Eine allgemeine, symmetrische Formulierung des Dekompositionsprinzips für duale Paare nichtlinearer minmax- und maxmin-Probleme

W. Oettli

[47]

Zusammenfassung: Es wird eine allgemeine Formulierung des Dekompositionsverfahrens zur Lösung konvex-konkaver Sattelpunktprobleme angegeben.

Januar 1973

Universität Mannheim, Fakultät für Mathematik und Informatik,
D-68 Mannheim, Schloss.

Im folgenden unternehmen wir den Versuch, das Prinzip der Dekomposition nichtlinearer, konvexer Programme, wie es von Dantzig [1], Benders [2], Huard [3], Kronsjö [4] und anderen [5] beschrieben wurde, auf allgemeine konvex-konkave minmax-Probleme zu übertragen. Dabei erreichen wir eine symmetrische Formulierung, d.h. es spielt keine Rolle, ob wir das Verfahren auf das gegebene primale minmax-Problem oder das dazu duale maxmin-Problem ansetzen: Die Iterationsvorschrift ist die gleiche, und das Verfahren löst beide Probleme gleichzeitig. Durch Spezialisierung erhalten wir dann eine unsymmetrische, primale Variante, die die Verfahren von Dantzig und Huard als Spezialfall enthält. Wir interessieren uns besonders für die geometrische Interpretation. So zeigen wir etwa, dass das primale Dekompositionsverfahren zwischen dem Schnittverfahren zur Lösung des dualen Problemes und dem Verfahren der zulässigen Richtungen zur Lösung des primalen Problems eingeordnet werden kann. Wir interessieren uns nicht für die Anwendung des Dekompositionsprinzipes auf bestimmte struktuierte Probleme.

Im folgenden Abschnitt stellen wir einige Hilfsmittel zusammen und geben einen Ueberblick über bereits bekannte Resultate.

1. Hilfsmittel

Dualität: Sei $\varphi(x,y) : X \times Y \to \mathbb{R}$ eine auf einem cartesischen Produkt definierte reellwertige Funktion. Mittels der beiden Funktionen

$$(1) \qquad M(x) = \sup_{y \in Y} \varphi(x,y) : X \to \mathbb{R} \cup \{+\infty\} ,$$

$$(2) \qquad m(y) = \inf_{x \in X} \varphi(x,y) : Y \to \mathbb{R} \cup \{-\infty\}$$

können wir ein duales Paar von Optimierungsproblemen bilden (Stoer [6]),

$$(3) \qquad P : \inf \{M(x) \mid x \in X, M(x) < +\infty\} \text{ ("primal problem")},$$

$$(4) \qquad D : \sup \{m(y) \mid y \in Y, m(y) > -\infty\} \text{ ("dual problem")}.$$

Wir sagen, $\hat{x}$ sei eine Lösung von P, oder auch, $\hat{x}$ sei eine Lösung von $\inf\limits_X \sup\limits_Y \varphi(x,y)$, geschrieben

$$\hat{x} : \inf_X \sup_Y \varphi(x,y) \quad ,$$

wenn $\hat{x} \in X$ und $\sup\limits_{y \in Y} \varphi(\hat{x},y) \leq \sup\limits_{y \in Y} \varphi(x,y) \; \forall x \in X$, und wenn $\sup\limits_{y \in Y} \varphi(\hat{x},y) < +\infty$. Offensichtlich gilt die grundlegende Ungleichung

$$(5) \qquad M(x) \geq \varphi(x,y) \geq m(y) \quad \forall x \in X, \forall y \in Y .$$

$(\hat{x},\hat{y}) \in X \times Y$ heisst *Sattelpunkt* von $\varphi(x,y)$ über $X \times Y$, wenn

$$M(\hat{x}) = m(\hat{y}) \quad .$$

Dies ist gleichwertig mit

$$\varphi(\hat{x},y) \leq \varphi(\hat{x},\hat{y}) \leq \varphi(x,\hat{y}) \; \forall x \in X, \forall y \in Y \quad .$$

Wir zeigen in der Regel die Sattelpunkteigenschaft von $(\hat{x},\hat{y})$, indem wir die zu (5) entgegengesetzte Ungleichung $M(\hat{x}) \leq m(\hat{y})$ verfizieren. Aus (5) folgt sofort

$$(6) \qquad (\hat{x},\hat{y}) \text{ ist Sattelpunkt} \Rightarrow \hat{x} \text{ löst } P, \; \hat{y} \text{ löst } D .$$

Offenbar gilt in diesem Fall auch für jede andere Lösung $\tilde{x}$ bzw. $\tilde{y}$ von P bzw. D, dass $(\tilde{x},\tilde{y})$ ein Sattelpunkt ist. Wenn also φ *überhaupt* Sattelpunkte hat, so ist die Menge aller Sattelpunkte von der Form $\hat{X} \times \hat{Y}$, wobei $\hat{X}$ die Menge aller Lösungen von P und $\hat{Y}$ die Menge aller Lösungen von D ist. Ein sehr handliches und für unsere Zwecke ausreichendes Kriterium für die Existenz von Sattelpunkten gibt der

Satz von Kakutani. - Die Mengen $X \subseteq \mathbb{R}^k$ und $Y \subseteq \mathbb{R}^m$ seien nichtleer, konvex und kompakt. $\varphi(x,y)$ sei stetig in $X \times Y$ und konvex-konkav (d.h. konvex in x für festes y, und konkav in y für festes x). Dann hat φ einen Sattelpunkt über $X \times Y$.

Wir betrachten den Spezialfall, dass

(7) $$\varphi(x,y) = F(x) + y^T f(x)\ , \quad Y = \{y \in \mathbb{R}^m | y \geq 0\} \equiv \mathbb{R}^m_+ \ ,$$

wo $f : X \to \mathbb{R}^m$. Dann ist

$$M(x) = \begin{cases} F(x), \text{ wenn } f(x) \leq 0, \\ + \infty \quad \text{sonst,} \end{cases}$$

und P wird zu einem gewöhnlichen Programm der Form

(8) $$P : \min \ \{F(x) | x \in X,\ f(x) \leq 0\} \ .$$

φ gemäss (7) heisst *Lagrangefunktion* zum Programm (8). Wenn

$$\varphi(x,y) = c^T x + y^T b + y^T Ax\ , \quad X = \mathbb{R}^k_+\ , \quad Y = \mathbb{R}^m_+ \ ,$$

so werden P und D zu dualen linearen Programmen

$$P : \min \{c^T x \mid Ax + b \leq 0 \ , \ x \geq 0\} \quad ,$$

$$D : \max \{b^T y \mid A^T y + c \geq 0 \ , \ y \geq 0\} \ .$$

Wenn (8) ein konvexes Programm darstellt (d.h. wenn $X \subseteq \mathbb{R}^k$ eine konvexe Menge ist und wenn $F(x)$, $f_j(x)$ auf X konvexe Funktionen sind), wenn $\hat{x}$ eine Lösung von (8) ist, und wenn die Slater'sche Regularitätsvoraussetzung erfüllt ist, nämlich

(9) $$\exists\, x^o \in X : f(x^o) < 0 \ ,$$

so gelten die *Kuhn-Tucker-Bedingungen:*

(10) $$\exists\, \hat{y} \in \mathbb{R}^m_+ : F(\hat{x}) \leq F(x) + \hat{y}^T f(x) \ \forall\, x \in X \ .$$

Aus (10) folgt sofort, dass $M(\hat{x}) \leq m(\hat{y})$; $(\hat{x},\hat{y})$ ist somit ein Sattelpunkt von φ über $X \times \mathbb{R}^m_+$. Wenn umgekehrt $(\hat{x},\hat{y})$ ein Sattelpunkt ist, so ist $(\hat{x})$ Lösung von (8), und die "Multiplikatoren" $\hat{y}_j$ erfüllen zusammen mit $\hat{x}$ die Bedingung (10). Aus (9) folgt weiterhin, dass

(11) $$\sum_j \hat{y}_j \leq \frac{F(x^o) - F(\hat{x})}{\min_j |f_j(x^o)|} \equiv B$$

für alle $\hat{y}$, die (10) erfüllen. Es gehen also keine Sattelpunkte von φ verloren, wenn wir von vornherein die nichtkompakte Menge $Y = \mathbb{R}^m_+$ ersetzen durch die kompakte Menge $Y = \{y \in \mathbb{R}^m \mid 0 \leq y_j \leq B\}$. Die Zahl der Sattelpunkte nimmt auch nicht zu, wenn wir uns auf solche Sattelpunkte beschränken, deren x-Teil eine Lösung des ursprünglichen Problems (8) darstellt. Wir wollen mit dieser Bemerkung plausibel machen, dass die später zu stellende Forderung der Kompaktheit von Y den Fall (7) nicht a priori ausschliesst.

Um die Anwendbarkeit des Satzes von Kakutani sicherzustellen, werden wir im weiteren, falls nicht anders erwähnt, bei der Behandlung der Probleme (3), (4) stets die folgende *Voraussetzung* treffen:

(12) $\left\{\begin{array}{l} X \subseteq \mathbb{R}^k \text{ und } Y \subseteq \mathbb{R}^m \text{ seien nichtleer, kompakt und konvex.} \\ \varphi\ (x,y) \text{ sei stetig und konvex-konkav auf } X \times Y. \end{array}\right.$

Die Funktion $M(x)$ ist dann konvex (als sup einer Familie von konvexen Funktionen), ebenso ist $m(y)$ konkav. Ausserdem sind $M(x)$ und $m(y)$ stetig. Die Extrema in (1), (2) sowie in (3), (4) werden angenommen, und wir können durchweg inf sup φ ersetzen durch min max φ. Wir werden von $(M(x)$ übrigens nur die Halbstetigkeit nach unten benötigen,

$$\liminf_{x^n \to \hat{x}} M(x^n) \geq M(\hat{x}) \quad ,$$

und diese Eigenschaft ist auch dann gewährleistet, wenn Y nicht kompakt ist, wie im Falle (7).

Bei der Behandlung des Problems (8) werden wir stets die folgende *Voraussetzung* treffen:

(13) $\left\{\begin{array}{l} X \subseteq \mathbb{R}^k \text{ sei nichtleer, kompakt und konvex. } F(x) \text{ und } f_j(x) \\ (j = 1, \ldots, m) \text{ seien auf } X \text{ stetig und konvex.} \\ \qquad\qquad\qquad\qquad \text{Ferner sei (9) erfüllt.} \end{array}\right.$

Dann hat (8) mindestens eine Lösung, die Kuhn-Tucker-Bedingungen sind erfüllt, und φ hat eine kompakte Menge von Sattelpunkten. (13) impliziert offenbar (12), bis auf die Kompaktheit von Y.

Gelegentlich werden wir uns für die Differenzierbarkeit von $M(x)$ interessieren. Zusätzlich zur Generalvoraussetzung (12) sei $\varphi(x,y)$ streng konkav in y und stetig nach x differenzierbar, mit dem x-Gradienten $\nabla_x \varphi(x,y)$. Dann ist $M(x)$ stetig differenzierbar, und es ist

$$(14) \qquad \nabla M(x) = \nabla_x \varphi\,[x,y(x)] \;,$$

wo $y(x)$ so bestimmt wird, dass $\varphi[x,y(x)] = M(x)$.

Beweis: Aus der Stetigkeit und strengen Konkavität von φ folgt leicht, dass $y(x)$ eindeutig und stetig von x abhängt. Wir zeigen dies hier nicht. Damit ist der im (14) als Gradient von M angegebene Vektor stetig in x. Wir zeigen nun die totale Differenzierbarkeit. Sei $t(x) = \nabla_x \varphi\,[x,y(x)]$. Man hat wegen der für konvexe, differenzierbare Funktionen gültigen Stützungleichung, dass

$$\begin{aligned} M(\xi) - M(x) &= \sup_Y \varphi(\xi,y) - \varphi[x,y(x)] \\ &\geq \varphi[\xi,y(x)] - \varphi[x,y(x)] \\ &\geq (\xi - x)^T\, \nabla_x\varphi\,[x,y(x)] \quad [\text{Stützungleichung}] \\ &= (\xi - x)^T\, t(x) \;. \end{aligned}$$

Durch Vertauschung von x und folgt hieraus mittels der Schwarzschen Ungleichung und der Stetigkeit von $t(x)$

$$\begin{aligned} M(\xi) - M(x) &\leq (\xi - x)^T\, t(\xi) \\ &\leq (\xi - x)^T\, t(x) + |\xi - x| \cdot |t(\xi) - t(x)| \\ &\leq (\xi - x)^T\, t(x) + o(\xi - x) \;, \end{aligned}$$

$$\lim_{\xi \to x} \frac{o(\xi - x)}{|\xi - x|} = 0 \quad .$$

Beide Ungleichungen zusammen ergeben

$$|M(\xi) - M(x) - (\xi - x)^T t(x)| \leq o\ (\xi - x) \ .$$

Damit ist die Differenzierbarkeit erwiesen, $\nabla M(x) = t(x)$.

Dekomposition: Dantzig und Huard haben nun je ein iteratives Dekompositionsverfahren zur Lösung des konvexen Programms

$$\min \{F(x) | x \in X \ , \ f(x) \leq 0\}$$

unter der Voraussetzung (13)angegeben. Beide Verfahren lösen das Problem über eine alternierende Folge von vorgenannten Master-Programmen und Subprogrammen.

Das Verfahren von Dantzig: Man startet mit einem Punkt $x^o \in X$, der die Slater Bedingung (9) erfüllt. Zur Durchführung der n-ten Iteration benötigt man die bereits bekannten Punkte $x^o, \ldots, x^{n-1} \in X$.

Master-Programm: Man löse das Paar dualer linear Programme

$$(15) \quad \left\{ \begin{array}{l} \max \quad z \\ \quad \text{unter} \quad -z + \eta^T f(x^j) + F(x^j) \geq 0 \ (j = 0, \ldots, n-1) \ , \\ \qquad\qquad z \in \mathbb{R} \ , \ \eta \geq 0 \\ \qquad\qquad (\text{"dual master"}) \end{array} \right.$$

$$(16) \quad \left\{ \begin{array}{l} \min \quad \sum_{j=0}^{n-1} \lambda_j \ F(x^j) \\ \quad \text{unter} \quad \sum_{j=0}^{n-1} \lambda_j \ f(x^j) \leq 0 \\ \qquad\qquad \sum_{j=0}^{n-1} \lambda_j = 1 \ , \ \lambda_j \geq 0 \ (j = 0, \ldots, n-1) \ , \\ \qquad\qquad (\text{"primal master"}) \end{array} \right.$$

mit Lösungen η^n, λ^n. Setze

$$\xi^n = \sum_{j=0}^{n-1} \lambda_j^n \; x^j \quad .$$

Subprogramm: Man bestimme $x^n \in X$ als Lösung von

(17) $$\min_X \; \varphi\,(x, \eta^n) \quad .$$

Hierbei ist $\varphi(x,y) = F(x) + y^T f(x)$ die zu (8) gehörende Lagrange-funktion.Damit ist eine Iteration beendet. Unter den getroffenen Voraussetzungen ist das Verfahren theoretisch durchführbar, und es gilt: Die Folge (ξ^n, η^n) hat mindestens ein Häufungspunkt. Wenn $(\hat{\xi}, \hat{\eta})$ ein solcher Häufungspunkt ist, so ist $(\hat{\xi}, \hat{\eta})$ Sattelpunkt von φ auf $X \times \mathbb{R}_+^m$.

Das Verfahren von Huard setzt zusätzlich noch voraus, dass $F(x)$ und $f_j(x)$ stetig nach x differenzierbar sind. Für die n-te Iteration benötigt man wieder die Hilfspunkte $x^o, \ldots, x^{n-1} \in X$ (x^o so dass (9) erfüllt ist). Man wählt eine kompakte konvexe Menge $\Xi^n \subseteq X$ derart, dass $\{x^o, \ldots, x^{n-1}\} \subseteq \Xi^n$.

Master-Programm: Man bestimmt ξ^n als Lösung von

(18) $$\min \{F(\xi) \,|\, \xi \in \Xi^n, \; f(\xi) \le 0\}, \qquad (\text{"primal master"})$$

und η^n als zugehörige Kuhn-Tucker-Multiplikator:

(19) $$\eta^n \ge 0, \; F(\xi^n) \le F(\xi) + (\eta^n)^T f(\xi) \;\forall\; \xi \in \Xi^n \qquad (\text{"dual master"}).$$

Man wählt also (ξ^n, η^n) als Sattelpunkt von $\varphi(\xi,\eta)$ über $\Xi^n \times \mathbb{R}_+^m$.

Subprogramm: Man bestimmt x^{n+1} als Lösung von

$$(20) \qquad \min_X \; [\, \varphi(\xi^n, \eta^n) + (x - \xi^n)^T \, \nabla_x \, \varphi(\xi^n, \eta^n)] \quad .$$

Damit ist eine Runde beendet. Es gilt die gleiche Aussage wie beim Verfahren von Dantzig.

2. Das symmetrische Dekompositionsverfahren

Wir suchen unter Zugrundelegung der Generalvoraussetzung (12) Lösungen der beiden dualen Probleme

$$(21) \qquad P : \min_X \; \max_Y \; \varphi(x,y) \; [\, = \min_X M(x) \,] \quad ,$$

$$(22) \qquad D : \max_Y \; \min_X \; \varphi(x,y) \; [\, = \max_Y m(y) \,] \quad ,$$

d.h. also, wir suchen einen Sattelpunkt von $\varphi(x,y)$ über $X \times Y$. Bevor wir uns dem eigentlichen Dekompositionsverfahren zuwenden, erinnern wir an das Schnittverfahren, das wir im folgenden mehrfach heranziehen werden. Für das Schnittverfahren sind keine Konvexitätsvoraussetzungen notwendig.

Das Schnittverfahren für D: Für den Start wählt man eine Punkt $x^o \in X$ beliebig. Zur Durchführung der n-ten Iterationsrunde, die die Punkte $\eta^n \in Y$ und $x^n \in X$ liefert, benötigt man die bereits bekannten Punkte $x^o, \ldots, x^{n-1} \in X$. Man wählt eine Menge X^n derart, dass

$$\{x^o, \ldots, x^{n-1}\} \subseteq X^n \subseteq X \quad , \quad X^n \text{ kompakt.}$$

Man bestimmt zuerst η^n und dann x^n als Lösung der folgenden Teilprobleme:

(23) $\qquad \eta^n : \max\limits_{Y} \min\limits_{X^n} \varphi(x,y)$ $\qquad$ ("dual master") ,

(24) $\qquad x^n : \min\limits_{X} \varphi(x, \eta^n)$ $\qquad$ ("dual subproblem").

Damit ist eine Runde des Verfahrens beendet. (24) und (25) haben aus Stetigkeits- und Kompaktheitsgründen Lösungen.

Satz 1. - Sei $\hat{\eta}$ ein Häufungspunkt von η^n. Dann ist $\hat{\eta}$ Lösung von D (hierfür sind keine Konvexitätsvoraussetzungen nötig).

Beweis: Sei $\hat{m} = \max\limits_{Y} m(y)$. Wegen der Kompaktheit von $X \times Y$ existiert eine konvergente Teilfolge $(x^{\bar{n}}, \eta^{\bar{n}}) \to (\bar{x}, \hat{\eta}) \in X \times Y$. Wegen (24) und wegen der Definition von $m(y)$ gilt

$$\varphi(x^n, \eta^n) = m(\eta^m) \quad .$$

Hieraus folgt im limes, da φ stetig und m halbstetig nach oben ist,

$$\varphi(\bar{x}, \hat{\eta}) \leq m(\hat{\eta}) \quad .$$

Wegen $x^k \in X^n$ für alle $k < n$, wegen (23), wegen $X^n \subseteq X$, und wegen der Definition von $\hat{m}$ gilt

$$\varphi(x^k, \eta^n) \geq \min_{X^n} \varphi(x, \eta^n) = \max_{Y} \min_{X^n} \varphi(x,y) \geq \max_{Y} \min_{X} \varphi(x,y) = \hat{m}$$

für alle $k < n$. Setzt man insbesondere $k = \bar{n}$, $n = \overline{n+1}$, so erhält man

$$\varphi\left(x^{\bar{n}}, \eta^{\overline{n+1}}\right) \geq \hat{m} \; ,$$

und, wegen $x^{\bar{n}} \to \bar{x}$, $\eta^{\overline{n+1}} \to \hat{\eta}$, im limes

$$\varphi(\bar{x}, \hat{\eta}) \geq \hat{m} \; ,$$

$$m(\hat{\eta}) \geq \hat{m} \; .$$

Damit ist $\hat{\eta}$ Lösung von D.

Analog definiert man das

Schnittverfahren für P: $y^o \in Y$ beliebig; Y^n kompakt derart, dass

$$\{y^o, \ldots, y^{n-1}\} \subseteq Y^n \subseteq Y .$$

(25) $$\xi^n : \min_X \max_{Y^n} \varphi(x,y) \qquad \text{("primal master")},$$

(26) $$y^n : \max_Y \varphi(\xi^n, y) \qquad \text{("primal subproblem")} .$$

Jeder Häufungspunkt $\hat{\xi}$ von ξ^n löst P.

Das symmetrische Dekompositionsverfahren: Für den Start wählt man $x^o \in X$, $y^o \in Y$ beliebig. Zur Durchführung der n-ten Iterationsrunde benötigt man die bereits bekannten Hilfspunkte $x^o, \ldots, x^{n-1} \in X$ sowie $y^o, \ldots, y^{n-1} \in Y$. Man wählt X^n kompakt und Ξ^n kompakt, konvex derart, dass

$$\{x^o, \ldots, x^{n-1}\} \subseteq X^n \subseteq \Xi^n \subseteq X ,$$

sowie Y^n kompakt und H^n kompakt, konvex derart, dass

$$\{y^o, \ldots, y^{n-1}\} \subseteq Y^n \subseteq H^n \subseteq Y .$$

Man bestimmt zuerst $\xi^n \in \Xi^n$, $\eta^n \in H^n$ und dann $x^n \in X$, $y^n \in Y$ als Lösungen der folgenden Teilprobleme:

(27) $$\eta^n : \max_{H^n} \min_{X^n} \varphi(x,y) \qquad \text{("dual master")},$$

(28) $$\xi^n : \min_{\Xi^n} \max_{Y^n} \varphi(x,y) \qquad \text{("primal master")},$$

(29) $$x^n : \min_X \varphi(x, \eta^n) \qquad \text{("dual subproblem")},$$

(30) $\quad y^n : \max_Y \varphi(\xi^n, y)$ ("primal subproblem").

Damit ist eine Runde des Verfahrens beendet. Wegen der Stetigkeit von φ und der Kompaktheit der auftretenden Mengen haben die Probleme (27) - (30) Lösungen.

Satz 2. - Jeder Häufungspunkt $(\hat{\xi}, \hat{\eta})$ der durch (27) - (30) erzeugten Folge (ξ^n, η^n) ist ein Sattelpunkt von φ auf $X \times Y$.

Beweis: Wegen der Kompaktheit von $X \times Y \times X \times Y$ existiert eine konvergente Teilfolge $(x^{\bar{n}}, y^{\bar{n}}, \xi^{\bar{n}}, \eta^{\bar{n}}) \to (\bar{x}, \bar{y}, \hat{\xi}, \hat{\eta}) \in X \times Y \times X \times Y$. Es gilt für alle $k < n$

$$\begin{aligned} \varphi(x^k, \eta^n) &\geq \min_{X^n} \varphi(x, \eta^n) && [\text{wegen } x^k \in X^n \;\forall\, k < n] \\ &= \max_{H^n} \min_{X^n} \varphi(x,y) && [\text{wegen } (27)] \\ &\geq \max_{H^n} \min_{\Xi^n} \varphi(x,y) && [\text{wegen } X^n \subseteq \Xi^n]\,. \end{aligned}$$

Analog für alle $k < n$

$$\varphi(\xi^n, y^k) \leq \min_{\Xi^n} \max_{H^n} \varphi(x,y)\,.$$

Nach dem Satz von Kakutani ist

$$\min_{\Xi^n} \max_{H^n} \varphi(x,y) = \max_{H^n} \min_{\Xi^n} \varphi(x,y).$$

Somit gilt

$$\varphi(\xi^n, y^k) \leq \varphi(x^k, \eta^n) \;\forall\, k < n\,.$$

Insbesondere wegen $\bar{n} < \overline{n+1}$

$$\varphi(\xi^{\overline{n+1}}, y^{\bar{n}}) \leq \varphi(x^{\bar{n}}, \eta^{\overline{n+1}})\,,$$

und im limes

$$\varphi(\hat{\xi}, \bar{y}) \leq \varphi(\bar{x}, \hat{\eta}) \quad .$$

Nach (29), (30) gilt

$$m(\eta^n) = \varphi(x^n, \eta^n), \; M(\xi^n) = \varphi(\xi^n, y^n) \quad ,$$

somit im limes für die Teilfolge, unter Verwendung der Halbstetigkeit von m und M,

$$m(\hat{\eta}) \geq \varphi(\bar{x}, \hat{\eta}), \; M(\hat{\xi}) \leq \varphi(\hat{\xi}, \bar{y}) \; .$$

Insgesamt ergibt sich damit

$$M(\hat{\xi}) \leq \varphi(\hat{\xi}, \bar{y}) \leq \varphi(\bar{x}, \hat{\eta}) \leq m(\hat{\eta}) \; ;$$

$(\hat{\xi}, \hat{\eta})$ ist also ein Sattelpunkt.

Man sieht, dass Satz 2 gültig bleibt, wenn man die Extremalprobleme (27) - (30) nur bis auf ϵ_n genau löst, $\epsilon_n \to 0$.

Eine naheliegende Wahl von X^n und Ξ^n, die die geforderten Bedingungen erfüllt, ist wohl $X^n = \{x^o, \ldots, x^{n-1}\}$, $\Xi^n = [x^o, \ldots, x^{n-1}]$ (die konvexe Hülle). Wir diskutieren kurz die geometrischen Aspekte verschiedener anderer Wahlmöglichkeiten für X^n, Ξ^n, Y^n, H^n.

1) Setzt man $H^n = Y$, $\Xi^n = X$, so erhält man zwei parallele, unverknüpfte Verfahren, nämlich das duale Schnittverfahren (23), (24) und das primale Schnittverfahren (25), (26). Aus Satz 1 folgt dann (ohne Konvexitätsvoraussetzungen), dass $\hat{\eta}$ eine Lösung von D und $\hat{\xi}$ eine Lösung von P darstellt (ohne notwendig ein Sattelpunkt zu sein).

Dieser Sonderfall des Verfahrens lässt sich am bequemsten unter dem Gesichtspunkt der "Approximation von aussen" interpretieren: Die über X zu minimierende Funktion $M(x)$ wird durch eine Folge von Funktionen $M^n(x) = \max_{Y^n} \varphi(x,y)$ von unten (aussen) her approximiert. Analog für $m(y)$.

2) Setzt man $X^n = \Xi^n$, $Y^n = H^n$, so sind (27), (28) duale Probleme, die besagen, dass (ξ^n, η^n) ein Sattelpunkt von φ über $\Xi^n, \times H^n$ sein soll. Die Menge $X \times Y$ wird durch $\Xi^n \times H^n$ von innen her approximiert.

3) Setzt man $\Xi^n = X^n = X$ (die Bestimmung von x^n erübrigt sich in diesem Fall), so erhält man ein unsymmetrisches Dekompositionsverfahren - wir nennen es das *duale Dekompositionsverfahren* (obwohl $(\hat{\xi}, \hat{\eta})$ natürlich wegen Satz 2 immer noch beide Probleme P und D lösen):

$$\xi^n : \min_{X} \max_{Y^n} \varphi(x,y) \qquad (\text{"primal master"}),$$

$$\eta^n : \max_{H^n} \min_{X} \varphi(x,y) \qquad (\text{"dual master"}),$$

$$y^n : \max_{Y} \varphi(\xi^n, y) \qquad (\text{"primal subproblem"}) .$$

Das unsymmetrische Dekompositionsverfahren kombiniert die beiden Aspekte der Approximation von aussen und der Approximation von innen. Statt $M(x)$ minimiert man $M^n(x)$ über X, und statt über Y maximiert man $m(y)$ über H^n. Das Subproblem dient der Bestimmung von $M^{n+1}(x)$ (Schnittgeneration) und der Bestimmung von H^{n+1} (Punktgeneration).

4) Setzt man $H^n = Y^n = Y$, so erhält man ein unsymmetrisches, primales Dekompositionsverfahren, das wir nochmals besonders herausstellen, da wir es noch mehrfach abändern wollen.

3. Das primale Dekompositionsverfahren

Man wählt $x^o \in X$ beliebig. Iterationsvorschrift:

(31) $\xi^n : \min_{\Xi^n} \max_{Y} \varphi(x,y)$ ("primal master") ,

(32) $\eta^n : \max_{Y} \min_{X^n} \varphi(x,y)$ ("dual master") ,

(33) $x^n : \min_{X} \varphi(x, \eta^n)$ ("dual subproblem^n").

Hierbei wird X^n kompakt und Ξ^n kompakt, konvex derart gewählt, dass

$$\{x^o, \ldots, x^{n-1}\} \subseteq X^n \subseteq \Xi^n \subseteq X .$$

Für die so erzeugte Folge (ξ^n, η^n) gilt

Satz 3. – Jeder Häufungspunkt $(\hat{\xi}, \hat{\eta})$ der Folge (ξ^n, η^n) ist ein Sattelpunkt von $\varphi(x,y)$ über $X \times Y$.

Der Beweis ergibt sich, wie erwähnt, aus Satz 2 durch Spezialisierung. Wir betrachten noch kurz das Verfahren (31) - (33) für den *Fall, dass* φ *eine Lagrangefunktion ist,*

$$\varphi(x,y) = F(x) + y^T f(x), \; Y = \{y \mid y \geq 0\} \equiv \mathbb{R}^m_+ ,$$

wobei wir $X^n = \Xi^n$ (kompakt, konvex) setzen. Dann ist (31), (32) sinngemäss so zu interpretieren, dass (ξ^n, η^n) einen Sattelpunkt von φ auf $\Xi^n \times \mathbb{R}^m_+$ darstellt, d.h. ξ^n löst das konvexe Programm

(34) $\min \{F(\xi) \mid \xi \in \Xi^n, \; f(\xi) \leq 0\}$,

und η^n erfüllt mit ξ^n die Kuhn-Tucker-Bedingung für dieses Programm,

(35) $\eta^n \geq 0, \quad F(\xi^n) \leq \varphi(\xi, \eta^n) \; \forall \, \xi \in \Xi^n$.

Weil Y nicht kompakt ist, müsser wir jetzt zusätzlich für den Startpunkt $x^o \in X$ fordern, dass $f(x^o) < 0$. Diese Forderung stellt wegen $x^o \in \Xi^n$ $(\forall n)$ sicher, 1) dass (34) eine Lösung ξ^n hat (weil der zulässige Bereich nicht leer ist); 2) dass (35) eine Lösung η^n hat (weil die Slater-Bedingung erfüllt ist); 3) dass die η^n gleichmässig beschränkt sind (man entnimmt dies leicht aus (11), wenn man berücksichtigt, dass $\Xi^n \subseteq X$). Das durch (34), (35) und (33) definierte Verfahren ist also theoretisch durchführbar, und die beschränkte Folge (ξ^n, η^n) hat mindestens einen Häufungspunkt $(\hat{\xi}, \hat{\eta})$. $\hat{\xi}$ löst dann (8), und $\hat{\eta}$ erfüllt zusammen mit $\hat{\xi}$ die Bedingungen (10).

Wir diskutieren im folgenden noch einige Varianten von (31) - (33), aus denen sich etwa die Verfahren von Dantzig und Huard ergeben. Es wäre ein leichtes, das allgemeine Modell so zu formulieren, dass diese Varianten sich durch blasse Spezialisierung ergeben; wir haben darauf verzichtet, um die Darstellung nicht unnötig zu komplizieren.

4. Varianten des primalen Dekompositionsverfahrens

Eine Variante, wenn φ in y streng konkav ist. Wenn φ in y sogar streng konkav ist, so lässt sich das Verfahren (31) - (33) bedeutend vereinfachen, und zwar wie folgt: Für die Durchführung des allgemeinen Schrittes benötigt man nur ξ^{n-1} und x^{n-1} (für den Start wählt man $\xi^o \in X$ und $x^o \in X$ beliebig). Sei $[\xi^{n-1}, x^{n-1}]$ die abgeschlossene Verbindungsstrecke zwischen ξ^{n-1} und x^{n-1}. Man berechnet der Reihe nach $\xi^n \in [\xi^{n-1}, x^{n-1}]$, $\eta^n \in Y$, $x^n \in X$ gemäss:

(36) $\xi^n : \min_{[\xi^{n-1}, x^{n-1}]} \max_Y \varphi(x,y)$,

(37) $\eta^n : \max_Y \varphi(\xi^n, \eta)$,

(38) $x^n : \min_X \varphi(x, \eta^n)$.

Für die so erzeugte Folge (ξ^n, η^n) gilt wieder die Aussage von Satz 3, falls φ zusätzlich streng konkav in y ist.

Beweis: Konvergiere $(x^{\bar{n}}, \xi^{\bar{n}}, \eta^{\bar{n}}) \to (\bar{x}, \hat{\xi}, \hat{\eta})$. Aus (36) folgt

(39) $M(\xi^{n+1}) \leq M(\xi) \forall \xi \in [\xi^n, x^n]$.

Somit insbesondere $M(\xi^{n+1}) \leq M(\xi^n)$, und wegen dieser Monotonie

$$\lim_{\bar{n}\to\infty} M(\xi^{\bar{n}+1}) = \lim_{\bar{n}\to\infty} M(\xi^{\bar{n}}) = M(\hat{\xi}) .$$

Dies, zusammen mit (39), ergibt für $\bar{n} \to \infty$

$$M(\hat{\xi}) \leq M(\xi) \forall \xi \in [\hat{\xi}, \bar{x}] .$$

$\hat{\xi}$ löst also $\min_{[\hat{\xi},\bar{x}]} \max_Y \varphi(x,y)$. Nach dem Satz von Kakutani existiert ein $\tilde{\eta} \in Y$ derart, dass $(\hat{\xi}, \tilde{\eta})$ einen Sattelpunkt von φ auf $[\hat{\xi}, \bar{x}] \times Y$ darstellt, insbesondere also gilt

$$\varphi(\hat{\xi}, \tilde{\eta}) \geq \varphi(\hat{\xi}, \eta) \forall \eta \in Y.$$

Aus (37) folgt

$$M(\xi^n) = \varphi(\xi^n, \eta^n) \geq \varphi(\xi^n, \eta) \forall \eta \in Y$$

und im limes

$$M(\hat{\xi}) = \varphi(\hat{\xi}, \hat{\eta}) \geq \varphi(\hat{\xi}, \eta) \;\forall\, \eta \in Y.$$

Wegen der strengen Konkavität ist der Maximumpunkt von $\varphi(\hat{\xi}, \eta)$ eindeutig bestimmt, und es folgt

$$\tilde{\eta} = \hat{\eta} \; .$$

Somit ist $(\hat{\xi}, \hat{\eta})$ Sattelpunkt auf $[\hat{\xi}, \bar{x}] \times Y$, insbesondere gilt also

$$\varphi(\hat{\xi}, \hat{\eta}) \leq \varphi(\xi, \hat{\eta}) \;\forall\, \xi \in [\hat{\xi}, \bar{x}] \; .$$

Hieraus folgt

$$\varphi(\hat{\xi}, \hat{\eta}) \leq \varphi(\bar{x}, \hat{\eta}) \; .$$

Aus (38) folgt im limes

$$\varphi(\bar{x}, \hat{\eta}) = m(\hat{\eta}) \; .$$

Somit gilt

$$M(\hat{\xi}) = \varphi(\hat{\xi}, \hat{\eta}) \leq \varphi(\bar{x}, \hat{\eta}) = m(\hat{\eta}) \; ,$$

und $(\hat{\xi}, \hat{\eta})$ ist ein Sattelpunkt.

Eine Variante, wenn φ stetig nach x differenzierbar ist. Wenn φ stetig (in beiden Argumenten) nach x differenzierbar ist, so kann man (31) - (33) wie folgt modifizieren (Ξ^n wie zuvor):

(40) $$\xi^n : \min_{\Xi^n} \max_{Y} \varphi(x,y) \; ,$$

(41) $$\eta^n : \max_{Y} \min_{\Xi^n} \varphi(x,y) \; ,$$

(42) $$x^n : \min_{X} \psi^n(x), \text{ wobei } \psi^n(x) = \varphi(\xi^n, \eta^n) + (x - \xi^n)^T \nabla_x \varphi(\xi^n, \eta^n) \; .$$

Es gilt wieder die Aussage von Satz 3.

Beweis: Konvergiere $(x^{\bar{n}}, \xi^{\bar{n}}, \eta^{\bar{n}}) \to (\bar{x}, \hat{\xi}, \hat{\eta})$. Aus (40), (41) folgt nach dem Satz von Kakutani, dass (ξ^n, η^n) einen Sattelpunkt über $\Xi^n \times Y$ darstellt, d.h. es gilt

(43) $$\varphi(\xi^n, \eta^n) \leq \varphi(\xi, \eta^n) \;\forall\; \xi \in \Xi^n \quad ,$$

(44) $$\varphi(\xi^n, \eta^n) \geq \varphi(\xi^n, \eta) \;\forall\; \eta \in Y.$$

Aus (43) folgt wegen der Konvexität von Ξ^n und der Differenzierbarkeit von φ, dass

$$(\xi - \xi^n)^T \nabla_x \varphi(\xi^n, \eta^n) \geq 0 \;\forall\; \xi \in \Xi^n \quad ,$$

somit

$$\psi^n(\xi) \geq \varphi(\xi^n, \eta^n) \;\forall\; \xi \in \Xi^n \quad ,$$

und insbesondere

$$\psi^n(x^k) \geq \varphi(\xi^n, \eta^n) \;\forall\; k < n \quad .$$

Mit $k = \bar{n}$, $n = \overline{n + 1}$ erhält man im limes

$$\hat{\psi}(\bar{x}) \equiv \varphi(\hat{\xi}, \hat{\eta}) + (\bar{x} - \hat{\xi})^T \nabla_x \varphi(\hat{\xi}, \hat{\eta}) \geq \varphi(\hat{\xi}, \hat{\eta}) \quad .$$

Aus $\psi^n(x) \leq \varphi(x, \eta^n) \;\forall\; x \in X$ (Stützungleichung für eine konvexe, differenzierbare Funktion) und aus (42) folgt

$$\psi^n(x^n) = \min_X \psi^n(x) \leq \min_X \varphi(x, \eta^n) = m(\eta^n) \quad ,$$

im lines

$$\hat{\psi}(\bar{x}) \leq m(\hat{\eta}) \quad .$$

Damit hat man

$$\varphi(\hat{\xi}, \hat{\eta}) \leq \hat{\psi}(\bar{x}) \leq m(\hat{\eta}) \quad .$$

Aus (44) folgt im limes

$$\varphi(\hat{\xi}, \hat{\eta}) = M(\hat{\xi}) \quad .$$

Damit hat man wieder die Sattelpunkteigenschaft $M(\hat{\xi}) \leq m(\hat{\eta})$.

Wenn φ eine Lagrangefunktion ist, so geht (40) - (42) über in das Verfahren (18) - (20) von Huard.

Das Verfahren der zulässigen Richtungen für P. Falls φ zugleich stetig nach x differenzierbar und streng konkav in y ist, so kann man die Verfahren (36) - (38) und (40) - (42) kombinieren. Man erhält

(45) $$\xi^n : \min_{[\xi^{n-1}, x^{n-1}]} \max_{Y} \varphi(x,y) \quad ,$$

(46) $$\eta^n : \max_{Y} \varphi(\xi^n, y) \quad ,$$

(47) $$x^n : \min_{X} \varphi(\xi^n, \eta^n) + (x - \xi^n)^T \nabla_x \varphi(\xi^n, \eta^n) \quad .$$

Dieses Verfahren erlaubt nun aber eine sehr einfache und wohlbekannte Interpretation. Wir wissen nämlich, dass in diesem Falle $M(x)$ stetig differenzierbar ist, und dass

$$\nabla M(x) = \nabla_x \varphi[x, y(x)] \, ,$$

wobei $y(x)$ so zu wählen ist, dass $\varphi[x,y(x)] = M(x)$. Aus (46) folgt aber, dass $\varphi(\xi^n, \eta^n) = M(\xi^n)$, und mithin

$$\nabla M(\xi^n) = \nabla_x \; \varphi(\xi^n, \eta^n) \quad .$$

Damit kann man (45) und (47) wie folgt formulieren:

$$(48) \qquad \xi^n : \min_{[\xi^{n-1}, x^{n-1}]} M(\xi) \quad ,$$

$$(49) \qquad x^n : \min_X \; M(\xi^n) + (x - \xi^n) \, \nabla M(\xi^n) \quad .$$

(48), (49) ist aber das bekannte *Verfahren der zulässigen Richtungen* von Frank und Wolfe [7] zur Minimierung einer stetig differenzierbaren konvexen Funktion $M(x)$ auf einer kompakten konvexen Menge X. Dieses Verfahren erweist sich damit als eine extreme Variante des Dekompositionsverfahrens, so wie andererseits das duale Schnittverfahren als extreme Variante im primalen Dekompositionsverfahren enthalten ist.

Eine Variante von Dantzig. Um das Verfahren von Dantzig zu erhalten, ändern wir den Schritt (31) im Verfahren (31) - (33) noch etwas ab. Sei $X^n = \{x^o, \ldots, x^{n-1}\}$. Mit dieser Wahl in X^n lässt sich (32) schreiben als

$$(50) \qquad \begin{cases} \max \; z \\ \quad z \leq \varphi(x^j, y) \qquad (j = 0, \ldots, n-1) \; , \\ \quad y \in Y, \; z \in \mathbb{R} \quad . \end{cases}$$

Sei η^n, z^n eine Lösung dieses konvexen Programms. Wegen der speziellen Form von (50) ist die Slater-Voraussetzung für (50) automatisch erfüllt, und es gelten die Kuhn-Tucker Bedingungen: Es existieren $\lambda_0 \geq 0, \ldots, \lambda_{n-1} \geq 0$ derart, dass

$$z^n \geq z + \sum_{j=0}^{n-1} \lambda_j \, [\varphi(x^j, y) - z] \quad \forall\, y \in Y \;, \forall\, z \in \mathbb{R} \;.$$

Hieraus folgt unschwer

$$(51) \qquad \begin{cases} \sum\limits_{j=0}^{n-1} \lambda_j = 1, \quad \lambda_j \geq 0 \quad (j = 0, \ldots, n-1) \;, \\ z^n \geq \sum\limits_{j=0}^{n-1} \lambda_j \; \varphi(x^j, y) \quad \forall\, y \in Y \;. \end{cases}$$

Bestimmt man jetzt ξ^n statt durch (31) durch

$$(52) \qquad \xi^n = \sum_{j=0}^{n-1} \lambda_j \; x^j \;,$$

so ist $\xi^n \in X$, und es gilt wegen der Konkavität von φ bez. y und wegen (51)

$$z^n \geq \varphi(\xi^n, y) \quad \forall\, y \in Y \;,$$

somit auch

$$z^n \geq \max_Y \; \varphi(\xi^n, y) = M(\xi^n) \;.$$

Wegen der Nebenbedingungen in (50) ergibt dies

$$\varphi(x^j, \eta^n) \geq M(\xi^n) \quad \forall\, j < n \;.$$

Damit kann man wiederum die Aussage von Satz 3 beweisen, wenn $(x^{\bar{n}}, \xi^{\bar{n}}, \eta^{\bar{n}}) \to (\bar{x}, \hat{\xi}, \hat{\eta})$. Setzt man nämlich $j = \bar{n}$, $n = \overline{n+1}$, so erhält man im limes

$$\varphi(\bar{x}, \hat{\eta}) \geq M(\hat{\xi}) \;.$$

Aus (33) folgt aber im limes

$$m(\hat{\eta}) = \varphi(\bar{x}, \hat{\eta}) .$$

Somit

$$M(\hat{\xi}) \leq m(\hat{\eta}) .$$

Das durch (50), (51), (52) und (33) gegebene Verfahren kann als Verallgemeinerung des Verfahrens von Dantzig angesehen werden. Ist nämlich

$$\varphi(x,y) = F(x) + y^T f(x) , \; Y = \{y | y \geq 0\} ,$$

so geht (50) über in das lineare Programm (15), und die λ_j, die die Kuhn-Tucker-Bedingungen für (15) erfüllen, sind identisch mit den Lösungen des dualen linearen Programms (16). Wegen der fehlenden Kompaktheit von Y benötigt man wieder die Voraussetzung $f(x^o) < 0$ die sicherstellt, dass (15) und (16) Lösungen haben, und dass die Lösungen η^n von (15) gleichmässig beschränkt sind.

Das Verfahren von Benders. Um noch kurz auf das Verfahren von Benders einzugehen, betrachten wir das Problem

(53) $$\min_{X \times V} \max_{Y} \psi(x, v, y) .$$

Hierbei ist X kompakt, V und Y sind kompakt, konvex. ψ ist stetig in (x, v, y) und konvex-konkav in (v,y) für beliebiges x. Nach dem Satz von Kakutani ist dann (53) äquivalent zu

(54) $$\min_{X} \left[\max_{Y} \min_{V} \psi(x, v, y)\right]$$

im folgenden Sinne: $(\hat{\xi}, \bar{v})$ löst (53) genau dann, wenn $\hat{\xi}$ (54) löst und $\bar{v}$ zusammen mit einem $\bar{y}$ ein Sattelpunkt von $\psi(\hat{\xi}, v, y)$ über

$V \times Y$ ist. Problem (54) hat die Form

(55) $$P : \min_X \max_Y \varphi(x,y), \quad \varphi(x,y) = \min_V \psi(x, v, y) .$$

Wir wenden nun auf (55) das primale Schnittverfahren an, also (25), (26), wobei wir $Y^n = \{y^o, \ldots, y^{n-1}\}$ wählen, und $\max_{Y^n} \varphi(x,y)$ als $\max_\nu \varphi(x,y^\nu)$ schreiben $(\nu = 0, \ldots, n-1)$. Das ergibt

(56) $$\xi^n : \min_X \left[\max_\nu \min_V \psi(x, v, y^\nu) \right] ,$$

(57) $$y^n : \max_Y \min_V \psi(\xi^n, v, y) .$$

Wir setzen noch

(58) $$v^n : \min_V \max_Y \psi(\xi^n, v, y) .$$

Sei nun $(\hat{\xi}, \bar{v}, \bar{y})$ ein Häufungspunkt von (ξ^n, v^n, y^n). Aus (56), (57) folgt gemäss Satz 1 (in seiner primalen Form), dass $\hat{\xi}$ eine Lösung von (55) und damit von (54) darstellt. Aus (57) und (58) folgt durch einen einfachen Grenzübergang und wegen des Satzes von Kakutani, dass $(\bar{v}, \bar{y})$ ein Sattelpunkt von $\psi(\hat{\xi}, v, y)$ ist. Damit ist $(\hat{\xi}, \bar{v})$ Lösung von (53).

Wir betrachten nun insbesondere den Fall, dass

(59) $$\psi(x, v, y) = \psi_1 (x,y) + \psi_2 (v,y).$$

Aus (57), (58) folgt, dass (v^n, y^n) ein Sattelpunkt von $\psi(\xi^n, v, y)$ ist, insbesondere

$$\psi(\xi^n, v^n, y^n) \leq \psi(\xi^n, v, y^n) \quad \forall v \in V ,$$

das heisst

$$\psi_2(v^n, y^n) \leq \psi_2(v, y^n) \quad \forall v \in V .$$

Es gibt also für alle ν

$$\min_V \psi_2(v, y^\nu) = \psi_2(v^\nu, y^\nu)$$

und somit für alle $x \in X$, alle ν

$$\min_V \psi(x, v, y^\nu) = \psi(x, v^\nu, y^\nu) .$$

Setzt man dies in (56) ein, so erhält man das Verfahren

(60) $$\xi^n : \min_X \max_\nu \psi(x, v^\nu, y^\nu) ,$$

(61) $$y^n : \max_Y \min_V \psi(\xi^n, v, y) ,$$

(62) $$v^n : \min_V \max_Y \psi(\xi^n, v, y) ,$$

Zur Lösung von (53) unter der speziellen Annahme (59). Dieses Verfahren wurde von Kronsjö für den Fall der Lagrangefunktion beschrieben. Hat man nämlich das Programm

$$\begin{aligned} \min \quad & F(x) + G(v) , \\ & f(x) + g(v) \leq 0 , \\ & x \in X , \ v \in V , \end{aligned}$$

so lässt sich dies in der Form (53) schreiben, wobei die Lagrangefunktion

$$\psi(x, v, y) = F(x) + G(v) + y^T[f(x) + g(v)]$$

der Bedingung (59) genügt; $Y = \{y | y \geq 0\}$. (61), (62) besagt dann, dass v^n eine Lösung von

$$\min \; F(\xi^n) + G(v)$$
$$f(\xi^n) + g(v) \le 0$$
$$v \in V$$

darstellt, mit y^n als zugehörigem Kuhn-Tucker-Multiplikator (Existenz und gleichmässige Beschränktheit der y^n müssen wieder durch eine zusätzliche Regularitätsvoraussetzung sichergestellt werden). Benders betrachtete ursprünglich den Fall, dass $G(v)$ und $g(v)$ linear sind:

(63) $$P : \min \; F(x) + c^T v$$
$$f(x) + A\,v \le 0$$
$$x \in X \;,\; v \ge 0 \quad .$$

(64) $$\xi^n : \min_X \; [\, F(x) + \max_\nu \; (y^\nu)^T f(x) \,] \; .$$

(65) $$v^n : \min \; c^T v$$
$$f(\xi^n) + A\,v \le 0$$
$$v \ge 0 \quad .$$

(66) $$y^n : \max \; y^T f(\xi^n)$$
$$A^T y + c \ge 0$$
$$y \ge 0 \quad .$$

Hierbei hat man wieder den Multiplikator y^n als Lösung des zu (65) dualen linearen Programms (66) ausgedrückt. Die vereinfachte Form von (64) gegenüber (60) rührt daher, dass gemäss der Dualitätstheorie für lineare Programme ,

$$(v^\nu)^T(A^Ty^\nu + c) = 0$$

("complementary slackness condition"), somit

$$\psi(x, v^\nu, y^\nu) = F(x) + c^Tv^\nu + (y^\nu)^T[f(x) + Av^\nu]$$
$$= F(x) + (y^\nu)^Tf(x) \quad .$$

Das Verfahren von Benders ist ein schönes Beispiel dafür, wie sich die zuvor beschriebenen Verfahren (hier: das primale Schnittverfahren) "erweitern" lassen, indem man die Funktion $\varphi(x,y)$ implizit als Extremumsfunktion definiert. Im Zusammenhang mit derartigen Erweiterungen ist die folgende Bemerkung nützlich.

Hilfssatz. - Seien X, V, Y, Z konvex, und sei die Funktion $F(x, v, y, z)$ konvex in (x,v) und konkav in (y,z). Dann ist

$$\varphi(x,y) = \inf_V \sup_Z F(x, v, y, z)$$

konvex in x und konkav in y.

Beweis: Wir zeigen zunächst die Konvexität in x (dabei unterdrücken wir die Abhängigkeit von y). $\psi(x,v) = \sup_Z F(x, v, z)$ ist konvex in (x,v) als Supremum einer Familie von konvexen Funktionen. Um zu zeigen, dass $\varphi(x) = \inf_V \psi(x,v)$ konvex ist, wählen wir $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \beta = 1$.

$$\alpha\,\varphi(x_1) + \beta\varphi\,(x_2) = \inf_{\substack{v_1 \in V \\ v_2 \in V}} [\,\alpha\,\psi(x_1, v_1) + \beta\,\psi(x_2, v_2)\,]$$

$$\geq \inf_{\substack{v_1 \in V \\ v_2 \in V}} \psi\,(\alpha\,x_1 + \beta x_2, \alpha v_1 + \beta v_2) \quad [\text{ Konvexität von } \psi\,]$$

$$= \inf_{v \in \alpha V + \beta V} \psi(\alpha x_1 + \beta x_2, v)$$

$$= \inf_{v \in V} \psi(\alpha x_1 + \beta x_2, v) \quad [\text{Konvexität von } V]$$

$$= \varphi(\alpha x_1 + \beta x_2) \ .$$

Damit ist φ konvex in x. Mit den gleichen Argumenten zeigt man die Konkavität in y.

Dekomposition bei grossen, strukturierten Problemen. Falls

$$\psi(x, v, z) = \sum_i \psi_i \ (x, v_i, y_i) \ ,$$

$$Y = \mathop{\times}_i Y_i, \quad V = \mathop{\times}_i V_i \ ,$$

so zerfallen (61), (62) in kleinere Teilprobleme der Form

$$\left.\begin{array}{l} \max\limits_{Y_i} \ \min\limits_{V_i} \\ \\ \min\limits_{V_i} \ \max\limits_{Y_i} \end{array}\right\} \quad \psi_i(\xi^n, v_i, y_i) \ .$$

Falls $\varphi(x,y) = \sum_i \varphi_i(x_i, y)$, $X = \mathop{\times}_i X_i$, so zerfällt (33) in kleinere Teilprobleme der Form

$$\min_{X_i} \ \varphi_i(x_i, \eta^n) \ .$$

Dies ist praktisch von erheblicher Tragweite.

LITERATUR

[1] G.B. DANTZIG: General convex objective forms. In: *Mathematical Methods in the Social Sciences* (K.J. Arrow, S. Karlin, P. Suppes ed.) pp. 151-158. Stanford University Press, Stanford, 1960.

[2] J.F. BENDERS: Partitioning procedures for solving mixed-variables programming problems. Numer. Math. 4 (1962), 238-252.

[3] P. BROISE, P. HUARD, J. SENTENAC: Décomposition des programmes mathématiques. Dunod, Paris, 1968.

[4] T.O.M. KRONSJÖ: Decomposition of a nonlinear convex separable economic system in primal and dual directions. In: *Optimization* (R. Fletcher ed.), pp. 85-97. Academic Press, London, 1969.

[5] D.M. HIMMELBLAU (editor): Decomposition of Large-Scale Problems. North-Holland Elsevier, Amsterdam (to appear).

[6] J. STOER: Duality in nonlinear programming and the minimax theorem. Numer. Math. 5 (1963), 371-379.

[7] M. FRANK, P. WOLFE: An algorithm for quadratic programming. Naval Res. Logist. Quart. 3 (1956), 95-110.