K^+(M_1, \ldots, M_{m+n})$. Die Bedingung ist aber auch notwendig. Sei nämlich $z \in K^-(M_1, \ldots, M_m)$, also

$$z = \sum_{i=1}^{m} \beta_i M_i \quad \text{mit} \quad \beta_i < 0, \ i = 1, \ldots, m. \tag{5.15}$$

Nach **V 1** gibt es Koeffizienten δ_i, so daß

$$z = \sum_{i=1}^{m+n} \delta_i M_i \quad \text{mit} \quad \delta_i \geqq 0, \ i = 1, \ldots, m+n. \tag{5.16}$$

Vergleicht man (5.15) mit (5.16), so folgt

$$\sum_{i=1}^{m} \beta_i M_i = \sum_{i=1}^{m+n} \delta_i M_i$$

oder

$$\sum_{i=1}^{m} (\beta_i - \delta_i) M_i = \sum_{i=m+1}^{m+n} \delta_i M_i. \tag{5.17}$$

Da $\beta_i < 0$ und $\delta_i \geqq 0$ sind, ist (5.17) mit der Bedingung (5.13) identisch.

Auf Grund der Sätze 2 und 3 können wir nun die recht allgemein gehaltene Voraussetzung **V 1** durch die äquivalente Voraussetzung **V 2** ersetzen.

V 2. In der Matrix M mit $m + n$ Spalten ($n \geqq 1$) sind — eventuell nach Umordnung — die ersten m Spalten $M_1, \ldots, M_m$ linear unabhängig, und es gibt eine nichtnegative Linearkombination der Spalten $M_{m+1}, \ldots, M_{m+n}$, die in $K^-(M_1, \ldots, M_m)$ liegt.

Nachdem wir nun wissen, unter welcher Voraussetzung das Programm

$$\left\{ \begin{aligned} & \min q' y \\ & \text{bezüglich} \quad My = b - Ax \\ & \qquad\qquad\quad y \geqq 0 \end{aligned} \right. \tag{5.18}$$

zu jedem b, A, x eine zulässige Lösung hat, wollen wir fragen, unter welchen Bedingungen (5.18) eine endliche optimale Lösung hat. Das ist bekanntlich dann und nur dann der Fall, wenn sowohl das primäre Programm (5.18) als auch das dazu duale Programm

$$\left\{ \begin{aligned} & \max (b - Ax)' z \\ & \text{bezüglich} \quad M' z \leqq q \end{aligned} \right. \tag{5.19}$$

zulässige Lösungen haben. Ob (5.19) zulässige Lösungen hat oder nicht, hängt offenbar nur von M und q, jedoch nicht von A, b und x ab. Wir haben also wieder die schon im vorigen Paragraphen geschilderte Situation: Genügt M zwar der Voraussetzung **V** 2, aber lassen M und q keine zulässige Lösung von (5.19) zu, so hat (5.18) zu jedem A und b eine unendliche Lösung, und (1.7) wird damit sinnlos. Dazu läßt sich folgende Aussage beweisen:

Satz 4. *M habe $m+1$ Spalten und genüge* **V** 2. *Sei daher*

$$M_{m+1} = -\sum_{i=1}^{m} \alpha_i M_i \quad \text{mit} \quad \alpha_i > 0,\ i = 1, \ldots, m.$$

Das Programm (5.18) *hat genau dann eine endliche Lösung, wenn gilt:*

$$\sum_{i=1}^{m} \alpha_i q_i \geqq -q_{m+1}. \tag{5.20}$$

Beweis. Die Bedingung (5.20) ist notwendig. Denn wenn (5.18) eine endliche Lösung hat, dann existiert eine zulässige Lösung von (5.19), z. B. z_0, für die also

$$M_i' z_0 \leqq q_i, \qquad i = 1, \ldots, m+1 \tag{5.21}$$

gilt. Aus (5.21) folgt

$$M_{m+1}' z_0 = -\sum_{i=1}^{m} \alpha_i M_i' z_0 \leqq q_{m+1} \tag{5.22}$$

und

$$\sum_{i=1}^{m} \alpha_i M_i' z_0 \leqq \sum_{i=1}^{m} \alpha_i q_i. \tag{5.23}$$

Aus (5.22) und (5.23) folgt (5.20).

Die Bedingung (5.20) ist aber auch hinreichend. Wir zeigen: Hat (5.19) keine zulässige Lösung, dann ist (5.20) verletzt. Daraus folgt: Ist (5.20) erfüllt, dann muß (5.19) eine zulässige und daher wegen **V** 2 (5.18) eine endliche optimale Lösung haben.

Wir nehmen also an, (5.19) habe keine zulässige Lösung. Wegen der linearen Unabhängigkeit von $M_1, \ldots, M_m$ hat das System

$$M_i' z = q_i, \qquad i = 1, .., m$$

eine eindeutige Lösung z_0. Auf Grund unserer Annahme gilt

$$M_{m+1}' z_0 > q_{m+1}.$$

Daraus folgt

$$-\sum_{i=1}^{m} \alpha_i M_i' z_0 = -\sum_{i=1}^{m} \alpha_i q_i > q_{m+1}.$$

Also ist (5.20) verletzt. Damit ist der Satz bewiesen.

Dieser Satz läßt sich nur teilweise für Matrizen M mit mehr als $m+1$ Spalten verallgemeinern.

Satz 5. *M habe $m+n$ Spalten ($n > 1$) und genüge* **V** 2. *Es gelte daher*

$$\sum_{i=m+1}^{m+n} \beta_i M_i = -\sum_{i=1}^{m} \alpha_i M_i, \quad \beta_i \geqq 0,\ \alpha_i > 0.$$

Dafür, daß (5.18) *eine endliche Lösung hat, ist folgende Bedingung notwendig:*

$$\sum_{i=1}^{m} \alpha_i q_i \geqq -\sum_{i=m+1}^{m+n} \beta_i q_i . \tag{5.24}$$

Beweis. Hat (5.18) eine endliche Lösung, so hat (5.19) eine zulässige Lösung z_0:

$$M_i' z_0 \leqq q_i, \qquad i = 1, \ldots, m+n .$$

Daraus folgt

$$\sum_{i=m+1}^{m+n} \beta_i M_i' z_0 \leqq \sum_{i=m+1}^{m+n} \beta_i q_i$$

und

$$\sum_{i=m+1}^{m+n} \beta_i M_i' z_0 = -\sum_{i=1}^{m} \alpha_i M_i' z_0 \geqq -\sum_{i=1}^{m} \alpha_i q_i ,$$

also

$$-\sum_{i=1}^{m} \alpha_i q_i \leqq \sum_{i=m+1}^{m+n} \beta_i q_i ,$$

was zu beweisen war.

Daß Bedingung (5.24) im allgemeinen nicht hinreichend ist, zeigt das folgende Beispiel.

Sei
$$M = \begin{pmatrix} 1 & 1 & -1 & -1 \\ -1 & 1 & 2 & -2 \end{pmatrix} = (M_1, M_2, M_3, M_4) .$$

Offenbar gilt

$$M_3 + M_4 = -(M_1 + M_2) .$$

Nach der Bezeichnungsweise in Satz 5 ist also

$$\beta_3 = \beta_4 = \alpha_1 = \alpha_2 = 1 ,$$

und M genügt **V 2.** Sei ferner

$$q_1 = q_2 = q_3 = 1 , \; q_4 = -2 .$$

Folglich ist

$$\begin{aligned} \alpha_1 q_1 + \alpha_2 q_2 &= 2, \\ \beta_3 q_3 + \beta_4 q_4 &= -1 . \end{aligned}$$

Damit ist (5.24) erfüllt. Die Restriktionen von (5.19) lauten hier

$$\begin{array}{rlrl} z_1 - z_2 \leqq 1, & \text{oder auch} & 1.\; z_1 \leqq & 1 + z_2, \\ z_1 + z_2 \leqq 1, & & 2.\; z_1 \leqq & 1 - z_2, \\ -z_1 + 2z_2 \leqq 1, & & 3.\; z_1 \geqq & -1 + 2z_2, \\ -z_1 - 2z_2 \leqq -2, & & 4.\; z_1 \geqq & 2 - 2z_2. \end{array}$$

Sei $z_2 \geqq 1$. Dann folgt aus den Restriktionen

$$\begin{aligned} &2.\; z_1 \leqq 1 - z_2 \leqq 0, \\ &3.\; z_1 \geqq -1 + 2z_2 \geqq 1 . \end{aligned}$$

Für $z_2 < 1$ ergibt sich aus

$$\begin{aligned} &2.\; z_1 \leqq 1 - z_2, \\ &4.\; z_1 \geqq 2(1 - z_2) > 1 - z_2 . \end{aligned}$$

Damit würde gelten

$$z_1 > 1 - z_2 \geqq z_1 .$$

In beiden Fällen ($z_2 \geqq 1$ und $z_2 < 1$) kommen wir also auf Grund der Restriktionen zu einem Widerspruch. Daher hat hier (5.19) keine zulässige und folglich (5.18) keine endliche Lösung.

Auf Grund dieser Sachlage wird man sich in den in Satz 5 behandelten Fällen von Mal zu Mal eine geeignete hinreichende Bedingung verschaffen müssen (wie das auch im vorigen Paragraphen in Satz 4 geschah), sofern nicht die offensichtlich hinreichende Bedingung $q \geqq 0$ erfüllt ist.

Wir haben bereits erwähnt, daß das Problem (1.7) unter der Voraussetzung **V2** sinnlos ist, wenn (5.19) keine zulässige Lösung hat. Daher machen wir in Zukunft die Voraussetzung

V3. In (1.7) erfüllt die Matrix M die Voraussetzung **V2**, und (5.19) hat eine zulässige Lösung.

Es zeigt sich nun, daß man mit **V3** mehr als nur die Endlichkeit der Lösung von (5.18) erreicht. Dazu schreiben wir die Zielfunktion von (1.7) in der Form

$$\bar{c}'x + \int (b - A\,x)'\, z_{\text{opt}}\, dP , \tag{5.25}$$

wobei P die Verteilung von (A, b, c) und $(b - A\,x)' z_{\text{opt}}$ zu jedem (A, b) der Optimalwert der Zielfunktion in (5.19) und daher gleich dem Optimalwert von $q'y$ in (5.18) ist, sofern **V3** gilt. z_{opt} hängt natürlich von A, b ab. Im übrigen sei daran erinnert, daß nach Voraussetzung die Erwartungswerte $\bar{A}, \bar{b}, \bar{c}$ existieren. Damit beweisen wir

Satz 6. *Unter der Voraussetzung* **V3** *ist die Funktion* (5.25) *für jedes* $x \in R_n$, *endlich.*

Beweis. Wir zeigen zunächst, daß der zulässige Bereich von (5.19) beschränkt ist. Sei z_0 zulässig in (5.19). Es gilt also

$$M_i' z_0 \leqq q_i \qquad i = 1, \ldots, m + n .$$

Wäre nun der zulässige Bereich von (5.19) unbeschränkt, so gäbe es ein $\Delta z_0 \neq 0$, so daß

$$M_i \Delta z_0 \leqq 0 \qquad i = 1, \ldots, m + n . \tag{5.26}$$

Nach **V3** existieren $\beta_i \geqq 0$, $\alpha_i > 0$ derart, daß

$$\sum_{i=m+1}^{m+n} \beta_i M_i = - \sum_{i=1}^{m} \alpha_i M_i \tag{5.27}$$

ist. Aus **V3** folgt außerdem, daß für mindestens ein $i \leqq m$

$$M_i' \Delta z_0 < 0 \tag{5.28}$$

gilt. Aus (5.26) bis (5.28) folgt

$$0 \geqq \sum_{i=m+1}^{m+n} \beta_i M_i' \Delta z_0 = - \sum_{i=1}^{m} \alpha_i M_i' \Delta z_0 > 0 .$$

Wegen dieses Widerspruches kann es ein solches Δz_0 nicht geben. Also ist der Bereich beschränkt. Folglich gibt es eine Zahl $\gamma \geqq 0$ derart, daß für jedes zulässige z gilt:

$$|z_i| \leqq \gamma \qquad i = 1, \ldots, m.$$

Damit gilt für (5.25) mit $z = z_{\text{opt}}$

$$(5.29)\qquad \begin{aligned} \left|\int (b - A\,x)'\,z\,dP\right| &\leqq \int |(b - A\,x)'\,z|\,dP \\ &= \int \Big|\sum_{i=1}^{m} \Big(b_i - \sum_{j=1}^{n'} A_{ij}\,x_j\Big)\,z_i\Big|\,dP \\ &\leqq \gamma \sum_{i=1}^{m} \int \Big|b_i - \sum_{j=1}^{n'} A_{ij}\,x_j\Big|\,dP. \end{aligned}$$

Da nach Voraussetzung für jedes feste $x \in R_n$ jedes der Integrale unter dem Summenzeichen existiert, d. h. einen endlichen Wert hat, ist der Satz bewiesen.

Da wir nun wissen, daß (5.25) zu jedem x einen endlichen Wert hat, können wir leicht folgenden Satz beweisen:

Satz 7. *Unter der Voraussetzung* **V3** *ist die Zielfunktion von* (1.7) *konvex in* x.

Beweis. Wir schreiben die Zielfunktion von (1.7) in der Form

$$\bar{c}'\,x + \int Q(x, A, b)\,dP$$

mit

$$Q(x, A, b) = \min q'\,y$$

bezüglich

$$\begin{aligned} M\,y &= b - A\,x \\ y &\geqq 0. \end{aligned}$$

Wir brauchen nur zu zeigen, daß $\int Q(x, A, b)\,dP$ konvex in x ist. Zunächst ist $Q(x, A, b)$ konvex in x. Seien x^1 und x^2 fest vorgegeben. Dann existieren nach **V3** endliche optimale Lösungen y^1 und y^2 mit

$$\left.\begin{aligned} Q(x^i, A, b) &= q'\,y^i \\ M\,y^i &= b - A\,x^i \\ y^i &\geqq 0 \end{aligned}\right\}\quad i = 1, 2.$$

Sei nun

$$\hat{x} = \lambda\,x^1 + (1 - \lambda)\,x^2, \quad 0 < \lambda < 1.$$

Dann gilt

$$\begin{aligned} M(\lambda\,y^1 + (1 - \lambda)\,y^2) &= b - A(\lambda\,x^1 + (1 - \lambda)\,x^2) \\ \lambda\,y^1 + (1 - \lambda)\,y^2 &\geqq 0. \end{aligned}$$

Folglich gilt

$$(5.30)\qquad Q(\hat{x}, A, b) \leqq \lambda\,q'\,y^1 + (1 - \lambda)\,q'\,y^2 = \lambda\,Q(x^1, A, b) + (1 - \lambda)\,Q(x^2, A, b).$$

Also ist $Q(x, A, b)$ konvex in x für jedes (A, b).

Aus (5.30) folgt nach einem bekannten Satz der Integrationstheorie

$$\int Q(\lambda\,x^1 + (1 - \lambda)\,x^2, A, b)\,dP \leqq \lambda \int Q(x^1, A, b)\,dP + (1 - \lambda) \int Q(x^2, A, b)\,dP,$$

und das wollten wir zeigen.

§ 6. Zur Differenzierbarkeit der Zielfunktion

Wir wissen nun, daß es sich bei (1.7) unter vernünftigen Voraussetzungen um ein konvexes Programm handelt. Zum Lösen solcher Programme ist es sehr nützlich, daß die Zielfunktion differenzierbar ist. Unter einer zusätzlichen Voraussetzung über die Wahrscheinlichkeitsverteilung $P(A, b)$ wollen wir das zunächst für den in § 4 behandelten Sonderfall beweisen.

Satz 1. *Sei in* (1.7) $M = (E, -E)$, *und sei* **V3** *erfüllt* (E *ist die* m*-reihige Einheitsmatrix*). *Ist das Wahrscheinlichkeitsmaß* $P(A, b)$ *absolut stetig in bezug auf das Lebesgue-Maß* $\mu(A, b)$, *in Zeichen* $P \ll \mu$, *dann ist die Zielfunktion in* (1.7) *überall stetig differenzierbar nach* x_i, $i = 1, \ldots, n$.

Beweis. Das Notprogramm hat das duale Programm

$$(6.1)\quad \left\{\begin{array}{ll} & \max \sum_{i=1}^{m} (b_i - A_i x) z_i \\ \text{bezüglich} & \\ & -q_i^- \leqq z_i \leqq q_i^+ \qquad i = 1, \ldots, m. \end{array}\right.$$

Darin ist A_i die i-te Zeile von A. Nach Voraussetzung **V3** hat (6.1) eine zulässige Lösung; also gilt

$$\tilde{q}_i = q_i^+ + q_i^- \geqq 0, \qquad i = 1, \ldots, m.$$

Sei z^* eine optimale Lösung bei gegebenem A, b, x. Da wir in (6.1) maximieren, muß gelten:

$$(6.2)\qquad z_i^* = \begin{cases} q_i^+ & \text{falls} \quad b_i - A_i x > 0 \\ -q_i^- & \text{falls} \quad b_i - A_i x < 0 \end{cases} \qquad i = 1, \ldots, m.$$

In (6.2) sind die Komponenten von z^* noch unbestimmt, für die $b_i - A_i x = 0$ gilt. Da die Punkte (A, b), für die $b_i - A_i x = 0$ gilt, bei festem x eine μ-Nullmenge und daher nach Voraussetzung eine P-Nullmenge bilden, können wir (6.2) ergänzen zu

$$z_i^* = \begin{cases} q_i^+ & \text{falls} \quad b_i - A_i x > 0 \\ -q_i^- & \text{falls} \quad b_i - A_i x \leqq 0 \end{cases} \qquad i = 1, \ldots, m.$$

Die Zielfunktion von (1.7) hat damit folgende Gestalt:

$$\begin{aligned}(6.3)\qquad \Phi(x) &= \int c' x \, dP(A, b, c) + \\ &+ \sum_{i=1}^{m} \Big\{ \int\limits_{b_i - A_i x > 0} (b_i - A_i x) q_i^+ \, dP(A, b) - \int\limits_{b_i - A_i x \leqq 0} (b_i - A_i x) q_i^- \, dP(A, b) \Big\} \\ &= \bar{c}' x + \sum_{i=1}^{m} (\bar{b}_i - \bar{A}_i x) q_i^+ - \sum_{i=1}^{m} \tilde{q}_i \int\limits_{b_i - A_i x \leqq 0} (b_i - A_i x) \, dP\end{aligned}$$

Wenn jedes der Integrale $J_i(x) = \int\limits_{b_i - A_i x \leqq 0} (b_i - A_i x) \, dP$ stetig nach $x_j, j = 1, \ldots, n$, differenzierbar ist, dann gilt dasselbe offenbar für $\Phi(x)$. Sei $\Delta x_j \neq 0$ und

$$\Delta x = (0, \ldots, \Delta x_j, \ldots, 0)'.$$

Wir definieren

$$\begin{aligned} \mathfrak{A} &= \{A, b \mid b_i - A_i x \leqq A_i \Delta x\} \\ \mathfrak{B} &= \{A, b \mid b_i - A_i x \leqq 0\}. \end{aligned}$$

Damit ist

$$\begin{aligned}J_i(x+\Delta x)-J_i(x)&=\int_{\mathfrak{A}}(b_i-A_ix-A_i\Delta x)\,dP-\int_{\mathfrak{B}}(b_i-A_ix)\,dP\\&=-\int_{\mathfrak{A}\cap\mathfrak{B}}A_{ij}\Delta x_j\,dP+\int_{\mathfrak{A}\cap\overline{\mathfrak{B}}}(b_j-A_ix-A_i\Delta x)\,dP-\\&-\int_{\overline{\mathfrak{A}}\cap\mathfrak{B}}(b_i-A_ix)\,dP=-\int_{\mathfrak{B}}A_{ij}\Delta x_j\,dP+\\&+\int_{\mathfrak{A}\cap\overline{\mathfrak{B}}}(b_i-A_ix-A_i\Delta x)\,dP-\int_{\overline{\mathfrak{A}}\cap\mathfrak{B}}(b_i-A_ix-A_i\Delta x)\,dP.\end{aligned}$$

Daher ist

$$\frac{J_i(x+\Delta x)-J_i(x)}{\Delta x_j}=-\int_{\mathfrak{B}}A_{ij}\,dP+R(\Delta x_j)$$

mit

$$R(\Delta x_j)=\int_{\mathfrak{A}\cap\overline{\mathfrak{B}}}\left(\frac{b_i-A_ix}{\Delta x_j}-A_{ij}\right)dP-\int_{\overline{\mathfrak{A}}\cap\mathfrak{B}}\left(\frac{b_i-A_ix}{\Delta x_j}-A_{ij}\right)dP.$$

Offenbar ist

$$\text{(6.4)}\qquad\begin{cases}\overline{\mathfrak{A}}\cap\mathfrak{B}=\{A,b\,|\,A_{ij}\Delta x_j<b_i-A_ix\leqq 0\}\\ \mathfrak{A}\cap\overline{\mathfrak{B}}=\{A,b\,|\,0<b_i-A_ix\leqq A_{ij}\Delta x_j\}.\end{cases}$$

Auf Grund von (6.4) können wir nun $R(\Delta x_j)$ abschätzen, wobei wir zwei Fälle unterscheiden:

a) $\Delta x_j>0$.

Dann gilt

$$R(\Delta x_j)\leqq\int_{\mathfrak{A}\cap\overline{\mathfrak{B}}}(A_{ij}-A_{ij})\,dP-\int_{\overline{\mathfrak{A}}\cap\mathfrak{B}}(A_{ij}-A_{ij})\,dP=0$$

$$R(\Delta x_j)\geqq\int_{\mathfrak{A}\cap\overline{\mathfrak{B}}}(0-A_{ij})\,dP-\int_{\overline{\mathfrak{A}}\cap\mathfrak{B}}(0-A_{ij})\,dP=-\int_{\mathfrak{A}\cap\overline{\mathfrak{B}}}A_{ij}\,dP+\int_{\overline{\mathfrak{A}}\cap\mathfrak{B}}A_{ij}\,dP.$$

b) $\Delta x_j<0$.

$$R(\Delta x_j)\leqq\int_{\mathfrak{A}\cap\overline{\mathfrak{B}}}(0-A_{ij})\,dP-\int_{\overline{\mathfrak{A}}\cap\mathfrak{B}}(0-A_{ij})\,dP=-\int_{\overline{\mathfrak{A}}\cap\mathfrak{B}}A_{ij}\,dP+\int_{\mathfrak{A}\cap\overline{\mathfrak{B}}}A_{ij}\,dP$$

$$R(\Delta x_j)\geqq\int_{\mathfrak{A}\cap\overline{\mathfrak{B}}}(A_{ij}-A_{ij})\,dP-\int_{\overline{\mathfrak{A}}\cap\mathfrak{B}}(A_{ij}-A_{ij})\,dP=0.$$

Nun müssen wir beachten, daß die in (6.4) definierten Mengen noch von Δx_j abhängen. Offenbar gilt

$$\lim_{\Delta x_j\to 0}\overline{\mathfrak{A}}\cap\mathfrak{B}=\{A,b\,|\,b_i-A_ix=0\}=\mathfrak{C}$$

$$\lim_{\Delta x_j\to 0}\mathfrak{A}\cap\overline{\mathfrak{B}}=\{A,b\,|\,0<b_i-A_ix\leqq 0\}=\mathfrak{D}.$$

$\mathfrak{C}$ ist eine Hyperebene des $R(A,b)$, und $\mathfrak{D}$ ist leer. Folglich gilt $\mu(\mathfrak{C})=\mu(\mathfrak{D})=0$ und nach Voraussetzung daher $P(\mathfrak{C})=P(\mathfrak{D})=0$. Daraus folgt sowohl unter a) als auch unter b) $\lim\limits_{\Delta x_j\to 0}R(\Delta x_j)=0$. Folglich ist $J_i(x)$ differenzierbar nach x_j, $j=1,\ldots,n$, und es gilt

$$\text{(6.5)}\qquad\frac{\partial J_i(x)}{\partial x_j}=-\int_{b_i-A_ix\leqq 0}A_{ij}\,dP.$$

Daß diese Ableitung stetig in x ist, sieht man sofort: Mit beliebigem Δx ist

$$\frac{\partial J_i(x+\Delta x)}{\partial x_j} - \frac{\partial J_i(x)}{\partial x_j} = - \int\limits_{b_i - A_i x \leqq A_i \Delta x} A_{ij}\, dP + \int\limits_{b_i - A_i x \leqq 0} A_{ij}\, dP = - \int\limits_{0 < b_i - A_i x \leqq A_i \Delta x} A_{ij}\, dP + \int\limits_{A_i \Delta x < b_i - A_i x \leqq 0} A_{ij}\, dP .$$

Für $|\Delta x| \to 0$ sind die Integrationsbereiche, wie wir sahen, P-Nullmengen, so daß die beiden Integrale gegen Null streben. Damit ist der Satz vollständig bewiesen.

Es liegt nun nahe zu vermuten, daß eine zu Satz 1 analoge Aussage für das allgemeine zweistufige Problem (1.7) richtig ist. Allerdings verläuft der Beweis hier anders. Wir benötigen dazu die folgenden Sätze der Integrationstheorie:

Satz 2. *Ist $\{f_n\}$ eine Folge integrierbarer Funktionen, die dem Maße nach — oder fast überall — gegen f konvergiert, und existiert eine integrierbare Funktion $g(x)$ derart, daß $|f_n(x)| \leqq |g(x)|$ fast überall gilt für $n = 1, 2, \ldots$, dann ist f integrierbar, und $\{f_n\}$ konvergiert im Mittel gegen f.*

Den *Beweis* dieses Satzes findet man z. B. bei HALMOS: „Measure Theory", S. 110. Auf diesen Satz stützt sich das folgende bekannte Theorem:

Satz 3. *Sei $f(x, t)$ für jedes t aus $\alpha \leqq t \leqq \beta$ integrierbar in bezug auf das beliebige Maß μ. Existiert für alle $x \in R_m$ mit Ausnahme einer μ-Nullmenge $\mathfrak{N}$ die partielle Ableitung $\frac{\partial f}{\partial t}(x, t_0)$ bei festem $t_0 \in (\alpha, \beta)$ und gibt es eine integrierbare, Funktion $h(x)$ derart, daß für alle hinreichend kleinen $|\Delta t| > 0$ stets*

$$\left| \frac{f(x, t_0 + \Delta t) - f(x, t_0)}{\Delta t} \right| \leqq h(x)$$

fast überall gilt, dann ist $\varphi(t) = \int\limits_{R_m} f(x, t)\, d\mu$ in t_0 differenzierbar, und es gilt

$$\varphi'(t_0) = \int\limits_{R_m - \mathfrak{N}} \frac{\partial f}{\partial t}(x, t_0)\, d\mu .$$

Nachdem wir nun die Hilfsmittel bereitgestellt haben, können wir den vermuteten Satz beweisen.

Satz 4. *Für das Problem* (1.7) *sei* **V3** *erfüllt. Ist $P(A, b) \ll \mu(A, b) - \mu(A, b)$ ist wieder das Lebesgue-Maß —, dann ist die Zielfunktion von* (1.7) *überall stetig differenzierbar nach x_j, $j = 1, \ldots, n$.*

Beweis. Die Zielfunktion in (1.7) lautet

$$E[c' x + Q(x, A, b)] = \bar{c}' x + E(Q(x, A, b)) .$$

mit

$$(6.6)\quad \left\{ \begin{array}{ll} & Q(x, A, b) = \min q' y \\ \text{bezüglich} & \\ & M y = b - A x \\ & y \geqq 0 . \end{array} \right.$$

Offenbar genügt es zu zeigen, daß $E(Q(x, A, b))$ stetig differenzierbar ist. Wir wählen ein festes x und zeigen zunächst, daß $E(Q(x, A, b))$ in x differenzierbar ist. Unter der Voraussetzung **V3** ist, wie wir wissen, die optimale Lösung von (6.6) zu jedem festen A, b und x endlich. Folglich gibt es eine zu dieser Lösung äquivalente Basislösung. Es existiert also eine Basis B aus m linear unabhängigen Spalten von M derart, daß der Vektor $\hat{y}$ aus den nicht zu B gehörenden Komponenten von y gleich dem Nullvektor ist, während für den Vektor $\tilde{y}$ aus den übrigen Komponenten gilt:

$$\tilde{y} = B^{-1}(b - A\,x) \geqq 0\,.$$

Setzen wir Nichtdegeneration voraus, so gilt

$$\tilde{y} = B^{-1}(b - A\,x) > 0\,.$$

In diesem Falle muß, da wir eine Optimallösung vor uns haben, das Simplex-Kriterium

$$\tilde{q}'\,B^{-1}M_i \leqq q_i \tag{6.7}$$

gelten für alle i ($\tilde{q}$ ist der Vektor aus den zu B gehörenden Komponenten von q). Wenn eine Basis der Bedingung (6.7) genügt, wollen wir sie eine Optimalbasis nennen.

Nun sondern wir aus M mit Hilfe der Bedingung (6.7) alle Optimalbasen aus und bezeichnen sie mit $B_1, \ldots, B_r$. Dazu bemerken wir, daß, wenn zu einem (A, b) die optimale Basislösung nicht als Basislösung einer dieser Optimalbasen dargestellt wird, diese Lösung degeneriert sein muß. Denn sei die entsprechende Basis $B \neq B_i$, $i = 1, \ldots, r$. Wäre nun $B^{-1}(b - A\,x) > 0$, so müßte (6.7) gelten, und folglich wäre B doch eine Optimalbasis entgegen der Annahme. Folglich muß mindestens eine Komponente von $B^{-1}(b - A\,x)$ gleich Null sein, was bedeutet, daß (A, b) in einem linearen Teilraum des Raumes $R(A, b)$, also in einer μ-Nullmenge und damit nach Voraussetzung in einer P-Nullmenge, liegt. Wir definieren folgende Mengen:

$$\begin{aligned}
&\mathfrak{B}_1 = \{A, b \mid B_1^{-1}(b - A\,x) > 0\}\\
&\mathfrak{B}_2 = \{A, b \mid B_2^{-1}(b - A\,x) > 0\} - \mathfrak{B}_1\\
&\mathfrak{B}_3 = \{A, b \mid B_3^{-1}(b - A\,x) > 0\} - (\mathfrak{B}_1 \cup \mathfrak{B}_2)\\
&\cdot\\
&\cdot\\
&\cdot\\
&\mathfrak{B}_r = \{A, b \mid B_r^{-1}(b - A\,x) > 0\} - \bigcup_{k=1}^{r-1} \mathfrak{B}_k\,.
\end{aligned}$$

Die Punkte (A, b), die zu keiner dieser Mengen $\mathfrak{B}_i$ gehören, führen, wie wir sahen, zu degenerierten Lösungen und gehören folglich einer von endlich vielen μ- und damit P-Nullmengen an.

Auf $\bigcup_{i=1}^{r} \mathfrak{B}_i$ gilt

$$Q(x, A, b) = \tilde{q}'\,B_i^{-1}(b - A\,x)\,, \quad \text{falls} \quad (A, b) \in \mathfrak{B}_i\,. \tag{6.8}$$

$\tilde{q}$ ist hier offenbar der Vektor aus den zu B_i gehörenden Komponenten von q. Sei

$$\Delta x = (0, \ldots, \Delta x_j, \ldots, 0)' \quad \text{und} \quad \Delta x_j \neq 0\,.$$

Offenbar ist $Q(x, A, b)$ in x_j differenzierbar, falls $(A, b) \in \mathfrak{B}_i$, $i = 1, \ldots, r$. Es gilt nach (6.8), wenn A_j die j-te Spalte von A ist,

$$\frac{\partial Q}{\partial x_j}(x, A, b) = -\,\tilde{q}'\,B_i^{-1}A_j \quad \text{falls} \quad (A, b) \in \mathfrak{B}_i\,,\ i = 1, \ldots, r\,. \tag{6.9}$$

Es existiert also für alle (A, b), die nicht in $\mathfrak{N} = R(A, b) - \bigcup_{i=1}^{r} \mathfrak{B}_i$ liegen, die partielle Ableitung $\frac{\partial Q}{\partial x_j}(x, A, b)$, und es gilt $P(\mathfrak{N}) = 0$.

Nun müssen wir die Differenzenquotienten abschätzen. Dazu definieren wir mit

$$d = b - A x$$

die Mengen

$$\mathfrak{B}_i' = \{d \mid B_i^{-1} d \geqq 0\} \qquad i = 1, \ldots, r.$$

Für diese Mengen gilt:

1. $\mathfrak{B}_i'$ ist konvex.

2. $\bigcup_{i=1}^{r} \mathfrak{B}_i' = R_m$, und jeder Randpunkt von $\mathfrak{B}_i'$ ist Element mindestens einer anderen Menge $\mathfrak{B}_j'$, $i \neq j$.

Die erste Behauptung ist offensichtlich richtig. Die zweite ergibt sich ohne weiteres aus der Voraussetzung **V3**.

Nun ist

$$b - A x - A_j \Delta x_j = d - A_j \Delta x_j = d^*(\Delta x_j).$$

Gehört d zu einer bestimmten Menge $\mathfrak{B}_{i_1}'$ so bestimmen wir das maximale $|\Delta x_j^1| \leqq |\Delta x_j|$ derart, daß $d^*(\Delta x_j^1)$ noch zu $\mathfrak{B}_{i_1}'$ gehört, wobei sign $\Delta x_j^1 =$ sign Δx_j sein soll. Entweder ist $\Delta x_j^1 = \Delta x_j$, oder $|\Delta x_j^1| < |\Delta x_j|$ und $d^*(\Delta x_j^1)$ ist ein Randpunkt von $\mathfrak{B}_{i_1}'$. Nach 2. gehört dann $d^*(\Delta x_j^1)$ zu einer anderen Menge $\mathfrak{B}_{i_2}'$, auf der wir die eben geschilderte Prozedur wiederholen. Wegen der Konvexität der $\mathfrak{B}_i'$ müssen wir nach $r_1 \leqq r$ Schritten zu dem Resultat $\Delta x_j^{r_1} = \Delta x_j$ gelangen, wobei also $d^*(\Delta x_j) \in \mathfrak{B}_{i_{r_1}}'$ ist. Beachten wir noch, daß auf dem Durchschnitt von $\mathfrak{B}_i'$ und $\mathfrak{B}_j'$ die den beiden Mengen entsprechenden Basislösungen den gleichen Wert der Zielfunktion $Q(x, A, b)$ ergeben müssen, dann ergibt sich für den Differenzenquotienten mit $\Delta x_i^0 = 0$

$$\begin{aligned}
\left|\frac{Q(x+\Delta x, A, b) - Q(x, A, b)}{\Delta x_j}\right| &= \frac{\left|\tilde{q}' B_{i_{r_1}}^{-1} d^*(\Delta x_j) - \tilde{q}' B_{i_1}^{-1} d^*(0)\right|}{|\Delta x_j|} \\
&= \frac{\left|\sum_{\nu=1}^{r_1} [\tilde{q}' B_{i_\nu}^{-1} d^*(\Delta x_j^\nu) - \tilde{q}' B_{i_\nu}^{-1} d^*(\Delta x_j^{\nu-1})]\right|}{|\Delta x_j|} \\
&= \frac{\left|\sum_{\nu=1}^{r_1} -\tilde{q}' B_{i_\nu}^{-1} (\Delta x_j^\nu - \Delta x_j^{\nu-1}) A_j\right|}{|\Delta x_j|} \\
&\leqq |A_j| \cdot \sum_{\nu=1}^{r_1} |\tilde{q}' B_{i_\nu}^{-1}| \leqq |A_j| \cdot \sum_{i=1}^{r} |\tilde{q}' B_i^{-1}| \\
&\leqq \sum_{k=1}^{m} |A_{kj}| \cdot \sum_{i=1}^{r} |\tilde{q}' B_i^{-1}|.
\end{aligned} \tag{6.10}$$

Da diese Abschätzung für jeden Punkt $(A, b) \in R(A, b)$ gilt, und da nach Voraussetzung die A_{ij} und daher

$$\sum_{i=1}^{r} |\tilde{q}' B_i^{-1}| \cdot \sum_{i=1}^{m} |A_{ij}|$$

integrierbar sind in bezug auf P, und da, wie wir wissen, $Q(x, A, b)$ für jedes x P-integrierbar ist, können wir Satz 3 anwenden und haben damit die Differenzierbarkeit von $Q(x) = E(Q(x, A, b))$ bewiesen. Es gilt dann nach Satz 3 und (6.9)

$$\frac{\partial Q}{\partial x_j}(x) = -\sum_{i=1}^{r} \int_{\mathfrak{B}_i} \tilde{q}' B_i^{-1} A_j \, dP, \quad j = 1, \ldots, n. \tag{6.11}$$

Um die Stetigkeit zu beweisen, müssen wir beachten, daß nach Definition $\mathfrak{B}_i$ von x abhängt. Mit einem beliebigen Vektor Δx ist daher

$$\frac{\partial Q}{\partial x_j}(x + \Delta x) - \frac{\partial Q}{\partial x_j}(x) = -\sum_{i=1}^{r} \left[\int_{\mathfrak{B}_i(x+\Delta x)} \tilde{q}' B_i^{-1} A_j \, dP - \int_{\mathfrak{B}_i(x)} \tilde{q}' B_i^{-1} A_j \, dP \right]. \tag{6.12}$$

Analog wie früher strebt die symmetrische Differenz der Mengen $\mathfrak{B}_i(x + \Delta x)$ und $\mathfrak{B}_i(x)$, also $(\mathfrak{B}_i(x + \Delta x) \cap \overline{\mathfrak{B}}_i(x)) \cup (\overline{\mathfrak{B}}_i(x + \Delta x) \cap \mathfrak{B}_i(x))$, mit $|\Delta x| \to 0$ gegen eine Menge vom μ-Maß Null, also auch gegen eine P-Nullmenge. Daraus folgt die Stetigkeit der Ableitung sofort.

§ 7. Ungleichungen

Man kann nun auf die Idee kommen — und das geschieht ja praktisch in den meisten linearen Programmen —, die zufälligen Variablen in (1.7) durch ihre Erwartungswerte zu ersetzen. Wenn das Optimum des dadurch entstehenden linearen Programms endlich ist, dann hat man zwar im allgemeinen keine Lösung des ursprünglichen Problems (1.7). Man erhält jedoch dann die Möglichkeit, den Optimalwert der Zielfunktion von (1.7) nach oben und unten abzuschätzen. Dazu beweisen wir folgenden

Satz 1. *Sei $x \in R_m$ ein Vektor zufälliger Variabler, deren Erwartungswerte existieren. Sei $y \in \mathfrak{B}$ ein Vektor von Parametern, und sei $f(x, y)$ für jedes (x, y) endlich und für jedes feste y konvex in x. Existiert $E[f(x, y)]$ für jedes y, und existiert $f(\bar{x}, \tilde{y}) = \min_{y \in \mathfrak{B}} f(\bar{x}, y)$ mit $\bar{x} = E(x)$, dann gilt*

$$f(\bar{x}, \tilde{y}) \leqq \inf_{y \in \mathfrak{B}} E f(x, y) \leqq E f(x, \tilde{y}).$$

Beweis. Sei $Q_n = \{x \mid -n \leqq x_j < n;\ j = 1, \ldots, m\}$. Wir definieren eine Folge vektorwertiger Funktionen $\{x^{n, \nu(n)}\}$ folgendermaßen:

$$x_j^{n, \nu(n)} = \begin{cases} \dfrac{i}{2^{\nu(n)}} & \text{falls } \dfrac{i}{2^{\nu(n)}} \leqq x_j < \dfrac{i+1}{2^{\nu(n)}};\ x \in Q_n; \\ & \qquad i = -n \cdot 2^{\nu(n)}, \ldots, n \cdot 2^{\nu(n)} - 1 \\ 0 & \text{sonst} \end{cases} \tag{7.1}$$

für $j = 1, \ldots, m$. Über $\nu(n)$ setzen wir vorläufig nur voraus, daß $\nu(n)$ eine natürliche Zahl ist, und daß $\nu(n+1) > \nu(n)$ gilt. Offenbar konvergiert $\{x^{n, \nu(n)}\}$ überall gegen x für $n \to \infty$. Ferner gilt

$$|x_j^{n, \nu(n)} - x_j| < \frac{1}{2^{\nu(n)}} \qquad \text{falls} \quad x \in Q_n$$

und

$$|x_j^{n, \nu(n)} - x_j| = |x_j| \quad \text{sonst.}$$

Also gilt überall

$$|x_j^{n,\nu(n)}| \leqq |x_j| + 1 .$$

Da nach Voraussetzung $|x_j| + 1$ integrierbar ist, konvergiert nach § 6, Satz 2 $\{x_j^{n,\nu(n)}\}$ im Mittel gegen x_j, d. h. $\{\int x_j^{n,\nu(n)}\,dP\}$ konvergiert gegen $\int x\,dP = \bar{x}$.

Zu festem $y \in \mathfrak{B}$ definieren wir nun die Treppenfunktionen

$$f^n(x, y) = \begin{cases} f(x^{n,\nu(n)}, y) & \text{falls} \quad x \in Q_n \\ f(0, y) & \text{sonst.} \end{cases} \tag{7.2}$$

Auf Grund der Voraussetzungen ist $f(x, y)$ überall stetig in x. Daher konvergiert $\{f^n(x, y)\}$ überall in R_m gegen $f(x, y)$ für $n \to \infty$. Ferner kann man wegen der Stetigkeit zu jedem n die Zahl $\nu(n)$ so groß wählen, daß zu festem $\varepsilon > 0$ gilt:

$$|f(x, y) - f^n(x, y)| \leqq \varepsilon \qquad \text{falls} \quad x \in Q_n .$$

Außerdem gilt

$$|f^n(x, y) - f(x, y)| \leqq |f(x, y)| + |f(0, y)| \qquad \text{falls} \quad x \notin Q_n .$$

Damit haben wir bei passender Wahl von $\nu(n)$ allgemein

$$|f^n(x, y)| \leqq |f(x, y)| + |f(0, y)| + \varepsilon .$$

Da nach Voraussetzung $|f(x, y)|$ P-integrierbar ist und das gleiche für die Konstante $|f(0, y)| + \varepsilon$ gilt, konvergiert wieder nach § 6, Satz 2 $\{f^n(x, y)\}$ im Mittel gegen $f(x, y)$.

Durch die Definition von x^n (die Zahl $\nu(n)$ können wir nun fortlassen, da sie durch n bestimmt ist) wird der Quader Q_n in endlich viele, z. B. $r(n)$, Quader Q_n^j, $j = 1, \ldots, r(n)$, unterteilt, aus denen nach (7.1) je ein Eckpunkt $x^{n,j}$ ausgewählt wird. Die Menge der Punkte x, die nicht zu Q_n gehören, bezeichnen wir mit Q_n^0 und setzen $x^{n,0} = 0$.

Wegen der Konvexität von $f(x, y)$ in x gilt dann

$$\begin{aligned} \int f^n(x, y)\,dP &= \sum_{j=0}^{r(n)} f(x^{n,j}, y)\, P(Q_n^j) \\ &\geqq f\Big(\sum_{j=0}^{r(n)} x^{n,j} P(Q_n^j), y\Big) \\ &= f\Big(\int x^n\,dP, y\Big) . \end{aligned}$$

Daraus und aus der Stetigkeit von f folgt dann

$$\begin{aligned} \int f(x, y)\,dP &= \lim_{n\to\infty} \int f^n(x, y)\,dP \geqq \lim_{n\to\infty} f\Big(\int x^n\,dP, y\Big) \\ &= f\Big(\lim_{n\to\infty} \int x^n\,dP, y\Big) = f\Big(\int x\,dP, y\Big) , \end{aligned}$$

also

$$f(E(x), y) \leqq E f(x, y) . \tag{7.3}$$

Nun gilt für jedes $y \in \mathfrak{B}$

$$\min_{y \in \mathfrak{B}} f(E x, y) = f(E x, \tilde{y}) \leqq f(E x, y) .$$

Nach (7.3) gilt dann für jedes $y \in \mathfrak{B}$

$$f(Ex, \tilde{y}) \leqq E f(x, y)$$

und folglich

$$f(Ex, \tilde{y}) \leqq \inf_{y \in \mathfrak{B}} E f(x, y).$$

Das ist der erste Teil der Behauptung. Da der zweite Teil

$$\inf_{y \in \mathfrak{B}} E f(x, y) \leqq E f(x, \tilde{y})$$

trivial ist, ist der Satz bewiesen.

Diesen Satz wollen wir nun auf unser Problem (1.7) anwenden. Dazu bezeichnen wir die Zielfunktion mit

$$Q(x, A, b, c) = c' x + \min (q' y \mid M y = b - A x, y \geqq 0).$$

Wir haben vorausgesetzt, daß $\bar{A}$, $\bar{b}$, $\bar{c}$ existieren und wissen, daß unter der Voraussetzung **V 3** die Funktion $Q(x, A, b, c)$ für jedes (x, A, b, c) endlich ist, und daß $E\, Q(x, A, b, c)$ für jedes x existiert. Nun zeigen wir

Satz 2. *Sei* **V 3** *erfüllt. Dann ist* $Q(x, A, b, c)$ *konvex in* (A, b, c) *für jedes* x.

Beweis. Wir brauchen nur zu zeigen, daß

$$Q(x, A, b) = \min (q' y \mid M y = b - A x, y \geqq 0) \tag{7.4}$$

konvex in A, b ist. Dazu wählen wir ein festes x, zwei Punkte (A^1, b^1), (A^2, b^2) und ein λ mit $0 < \lambda < 1$. Zu (A^i, b^i) mögen die Lösungen $y^i (i = 1, 2)$ gehören. Nun betrachten wir $(\tilde{A}, \tilde{b}) = \lambda (A^1, b^1) + (1 - \lambda)(A^2, b^2)$. Für $\tilde{y} = \lambda y^1 + (1 - \lambda) y^2$ gilt, da die y^i Lösungen von (7.4) mit $(A, b) = (A^i, b^i)$, $i = 1, 2$, sind,

$$M \tilde{y} = \lambda (b^1 - A^1 x) + (1 - \lambda)(b^2 - A^2 x) = \tilde{b} - \tilde{A} x$$

und

$$\tilde{y} \geqq 0.$$

Folglich gilt

$$q' \tilde{y} \geqq Q(x, \tilde{A}, \tilde{b}).$$

Das bedeutet aber

$$Q(x, \tilde{A}, \tilde{b}) \leqq \lambda\, Q(x, A^1, b^1) + (1 - \lambda)\, Q(x, A^2, b^2),$$

womit die Behauptung bewiesen ist.

Da nun alle Voraussetzungen gegeben sind, können wir ohne weiteres Satz 1 anwenden, und erhalten

Satz 3. *In* (1.7) *sei die Voraussetzung* **V 3** *erfüllt. Ferner existiere*

$$\min_{x \geqq 0} Q(x, \bar{A}, \bar{b}, \bar{c}) = Q(\tilde{x}, \bar{A}, \bar{b}, \bar{c}).$$

Dann gilt

$$\min_{x \geqq 0} Q(x, \bar{A}, \bar{b}, \bar{c}) \leqq \inf_{x \geqq 0} E\, Q(x, A, b, c) \leqq E\, Q(\tilde{x}, A, b, c).$$

Literatur

[1] HALMOS, P. R.: Measure theory. Sixth printing. Princeton, N. J.: D. van Nostrand Company, Inc. 1959.

[2] RICHTER, H. R.: Wahrscheinlichkeitstheorie. Berlin-Göttingen-Heidelberg: Springer 1956.

[3] KÜNZI, H. P., u. W. KRELLE: Nichtlineare Programmierung. Berlin-Göttingen-Heidelberg: Springer 1962.

[4] DANTZIG, G. B.: Linear programming and extensions. Princeton: University Press 1963.

[5] — Linear programming under uncertainty. Management Sci. **1**, 197—206 (1955).

[6] ELMAGHRABY, S. E.: An approach to linear programming under uncertainty. Operations Res. **7**, 208—216 (1959).

[7] MADANSKY, A.: Inequalities for stochastic linear programming problems. Management Sci. **6**, 197—204 (1960).

[8] — Methods of solution of linear programs under uncertainty. Operations Res. **10**, 463—471 (1962).

[9] WETS, R.: Programming under uncertainty: the complete problem. Boeing Scientific Research Laboratories, Mathematics Research Laboratory, Mathematical Note No. 370, October 1964.

[10] — Programming under uncertainty: the equivalent convex program. J. SIAM Appl. Math. **14**, 89—105 (1966).

[11] WALKUP, D. W., and R. WETS: Stochastic programs with recourse. Boeing document D 1-82-0551, Boeing Scientific Research Laboratories, Seattle, Washington (1966).

Seminar für Angewandte Mathematik
und Statistik der Universität
CH-8006 Zürich (Schweiz)