Decomposition Schemes for Finding Saddle Points of Quasi-Convex-Concave Functions

Werner Oettli

Nr.77 (1987)

Decomposition Schemes for Finding Saddle Points of Quasi-Convex-Concave Functions

Werner Oettli

1. Introduction

Decomposition methods for finding saddle points of a function $\varphi : X \times Y \to I\!R$ are characterized by an alternating succession of master programs and subprograms [7]. The master programs determine the proper iteration points, which are approximate saddle points over a subset $X^n \times Y^n$ of the original domain. The subprograms calculate auxiliary points, which serve to update the subset under consideration. For certain structured problems the subprograms may decompose; this fact accounts for the name and the practical importance of decomposition methods, but is not essential for their mathematical theory of convergence, which is our main concern here.

In [12] a symmetric decomposition scheme for finding convex-concave saddle points has been described, which subsumed several previously known methods. More recently in [1] another general scheme has been described, which introduced regularisation in solving the subprograms. The present paper synthesizes these two approaches and may also be viewed as a survey of some classical decomposition methods.

We work essentially within the setting of Sion's celebrated minimax theorem [15], i.e. the function φ considered will be quasi-convex-concave. The compactness assumption for the underlying domain, which has already been relaxed in [1], will be further weakened. We also admit more general regularizing functions than in [1]. The procedure is described in such a way that the extension to Nash equilibrium points in n-person games is straightforward.

We recall that a function $f : X \to I\!R$ is called quasiconvex iff the level sets $\{x \in X | f(x) \leq \alpha\}$ are convex for all $\alpha \in I\!R$. Furthermore f is called quasiconcave iff $-f$ is quasiconvex. A function $\varphi : X \times Y \to I\!R$ is called quasi-convex-concave iff $\varphi(\cdot, y) : X \to I\!R$ is quasiconvex for all $y \in Y$ and $\varphi(x, \cdot) : Y \to I\!R$ is quasiconcave for all $x \in X$. We denote by $conv A$ the convex hull of the set A.

2. Preliminaries

Let there be given two nonvoid sets X, Y and a function $\varphi(x,y) : X \times Y \to I\!R$.

A point $(\xi^*, \eta^*) \in X \times Y$ is called a saddle point of φ on $X \times Y$ iff

$$\varphi(\xi^*, y) \le \varphi(x, \eta^*) \quad \forall (x,y) \in X \times Y. \tag{1}$$

Henceforth we shall use the functions

$$M(x) := \sup_{y \in Y} \varphi(x,y) : X \to I\!R \cup \{+\infty\}, \tag{2}$$

$$m(y) := \inf_{x \in X} \varphi(x,y) : Y \to I\!R \cup \{-\infty\}. \tag{3}$$

Condition (1) can then be written as $M(\xi^*) \le m(\eta^*)$, and this inequality can only be satisfied as an equality, since one has always $m(y) \le \varphi(x,y) \le M(x)$ for arbitrary $(x,y) \in X \times Y$. If (ξ^*, η^*) is a saddle point, then ξ^* solves the primal problem

(P): $$\inf\{M(x) | x \in X\}, \tag{4}$$

and η^* solves the dual problem

(D): $$\sup\{m(y) | y \in Y\}, \tag{5}$$

and the extreme values are equal. Hence if the set of saddle points is nonempty, then it consists of all the pairs (ξ^*, η^*) where ξ^* solves (P) and η^* solves (D).

Let us now assume

(H1) X and Y are nonvoid closed convex sets in some normed linear spaces;

$\varphi : X \times Y \to I\!R$ is quasi-convex-concave and continuous.

Under (H1) the function $M(\cdot)$, being the supremum of a family of lower semicontinuous functions, is again lower semicontinuous. Likewise $m(\cdot)$ is upper semicontinuous. By Sion's minimax theorem [15, 9] φ has a saddle point over $X \times Y$ if, in addition to (H1), X and Y are compact. The latter requirement can be weakened as follows:

(H2) We are given a nonempty finite subset $Z^0 \subset X \times Y$ such that the set

$S := \{(\xi, \eta) \in X \times Y | \varphi(\xi, y) \le \varphi(x, \eta) \quad \forall (x,y) \in Z^0\}$ is compact.

Lemma 1. Under (H1) and (H2) φ has saddle points on $X \times Y$, and all saddle points lie in S.

Proof.

Assume that there exists no saddle point $(\xi^*, \eta^*) \in X \times Y$ satisfying (1). Consider the sets

$$S(x,y) := \{(\xi, \eta) \in X \times Y | \varphi(\xi, y) \le \varphi(x, \eta)\}, \ (x,y) \in X \times Y.$$

Then the family of closed sets $\{S(x,y)|(x,y) \in X \times Y\}$ has empty intersection over S. Since $S = \bigcap\{S(x,y)|(x,y) \in Z^0\}$ is compact by (H2), there exist finite subsets $X^1 \subset X$ and $Y^1 \subset Y$ with $Z^0 \subset X^1 \times Y^1$ such that the family $\{S(x,y)|(x,y) \in X^1 \times Y^1\}$ has empty intersection over $X \times Y$. Let $\Xi^1 := convX^1$ and $H^1 := convY^1$. These sets are convex and compact, and by Sion's original result there exists a saddle point $(\xi^1, \eta^1) \in \Xi^1 \times H^1$ of φ over $\Xi^1 \times H^1$. But then

$$(\xi^1, \eta^1) \in \bigcap\{S(x,y)|(x,y) \in X^1 \times Y^1\} \neq \emptyset,$$

a contradiction. Hence φ has saddle points over $X \times Y$.
If (ξ^*, η^*) is a saddle point over $X \times Y$, then clearly $\varphi(\xi^*, y) \leq \varphi(x, \eta^*)$ for all $(x,y) \in Z^0$, hence $(\xi^*, \eta^*) \in S$. q.e.d.

Remark. A closer inspection of the proof of the lemma shows that in (H2) the requirement of Z^0 being finite can be replaced by the requirement that $Z^0 = X^0 \times Y^0$ with $convX^0$ and $convY^0$ being compact. In particular, if Y itself is compact and $\{\xi \in X|M(\xi) \leq M(\overline{x})\}$ is compact for some $\overline{x} \in X$, then (H2) is satisfied with $Z^0 := \{\overline{x}\} \times Y$, since in this case

$$S = \{(\xi, \eta) \in X \times Y|M(\xi) \leq \varphi(\overline{x}, \eta)\} \subset \{\xi \in X|M(\xi) \leq M(\overline{x})\} \times Y,$$

and S is compact.

Example. We go through an example in detail to illustrate the usefulness of hypothesis (H2). We assume in addition to (H1) that Y is a cone, and that $\varphi(x,y) := f(x) + g(x,y)$ with $g(x,\cdot)$ positively homogeneous in Y for every $x \in X$. Then

$$M(x) := \begin{cases} f(x) & \text{if } g(x,y) \leq 0 \quad \forall y \in Y \\ +\infty & \text{else,} \end{cases}$$

and the primal problem (4) becomes

(P'): $$\inf\{f(x)|x \in X,\ g(x,y) \leq 0 \quad \forall y \in Y\}. \tag{6}$$

We assume further that $f(\cdot)$ is inf-compact, meaning that the level sets $\{x \in X|f(x) \leq \alpha\}$ are compact for all $\alpha \in I\!R$, and we assume that the following Slater-type regularity assumption (RA) is satisfied:

(RA) Y is locally compact, and there exists a finite subset $X^0 \subset X$ such that

$$\min_{x \in X^0} g(x,y) < 0 \quad \forall y \in Y \setminus \{0_Y\}.$$

Then hypothesis (H2) is satisfied with $Z^0 := X^0 \times \{0_Y\}$.

Indeed: From local compactness of Y follows the existence of a compact subset $B \subset Y \setminus \{0_Y\}$ such that $Y = I\!R_+ \cdot B$, and since $g(x, \cdot)$ is upper semicontinuous (RA) implies the existence of $\delta > 0$ such that $\min_{x \in X^0} g(x, y) \leq -\delta \quad \forall y \in B$. Let $\beta \in I\!R$ be arbitrary, and choose $k \in I\!R$ such that $k \geq \max_{x \in X^0} f(x)$ and $k \geq \beta$. Since $g(x, \cdot)$ is positively homogeneous it follows for all $\lambda \geq 0$:

$$\min_{x \in X^0} (f(x) + g(x, y)) \leq k - \lambda\delta \quad \forall y \in \lambda B.$$

In particular if $\lambda > \delta^{-1}(k - \beta) =: r_0$ we have $\min_{x \in X^0} \varphi(x, y) < \beta \quad \forall y \in \lambda B$. Hence for $\lambda > r_0$ the set λB is disjoint with the level set $\{y \in Y \mid \min_{x \in X^0} \varphi(x, y) \geq \beta\}$; the latter is therefore contained in the compact set $[0, r_0] \cdot B$ and is itself compact. Now choose $Z^0 := X^0 \times \{0\}$. Then

$$\begin{aligned} S = \; & \{(\xi, \eta) \in X \times Y \mid \varphi(\xi, 0) \leq \varphi(x, \eta) \quad \forall x \in X^0\} \\ \subset \; & \{\xi \in X \mid f(\xi) \leq \alpha^0\} \times \{\eta \in Y \mid \min_{x \in X^0} \varphi(x, \eta) \geq \beta^0\}, \end{aligned}$$

where $\alpha^0 := \sup_Y \min_{X^0} \varphi(x, y)$ and $\beta^0 := \inf_X f(x)$. α^0 and β^0 are finite since the functions being extremized are continuous and have compact level sets. Now S, being contained in a compact set, is itself compact, and so (H2) is satisfied. q.e.d.

Hence under the assumptions of the example φ has a saddle point (ξ^*, η^*). By what has been said previously this implies that ξ^* solves (P') and that $f(\xi^*) \leq f(x) + g(x, \eta^*)$ $\forall x \in X$.

3. Decomposition Principle

For notational simplicity it is convenient to represent the saddle point problem (1) as a variational inequality problem. Set $Z := X \times Y$ and define $\Phi(\cdot, \cdot) : Z \times Z \to I\!R$ by

$$\Phi(z, \varsigma) := \varphi(x, \eta) - \varphi(\xi, y), \; z := (x, y), \; \varsigma := (\xi, \eta). \tag{7}$$

Then the problem of finding (ξ^*, η^*) satisfying (1) is equivalent to finding $\varsigma^* = (\xi^*, \eta^*)$ satisfying

$$\varsigma^* \in Z, \; \Phi(z, \varsigma^*) \geq 0 \quad \forall z \in Z. \tag{8}$$

Note that $\Phi(\varsigma, \varsigma) = 0$ for all $\varsigma \in Z$. Moreover Φ is continuous from (H1). But $\Phi(\cdot, \varsigma)$ is not necessarily quasiconvex (unless φ is convex-concave, in which case $\Phi(\cdot, \varsigma)$ is convex). Note that ς^* satisfying (8) maximizes the function $\inf_{z \in Z} \Phi(z, \cdot)$ over Z. Under (H1) and (H2) problem (8) has a solution. Moreover any solution of (8) is in S. This result obviously remains true if in (8) we replace the set Z by a closed, convex product set $\Omega = \Xi \times H$ with $Z^0 \subset \Omega \subset Z$.

Several other assumptions which we have to make are collected in the following hypothesis (H3), where Φ, S, Z have the same meaning as before.

(H3) a) We are given a lower semicontinuous function $H : Z \times Z \to I\!R \cup \{+\infty\}$ with $H(z,\varsigma) \geq 0$ for all $(z,\varsigma) \in Z \times Z$ and $H(z,z) = 0$ for all $z \in Z$.

b) The function $\varsigma \mapsto \inf_{z \in Z} (\Phi + H)(z,\varsigma)$ is upper semicontinuous.

c) $(\Phi + H)(z,\varsigma) \geq 0$ for all $z \in Z$ implies $\Phi(z,\varsigma) \geq 0$ for all $z \in Z$.

d) There exists a compact set $K \subset Z$ such that $\{z \in Z | (\Phi + H)(z,\varsigma) \leq 0\} \subset K$ for all $\varsigma \in S$.

Examples.

a) Let Z be finite-dimensional, and let φ be convex-concave, which implies that $\Phi(\cdot,\varsigma)$ is convex. Moreover let Φ be defined on an open neighbourhood of $Z \times Z$. Then from [13;theorem 24.7] it follows that the subgradients of $\Phi(\cdot,\varsigma)$ are uniformly bounded on the compact set $S \times S$. Hence

$$\Phi(z,\varsigma) = \Phi(z,\varsigma) - \Phi(\varsigma,\varsigma) \geq -k\|z-\varsigma\| \quad \forall z \in Z,\ \forall \varsigma \in S.$$

Now choose $H(z,\varsigma) := \|z-\varsigma\|^2$. Then $(\Phi + H)(z,\varsigma) \leq 0$ $(z \in Z,\ \varsigma \in S)$ implies $0 \geq -k\|z-\varsigma\| + \|z-\varsigma\|^2$, and thereby $\|z-\varsigma\| \leq k$. Hence there exists a set K as required in (H3)d). Moreover, if H is choosen this way, and $\Phi(\cdot,\varsigma)$ is convex, then (H3)a)–c) is trivially satisfied. If Φ is continuously differentiable, one can also choose

$$(\Phi + H)(z,\varsigma) := \Phi(\varsigma,\varsigma) + \langle \nabla_1 \Phi(\varsigma,\varsigma), z-\varsigma \rangle + k \cdot \|z-\varsigma\|^2,$$

with $k > 0$ so large that $H \geq 0$.

b) Let Z be finite-dimensional again and consider the choice

$$H(z,\varsigma) := \begin{cases} 0 & \text{for } \|z-\varsigma\| \leq \rho \\ +\infty & \text{else} \end{cases}$$

for some $\rho > 0$. Then (H3)a), b), d) are satisfied. Moreover (H3)c) is satisfied if every local minimum of $\Phi(\cdot,\varsigma)$ is a global minimum.

In what follows we always assume that (H1), (H2), (H3) hold. For ease of notation we describe the symmetric decomposition scheme in terms of Φ and Z above. From this basic model we derive by specialization various unsymmetric implementations.

The decomposition method

At the start we are given the finite subset $Z^0 \subset Z$ from assumption (H2). At the beginning of the n-th iteration we need the previously calculated auxiliary points $z^1, \cdots, z^{n-1} \in Z$. Choose a set Z^n such that $Z^0 \cup \{z^1, \cdots, z^{n-1}\} \subset Z^n \subset Z$. Choose a product set $\Omega^n = \Xi^n \times H^n$ closed, convex, such that $Z^n \subset \Omega^n \subset Z$.

Master program:

$$\text{Select the iteration point } \varsigma^n \in \Omega^n \text{ such that } \quad \Phi(z, \varsigma^n) \geq 0 \quad \forall z \in Z^n. \tag{9}$$

Subprogram:

$$\begin{gathered}\text{Select the auxiliary point } z^n \in Z \text{ such that} \\ (\Phi + H)(z^n, \varsigma^n) \leq (\Phi + H)(z, \varsigma^n) \quad \forall z \in Z.\end{gathered} \tag{10}$$

This completes the n-th iteration.

We convince ourselves that the above rules are consistent. The existence of ς^n satisfying (9) follows from the fact that φ has a saddle point on $\Xi^n \times H^n$. Moreover, since $Z^0 \subset Z^n$, we have $\varsigma^n \in S$. In addition, since $(\Phi + H)(\varsigma^n, \varsigma^n) = 0$, it follows from assumption (H3)d) that the lower semicontinuous function $(\Phi + H)(\cdot, \varsigma^n)$ assumes its minimum over Z within the compact set K. Hence z^n satisfying (10) exists, and moreover $z^n \in K$.

As a stopping criterion we introduce the quantity

$$\tau_n := (\Phi + H)(z^n, \varsigma^n). \tag{11}$$

It follows from (10) that $\tau_n \leq 0$ and

$$\tau_n \leq (\Phi + H)(z, \varsigma^n) \quad \forall z \in Z. \tag{12}$$

If $z^n = z^{k_0}$ for some $k_0 < n$, then $\tau_n = 0$, since (9) and $H \geq 0$ imply

$$0 \leq (\Phi + H)(z^{k_0}, \varsigma^n) = (\Phi + H)(z^n, \varsigma^n) = \tau_n \leq 0.$$

In this case it follows from (12) and (H3)c) that ς^n is already an exact solution of (8), i.e. the algorithm terminates after finitely many steps. If this case does not occur, for the sequence $\{\varsigma^n\}$ generated by the above rules we have

Theorem 1. The sequence $\{\varsigma^n\}$ has cluster points, and every cluster point is a solution of (8). Moreover, for the quantity $\tau_n := (\Phi + H)(z^n, \varsigma^n)$ there holds $\lim_{n \to \infty} \tau_n = 0$.

Proof. Since $\varsigma^n \in S$, where S is compact, it follows that the sequence $\{\varsigma^n\}$ has cluster points. Let ς^* be an arbitrary cluster point of this sequence. Since $z^n \in K$, and K is compact, the sequence $\{(z^n, \varsigma^n)\}$ contains a subsequence $\{(z^{n(j)}, \varsigma^{n(j)})\}$ $(j \in I\!N)$ converging to (z^*, ς^*) for some $z^* \in K$. From (9) we have $0 \leq \Phi(z^k, \varsigma^n) \quad \forall k < n$. In particular, for $n := n(j)$, we obtain in the limit $0 \leq \Phi(z^k, \varsigma^*) \quad \forall k$. Now, for $k := n(j)$ we obtain in the limit

$$0 \leq \Phi(z^*, \varsigma^*).$$

From (10) and $H \geq 0$ follows $\Phi(z^n, \varsigma^n) \leq (\Phi + H)(z, \varsigma^n) \quad \forall z \in Z$. By (H3)b) this yields in the limit

$$\Phi(z^*, \varsigma^*) \leq (\Phi + H)(z, \varsigma^*) \quad \forall z \in Z.$$

Altogether we have

$$0 \leq (\Phi + H)(z, \varsigma^*) \quad \forall z \in Z. \tag{13}$$

From assumption (H3)c) it follows that

$$0 \leq \Phi(z, \varsigma^*) \quad \forall z \in Z.$$

Hence ς^* is a solution of (8). Moreover, it follows from (13) that

$$0 \leq (\Phi + H)(z^*, \varsigma^*).$$

But $0 \geq (\Phi + H)(z^n, \varsigma^n)$ for all n, as stated above, and due to the lower semicontinuity of $\Phi + H$ we obtain $(\Phi + H)(z^*, \varsigma^*) = 0$. In view of our compactness assumption this means that $\tau_n \to 0$ for the entire sequence $\{\tau_n\}$. q.e.d.

In the absence of further information about H beyond that given in (H3), the condition $-\varepsilon \leq \tau_n$ (where $\varepsilon > 0$ is given) may be used as a convenient stopping criterion to terminate the procedure.

<u>**Example**</u>. The cutting method for the problem

$$\max_{\varsigma \in Z} \big(\inf_{z \in Z} \Phi(z, \varsigma) \big),$$

where Z is compact and Φ is continuous, runs as follows:

$$\varsigma^n \text{solves } \max_{\varsigma \in Z} \big(\inf_{z \in Z^n} \Phi(z, \varsigma) \big),$$
$$z^n \text{solves } \min_{z \in Z} \Phi(z, \varsigma^n),$$

where $z^0 \in Z$ is arbitrary and $\{z^0, z^1, \cdots, z^{n-1}\} \subset Z^n \subset Z$. Clearly z^n, ς^n satisfify (9) and (10) with $\Omega^n = Z$ and $H \equiv 0$, provided (8) is solvable. If Φ is given by (7) and $Z^n = X^n \times Y^n$, then the cutting method with $\varsigma^n = (\xi^n, \eta^n)$, $z^n = (x^n, y^n)$ decomposes as follows:

$$\xi^n \text{ solves } \min_{\xi \in X} \big(\sup_{y \in Y^n} \varphi(\xi, y) \big), \tag{14}$$

$$\eta^n \text{ solves } \max_{\eta \in Y} \big(\inf_{x \in X^n} \varphi(x, \eta) \big), \tag{15}$$

$$x^n \text{ solves } \min_{x \in X} \varphi(x, \eta^n), \tag{16}$$

$$y^n \text{ solves } \max_{y \in Y} \varphi(\xi^n, y). \tag{17}$$

Hence we obtain two parallel, unrelated algorithms: The first one, involving ξ^n, y^n and given by (14), (17) is the cutting method for solving $\min_{\xi \in X} M(\xi)$. The second one, involving η^n, x^n and given by (15), (16) is the cutting method for solving $\max_{\eta \in Y} m(\eta)$.

In the remaining parts we shall restrict ourselves to the case where $Z^n = X^n \times Y^n$ for all n and

$$H(z, \varsigma) := F(x, \xi) + G(y, \eta)$$

with $F : X \times X \to I\!R$ and $G : Y \times Y \to I\!R$ continuous, nonnegative functions satisfying

$$F(x, x) = 0 \quad \forall x \in X, \; G(y, y) = 0 \quad \forall y \in Y,$$

$$\varphi(x, \eta) + F(x, \xi) \geq \varphi(\xi, \eta) \quad \forall x \in X \text{ implies } \varphi(x, \eta) \geq \varphi(\xi, \eta) \quad \forall x \in X, \tag{18}$$

$$\varphi(\xi, y) - G(y, \eta) \leq \varphi(\xi, \eta) \quad \forall y \in Y \text{ implies } \varphi(\xi, y) \leq \varphi(\xi, \eta) \quad \forall y \in Y. \tag{19}$$

Then (H3)a), b), c) are satisfied. We have

$$(\Phi + H)(z, \varsigma) = \varphi(x, \eta) + F(x, \xi) - (\varphi(\xi, y) - G(y, \eta)),$$

and therefore (9), (10) take the following form:

a) Select $(\xi^n, \eta^n) \in \Xi^n \times H^n$ such that

$$\varphi(\xi^n, y) \leq \varphi(x, \eta^n) \quad \forall x \in X^n, \; \forall y \in Y^n. \tag{20}$$

b) Select $x^n \in X$ such that

$$x^n \text{ solves } \min_{x \in X} (\varphi(x, \eta^n) + F(x, \xi^n)). \tag{21}$$

Select $y^n \in Y$ such that

$$y^n \text{ solves } \max_{y \in Y} (\varphi(\xi^n, y) - G(y, \eta^n)). \tag{22}$$

Here $X^n \subset X$ and $\Xi^n \subset X$ have to be chosen such that $X^0 \cup \{x^1, \cdots, x^{n-1}\} \subset X^n \subset \Xi^n$ with Ξ^n closed, convex. Similarly for Y^n and H^n.

If $H^n = Y^n = Y$ for all n, then the auxiliary points y^n are not needed for the procedure, and (22) becomes void. In this case, if only $\xi^n \in X^n$, we can use

$$\tau_n := \varphi(x^n, \eta^n) - \varphi(\xi^n, \eta^n) + F(x^n, \xi^n)$$

as stopping criterion.

We remark that if the problem is separable in the following sense:

$$\varphi(x,y) := \sum_{i,j} \varphi_{ij}(x_i, y_j), \; X = \prod_i X_i, \; Y = \prod_j Y_j,$$

$$F(x,\xi) := \sum_i F_i(x_i, \xi_i), \; G(y,\eta) := \sum_j G_j(y_j, \eta_j),$$

then the subprograms (21) and (22) disaggregate into subproblems of smaller size.

4. Variants

Let us consider in more detail a particular implementation of the procedure (20)–(22). We assume that φ is convex-concave and that $X^0 = \{x^0\}$ is a singleton. For the n-th iteration, given $x^0, \cdots, x^{n-1}$, we let

$$\Sigma^n := \{\lambda = (\lambda_0, \cdots, \lambda_{n-1}) \in \mathbb{R}^n | \lambda \geq 0, \; \sum_{i=0}^{n-1} \lambda_i = 1\},$$

and for $\lambda \in \Sigma^n$ and $y \in Y$ we define

$$\psi_n(\lambda, y) := \sum_{i=0}^{n-1} \lambda_i \varphi(x^i, y).$$

H^n is as before. The n-th iteration consists of the following steps:

$$\text{Determine } (\lambda^n, \eta^n) \text{ with } \lambda^n := (\lambda_0^n, \cdots, \lambda_{n-1}^n) \text{ as a saddle point of } \psi_n(\lambda, y) \text{ on } \Sigma^n \times H^n. \tag{23}$$

$$\text{Set } \xi^n := \sum_{i=0}^{n-1} \lambda_i^n x^i. \tag{24}$$

Determine x^n, y^n according to (21), (22).

Setting $X^n = \{x^0, \cdots, x^{n-1}\}$ and $\Xi^n := conv(x^0, \cdots, x^{n-1})$, we have $\xi^n \in \Xi^n$, and the couple (ξ^n, η^n) defined above satisfies (20), since

$$\psi_n(\lambda^n, y) \leq \psi_n(\lambda, \eta^n) \quad \forall \lambda \in \Sigma^n, \; \forall y \in H^n$$

implies

$$\varphi(\xi^n, y) \leq \varphi(x^i, \eta^n) \quad \forall i = 0, \cdots, n-1, \; \forall y \in H^n.$$

The existence of a saddle point of ψ_n over $\Sigma^n \times H^n$ is guaranteed, since the validity of (H2) for φ with regard to $X \times Y$ implies the validity of (H2) for ψ_n with regard to $\Sigma^n \times Y$.

Note that η^n can be found as a solution of

$$(\mathrm{D}_n): \qquad \max_{\eta \in H^n} \Big(\inf_{\lambda \in \Sigma^n} \sum_{i=0}^{n-1} \lambda_i \varphi(x^i, \eta) \Big) = \max_{\eta \in H^n} \Big(\min_{i=0,\cdots,n-1} \varphi(x^i, \eta) \Big),$$

and λ^n can be found as a solution of

$$(\mathrm{P}_n): \qquad \min_{\lambda\in\Sigma^n}\Big(\sup_{y\in H^n}\sum_{i=0}^{n-1}\lambda_i\varphi(x^i,y)\Big).$$

If one chooses $Y^n = H^n = Y$ for all n (which implies that y^n is superfluous), one obtains algorithm 2 of [1]. Provided that in addition $F = G = 0$ one obtains for algorithm (23), (24), (21) the following two-sided bounds for the saddle value $\varphi(\xi^*,\eta^*)$:

$$\varphi(x^n,\eta^n) = m(\eta^n) \le \varphi(\xi^*,\eta^*) \le M(\xi^n) \le \sup_{y\in Y}\psi_n(\lambda^n,y) = \psi_n(\lambda^n,\eta^n),$$

where the last inequality follows from the convexity of $\varphi(\cdot,y)$.

If $\varphi(x,y) := f(x) + g(x,y)$ with Y a cone and $g(x,\cdot)$ positively homogeneous in Y for every $x \in X$, then the primal problem, as already stated, becomes

$$(\mathrm{P'}): \qquad \inf\{f(x)|x\in X,\ g(x,y)\le 0 \quad \forall y\in Y\},$$

and with $Y^n = H^n = Y$ step (23) reads as follows:

$$(\mathrm{P}'_n):\ \lambda^n \text{solves} \quad \min_{\lambda\in\Sigma^n}\Big(\sup_{y\in Y}\sum_i\lambda_i(f(x^i)+g(x^i,y))\Big)$$
$$= \min\{\sum_i\lambda_i f(x^i)|\lambda\in\Sigma^n,\ \sum_i\lambda_i g(x^i,y)\le 0 \quad \forall y\in Y\},$$

$$(\mathrm{D}'_n):\ \eta^n \text{solves} \quad \max_{\eta\in Y}\Big(\inf_{\lambda\in\Sigma^n}\sum_i\lambda_i(f(x^i)+g(x^i,\eta))\Big)$$
$$= \max_{\eta\in Y}\big(\min_i(f(x^i)+g(x^i,\eta))\big).$$

Note that the convexity of $g(\cdot,y)$ implies that ξ^n is feasible for (P').

Finally, in case that $Y = I\!R^n_+$ and $\varphi(x,y) := f(x) + \langle y, g(x)\rangle$, $g: X \to I\!R^n$, (P') reads

$$(\mathrm{P''}): \qquad \inf\{f(x)|x\in X,\ g(x)\le 0\},$$

and (P'_n) and (D'_n) become a pair of dual linear programming problems:

$$(\mathrm{P}''_n): \qquad \min\{\sum_i\lambda_i f(x^i)|\lambda\in\Sigma^n,\ \sum_i\lambda_i g(x^i)\le 0\},$$

$$(\mathrm{D}''_n): \qquad \max\{\min_i(f(x^i)+\langle\eta,g(x^i)\rangle)|\eta\ge 0\}.$$

This method with $F = 0$ has been given by Dantzig [7, Ch.24.1] and – with a different motivation – in [16, Ch.14.4]. With $F = 0$, but φ arbitrary, algorithm (23), (24), (21) has been described in [12] as an extension of Dantzig's method for φ not necessarily being a classical Lagrangian.

Algorithm 1 of [1] is obtained if one treats y in the same way as x: One defines

$$\psi_n(\lambda,\mu) := \sum_{i,j}\lambda_i\mu_j\varphi(x^i,y^j),$$

where i, j run from 0 to $n-1$, one requires (λ^n, μ^n) to be a saddle point of $\psi_n(\lambda, \mu)$ over $\Sigma^n \times \Sigma^n$, and one sets

$$\xi^n := \sum_i \lambda_i^n x^i, \quad \eta^n := \sum_j \mu_j^n y^j.$$

The determination of (λ^n, μ^n) is a dual pair of linear programming problems:

$$\lambda^n \text{ solves } \min_{\lambda \in \Sigma^n} (\max_j \sum_i \lambda_i \varphi(x^i, y^j)),$$

$$\mu^n \text{ solves } \max_{\mu \in \Sigma^n} (\min_i \sum_j \mu_j \varphi(x^i, y^j)).$$

(20) is again satisfied with $X^n := \{x^0, \cdots, x^{n-1}\}$, $\Xi^n := conv(x^0, \cdots, x^{n-1})$, $Y^n := \{y^0, \cdots, y^{n-1}\}$, $H^n := conv(y^0, \cdots, y^{n-1})$.

Huard's method. For the case that φ is an ordinary Lagrangian function Huard [5] has given a modification of Dantzig's decomposition algorithm. We can generalize Huard's method for problem (8)

$$\varsigma^* \in Z, \; \Phi(z, \varsigma^*) \geq 0 \quad \forall z \in Z$$

as follows. We assume that we are given a continuous function $\tilde{H} : Z \times Z \to \mathbb{R}$, which satisfies all the requirements of (H3) with one exception: $H \geq 0$ is replaced by the weaker requirement that for every $a \in Z$

$$\Phi(z, \varsigma) \geq 0 \quad \forall z \in [\varsigma, a] \text{ implies } (\Phi + \tilde{H})(z, \varsigma) \geq 0 \quad \forall z \in [\varsigma, a]. \tag{25}$$

We assume for simplicity that Z is compact and choose Ω^n a convex, compact product set such that $\Omega^n \supset conv\{z^0, z^1, \cdots, z^{n-1}\}$. Then the algorithm reads:

$$\text{Choose } \varsigma^n \in \Omega^n \text{ such that } \Phi(z, \varsigma^n) \geq 0 \quad \forall z \in \Omega^n. \tag{26}$$

$$\text{Choose } z^n \in Z \text{ as a solution of } \min_{z \in Z} (\Phi + \tilde{H})(z, \varsigma^n). \tag{27}$$

Note that (26) is essentially a sharpening of (9) (now $Z^n = \Omega^n$), (27) remains practically the same as (10). The existence of ς^n and z^n with the required properties is ensured, and moreover they lie in a compact set. Any cluster point ς^* of the sequence $\{\varsigma^n\}$ is a solution of (8).

Indeed: There exists a subsequence, indexed by $n(j)$, such that $\varsigma^{n(j)} \to \varsigma^*$, $z^{n(j)} \to z^* \in Z$. From (26) follows in view of (25) that $(\Phi + \tilde{H})(z, \varsigma^n) \geq 0 \quad \forall z \in \Omega^n$. Hence in particular $(\Phi + \tilde{H})(z^k, \varsigma^n) \geq 0 \quad \forall k < n$. In the limit this gives

$$(\Phi + \tilde{H})(z^*, \varsigma^*) \geq 0.$$

From (27) follows in the limit that

$$(\Phi + \tilde{H})(z^*, \varsigma^*) \leq (\Phi + \tilde{H})(z, \varsigma^*) \quad \forall z \in Z.$$

Altogether we have $0 \leq (\Phi + \tilde{H})(z, \varsigma^*) \quad \forall z \in Z$, and in view of (H3)c) this implies

$$0 \leq \Phi(z, \varsigma^*) \quad \forall z \in Z.$$

Hence (8) is satisfied. q.e.d.

Suppose in particular that we choose $\Omega^n = conv(x^0, \cdots, x^{n-1}) \times Y$, in which case the need for calculating y^n disappears. Then we choose $\tilde{H}(z, \varsigma) := \tilde{F}(x, \xi, \eta)$ and subproblem (27) becomes to find

$$x^n \in X \text{ minimizing } \varphi(x, \eta^n) + \tilde{F}(x, \xi^n, \eta^n).$$

In particular, if φ is continuously differentiable, we may choose $\tilde{F}$ in such a way that

$$\varphi(x, \eta^n) + \tilde{F}(x, \xi^n, \eta^n) := \varphi(\xi^n, \eta^n) + \langle \nabla_1 \varphi(\xi^n, \eta^n), x - \xi^n \rangle + k\|x - \xi^n\|^2$$

for some $k \geq 0$. Then condition (25) is satisfied. Condition (H3)c) is satisfied if $\varphi(\cdot, \eta)$ is pseudoconvex.

5. Deletion of auxiliary points

It is an unpleasant feature of the methods described so far that the auxiliary points x^n, y^n have to be stored and used in all subsequent iterations. Here we want to obtain versions which allow for the deletion of auxiliary points. The crucial hypothesis which we need for this is (H4):

(H4) $\varphi(x, \cdot)$ is unimodal, i.e., for all $x \in X$ there exists at most one $y \in Y$ where $\varphi(x, \cdot)$ assumes its supremum over Y.

With this hypothesis the master program (20) in the decomposition method may be drastically simplified towards a method of feasible directions [4]. A first example of the method to be described below has been given in [12]. A more elaborate version for the quadratic case, together with an estimate of the rate of convergence, has been described in [14].

Besides (H4) we make in this section the following additional assumptions:

1) The set Y is compact.

2) For some $\overline{x} \in X$ the set $Q := \{\xi \in X | M(\xi) \leq M(\overline{x})\}$ is compact. Let $\tilde{S} := Q \times Y$.

3) There exists a compact set $K \subset X$ such that

$$\{x \in X | \varphi(x, \eta) + F(x, \xi) \leq \varphi(\xi, \eta)\} \subset K \text{ for all } (\xi, \eta) \in \tilde{S}.$$

From 1) follows the continuity of $M(\cdot)$, and from 1) and 2) follows the existence of saddle points of φ over $X \times Y$ – see the remark after lemma 1. Assumption 3) replaces (H3)d).

The modified iteration runs as follows: For the start we choose $\xi^0 \in Q$ and $x^0 \in K$ arbitrarily. At the beginning of the n-th iteration we are given $\xi^{n-1} \in Q$ and $x^{n-1} \in K$ calculated in the previous iteration. The n-th iteration consists of the following steps a) and b):

a) Select $(\xi^n, \eta^n) \in X \times Y$ such that

$$M(\xi^n) = \varphi(\xi^n, \eta^n),\ M(\xi^n) \leq \varphi(\xi^{n-1}, \eta^n),\ M(\xi^n) \leq \varphi(x^{n-1}, \eta^n). \tag{28}$$

b) Select $x^n \in X$ such that

$$x^n \text{ solves } \min_{x \in X} (\varphi(x, \eta^n) + F(x, \xi^n)). \tag{29}$$

The requirements under a) are consistent, since any saddle point (ξ^n, η^n) of φ over $[\xi^{n-1}, x^{n-1}] \times Y$ is a solution. If we choose (ξ^n, η^n) in this way, then the computation of (ξ^n, η^n) may be conceived as taking place in two stages. First we calculate ξ^n by minimizing $M(\cdot)$ over $[\xi^{n-1}, x^{n-1}]$. Then we calculate η^n by maximizing $\varphi(\xi^n, \cdot)$ over Y. Due to (H4), (ξ^n, η^n) so calculated is indeed a saddle point on $[\xi^{n-1}, x^{n-1}] \times Y$. Since $\xi^{n-1} \in Q$ and $M(\xi^n) \leq \varphi(\xi^{n-1}, \eta^n) \leq M(\xi^{n-1})$, it follows that $\xi^n \in Q$, too. Hence $(\xi^n, \eta^n) \in \tilde{S}$, and from assumption 3) follows the existence of x^n satisfying b); moreover $x^n \in K$. The sequence $\{(\xi^n, \eta^n)\}$ has cluster points, since it is contained in the compact set $\tilde{S}$.

Theorem 2. Every cluster point of the sequence $\{(\xi^n, \eta^n)\}$ generated by (28), (29) is a saddle point of φ on $X \times Y$.

Proof.

Let (ξ^*, η^*)be a cluster point of $\{(\xi^n, \eta^n)\}$. Due to the compactness of K there exists a subsequence, indexed by $n(j)$, such that

$$\xi^{n(j)} \rightarrow \xi^*,\ \eta^{n(j)} \rightarrow \eta^*,\ x^{n(j)} \rightarrow x^*,\ \eta^{n(j)+1} \rightarrow \overline{\eta}.$$

From (28) we obtain

$$M(\xi^n) = \varphi(\xi^n, \eta^n) \leq \varphi(\xi^{n-1}, \eta^n) \leq M(\xi^{n-1}).$$

Hence the entire sequences $\{M(\xi^n)\}$, $\{\varphi(\xi^n, \eta^n)\}$, and $\{\varphi(\xi^{n-1}, \eta^n)\}$ are decreasing and are converging towards the same value. Due to continuity we obtain then

$$M(\xi^*) = \varphi(\xi^*, \eta^*) = \varphi(\xi^*, \overline{\eta}), \tag{30}$$

and (H4) implies $\eta^* = \overline{\eta}$. Furthermore, (28) gives $M(\xi^n) \leq \varphi(x^{n-1}, \eta^n)$, and substituting $n := n(j)+1$ we obtain in the limit

$$M(\xi^*) \le \varphi(x^*, \overline{\eta}) = \varphi(x^*, \eta^*);$$

hence, from (30)

$$\varphi(\xi^*, \eta^*) \le \varphi(x^*, \eta^*). \tag{31}$$

From (29) it follows that

$$\varphi(x^n, \eta^n) + F(x^n, \xi^n) \le \varphi(x, \eta^n) + F(x, \xi^n) \quad \forall x \in X.$$

Since $F \ge 0$, this yields in the limit for the subsequence

$$\varphi(x^*, \eta^*) \le \varphi(x, \eta^*) + F(x, \xi^*) \quad \forall x \in X.$$

Using (31) we obtain

$$\varphi(\xi^*, \eta^*) \le \varphi(x, \eta^*) + F(x, \xi^*) \quad \forall x \in X.$$

Then from (18) it follows that

$$\varphi(\xi^*, \eta^*) \le \varphi(x, \eta^*) \quad \forall x \in X,$$

hence $M(\xi^*) \le m(\eta^*)$, and (ξ^*, η^*) is a saddle point. q.e.d.

Rate of convergence. Since the variant (28)–(29) is close to a method of feasible directions, it is natural that we can estimate the rate of convergence by adapting results for the latter class. We borrow from [4]. We specialize algorithm (28)–(29) as follows. $F(x, \xi) := \|x - \xi\|^2$, and (ξ^n, η^n) is chosen as a saddle point of φ on $[\xi^{n-1}, x^{n-1}] \times Y$. So we have

$$\xi^n \text{ solves } \min\{M(\xi) | \xi \in [\xi^{n-1}, x^{n-1}]\},$$
$$\eta^n \text{ satisfies } M(\xi^n) = \varphi(\xi^n, \eta^n),$$
$$x^n \text{ solves } \min\{\varphi(x, \eta^n) + \|x - \xi^n\|^2 | x \in X\}.$$

We assume in addition

1. $\varphi(\cdot, y)$ is convex;
2. there exist constants $0 < v \le V$ such that for all $\xi \in Q$
   a) $M(\xi + h) - M(\xi) \ge \varphi(\xi + h, \eta(\xi)) - \varphi(\xi, \eta(\xi)) + v\|h\|^2 \quad \forall h \in X - \xi,$
   b) $M(\xi + h) - M(\xi) \le \varphi(\xi + h, \eta(\xi)) - \varphi(\xi, \eta(\xi)) + V\|h\|^2 \quad \forall h \in X - \xi,$
   where $\eta(\xi)$ is (uniquely) determined by the requirement $\varphi(\xi, \eta(\xi)) = M(\xi)$.

Theorem 3. With $\Delta_n := M(\xi^n) - \inf_{x \in X} M(x)$ we have $\Delta_{n+1} \le \Delta_n (1 - \frac{\overline{v}}{\overline{V}})$, where $\overline{v} = \min\{1, v\}$, and $\overline{V} = \max\{1, V\}$.

Proof. Set $\tau_n := \varphi(x^n, \eta^n) - \varphi(\xi^n, \eta^n) + \|x^n - \xi^n\|^2$. Then from the definition of x^n

follows

$$\tau_n \leq \varphi(x,\eta^n) - \varphi(\xi^n,\eta^n) + \|x-\xi^n\|^2 \quad \forall x \in X,$$

hence

$$\tau_n \leq \varphi(\xi^n + h,\eta^n) - \varphi(\xi^n,\eta^n) + \|h\|^2 \quad \forall h \in X - \xi^n.$$

Since $\overline{v} \leq 1$, $\tilde{h} \in X - \xi^n$ implies $h := \overline{v}\tilde{h} \in X - \xi^n$. So we obtain, using the convexity of $\varphi(\cdot, y)$, that

$$\begin{aligned}\tau_n &\leq \overline{v} \cdot (\varphi(\xi^n + \tilde{h},\eta^n) - \varphi(\xi^n,\eta^n)) + \overline{v}^2\|\tilde{h}\|^2 \\ &\leq \overline{v} \cdot (\varphi(\xi^n + \tilde{h},\eta^n) - \varphi(\xi^n,\eta^n) + v\|\tilde{h}\|^2) \quad \forall \tilde{h} \in X - \xi^n.\end{aligned}$$

From assumption 2a) it follows that

$$\tau_n \leq \overline{v} \cdot (M(\xi^n + \tilde{h}) - M(\xi^n)) \quad \forall \tilde{h} \in X - \xi^n,$$

hence

$$\tau_n \leq \overline{v} \cdot (-\Delta_n).$$

Furthermore with $\overline{\lambda} := \frac{1}{\overline{V}}$ and $h^n := x^n - \xi^n$ we obtain

$$\begin{aligned}&M(\xi^n + \overline{\lambda}h^n) - M(\xi^n) \\ &\qquad \leq \varphi(\xi^n + \overline{\lambda}h^n,\eta^n) - \varphi(\xi^n,\eta^n) + V \cdot \overline{\lambda}^2 \cdot \|h^n\|^2 \quad \text{[from assumption 2b)]} \\ &\qquad \leq \overline{\lambda}(\varphi(x^n,\eta^n) - \varphi(\xi^n,\eta^n)) + V \cdot \overline{\lambda}^2 \cdot \|h^n\|^2 \quad \text{[from assumption 1)]} \\ &\qquad \leq \overline{\lambda} \cdot (\varphi(x^n,\eta^n) - \varphi(\xi^n,\eta^n) + \|h^n\|^2) \quad [\text{since}\,\overline{\lambda} \leq \frac{1}{V}] \\ &\qquad = \overline{\lambda} \cdot \tau_n.\end{aligned}$$

Since $\xi^n + \overline{\lambda}h^n \in [\xi^n, x^n]$ it follows from the definition of ξ^{n+1} that

$$M(\xi^{n+1}) - M(\xi^n) \leq \overline{\lambda}\tau_n.$$

Hence $\Delta_{n+1} - \Delta_n \leq \overline{\lambda} \cdot \overline{v}(-\Delta_n)$ and $\Delta_{n+1} \leq \Delta_n(1 - \frac{\overline{v}}{\overline{V}})$. q.e.d.

The same rate of convergence, but under somewhat different assumptions, has also been established in [14].

If we require in addition to the assumptions made for algorithm (28)–(29) that $\varphi(\cdot, y)$ is convex and $\varphi(x, \cdot)$ is strictly concave (thus sharpening (H4)), then algorithm (28)–(29) can be modified as follows: Given $\xi^{n-1} \in Q$ and $x^{n-1} \in K$ we define for $(x, y) \in [\xi^{n-1}, x^{n-1}] \times Y$:

$$\psi_n(x,y) := \lambda\varphi(\xi^{n-1},y) + (1-\lambda)\varphi(x^{n-1},y),$$

where $\lambda \in [0,1]$ is determined by $x = \lambda\xi^{n-1} + (1-\lambda)x^{n-1}$. The n-th iteration consists of the following steps:

a) Select $(\xi^n, \eta^n) \in X \times Y$ such that

$$(\xi^n, \eta^n) \text{ is a saddle point of } \psi_n(\xi, \eta) \text{ over } [\xi^{n-1}, x^{n-1}] \times Y.$$

b) Select $x^n \in X$ according to (29).

This is essentially algorithm 2 with the deletion rule from [1]. Again the sequence $\{(\xi^n, \eta^n)\}$ is contained in the compact set $\tilde{S}$. Every cluster point of the sequence $\{(\xi^n, \eta^n)\}$ is a saddle point of φ on $X \times Y$.

Indeed: Let (ξ^*, η^*) be a cluster point of (ξ^n, η^n). Due to the compactness of K there exists a subsequence, indexed by $n(j)$ $(j \in I\!N)$, such that

$$\xi^{n(j)} \to \xi^*, \ \eta^{n(j)} \to \eta^*, \ x^{n(j)} \to x^*, \ \eta^{n(j)+1} \to \overline{\eta}, \ \xi^{n(j)-1} \to \tilde{\xi}, \ x^{n(j)-1} \to \tilde{x}.$$

Then $\xi^* = \lambda^*\tilde{\xi} + (1-\lambda^*)\tilde{x}$ for some $\lambda^* \in [0,1]$. Let $\psi_*(\xi^*, y) := \lambda^*\varphi(\tilde{\xi}, y) + (1-\lambda^*)\varphi(\tilde{x}, y)$. From step a) follows in particular

$$\sup_{y \in Y} \psi_n(\xi^n, y) \le \psi_n(\xi^n, \eta^n) \le \psi_n(\xi^{n-1}, \eta^n) = \varphi(\xi^{n-1}, \eta^n).$$

From the convexity of $\varphi(\cdot, y)$ follows $\varphi(\xi^n, y) \le \psi_n(\xi^n, y)$. So we obtain

$$M(\xi^n) \le \sup_{y \in Y} \psi_n(\xi^n, y) \le \psi_n(\xi^n, \eta^n) \le \varphi(\xi^{n-1}, \eta^n) \le M(\xi^{n-1}).$$

Hence the entire sequences $\{M(\xi^n)\}$, $\{\sup_{y \in Y} \psi_n(\xi^n, y)\}$, $\{\psi_n(\xi^n, \eta^n)\}$ and $\{\varphi(\xi^{n-1}, \eta^n)\}$ are converging towards the same value. From continuity and $\varphi(\xi^*, y) \le \psi_*(\xi^*, y)$ we obtain then

$$M(\xi^*) = \sup_{y \in Y} \psi_*(\xi^*, y) = \psi_*(\xi^*, \eta^*) = \varphi(\xi^*, \overline{\eta}) \le \psi_*(\xi^*, \overline{\eta}).$$

Since $\psi_*(\xi^*, \cdot)$ is strictly concave and therefore satisfies (H4), this implies $\eta^* = \overline{\eta}$. So we obtain

$$M(\xi^*) = \varphi(\xi^*, \eta^*) = \psi_*(\xi^*, \eta^*).$$

Step a) gives furthermore

$$\psi_n(\xi^n, \eta^n) \le \psi_n(x^{n-1}, \eta^n) = \varphi(x^{n-1}, \eta^n).$$

Substituting $n := n(j) + 1$ we obtain in the limit that

$$\psi_*(\xi^*, \eta^*) \le \varphi(x^*, \overline{\eta}) = \varphi(x^*, \eta^*).$$

Hence

$$M(\xi^*) = \varphi(\xi^*, \eta^*) \le \varphi(x^*, \eta^*).$$

Since subprogram (29) remains unchanged the same argument as in the proof of theorem 2 shows then that $M(\xi^*) \le m(\eta^*)$, and (ξ^*, η^*) is a saddle point. q.e.d.

6. Extension to equilibrium problems

The notion of a (Nash-) equilibrium is of fundamental importance in the theory of non-cooperative n-person games. Let there be given a finite family of sets Z_i $(i \in I)$ and a corresponding family of functions $f_i : \prod_{j\in I} Z_j \to \mathbb{R}$ $(i \in I)$. We abbreviate

$Z := \prod_{i\in I} Z_i$, $Z_{\sim i} := \prod_{j\in I, j\neq i} Z_j$ $(i \in I)$, so that $Z = Z_i \times Z_{\sim i}$. Similary for $z =: (z_i)_{i\in I} \in Z$ we abbreviate $z_{\sim i} := (z_j)_{j\in I, j\neq i} \in Z_{\sim i}$, so that $z = (z_i, z_{\sim i})$.

A point $\varsigma^* \in Z$ with $\varsigma^* = (\varsigma_i^*)_{i\in I}$ is called an equilibrium point of the system of functions f_i iff for all $i \in I$

$$f_i(\varsigma_i^*, \varsigma_{\sim i}^*) \leq f_i(z_i, \varsigma_{\sim i}^*) \quad \forall z_i \in Z_i .$$

Let us assume that for all $i \in I$

1. the sets Z_i are nonempty, convex, compact, (32)
2. the functions $f_i(\cdot)$ are continuous on Z , (33)
3. the functions $f_i(\cdot, z_{\sim i})$ are quasiconvex on Z_i for each fixed $z_{\sim i} \in Z_{\sim i}$. (34)

Then there exists at least one equilibrium point – see [10, 11], and the remark below.

We define the function $\Phi : Z \times Z \to \mathbb{R}$ by means of

$$\Phi(z, \varsigma) := \sum_{i\in I} (f_i(z_i, \varsigma_{\sim i}) - f_i(\varsigma_i, \varsigma_{\sim i})),$$

where $z = (z_i)_{i\in I} \in Z$ and $\varsigma = (\varsigma_i)_{i\in I} \in Z$. Then $\Phi(\varsigma, \varsigma) = 0$ for all $\varsigma \in Z$. It can easily be seen that $\varsigma^* \in Z$ is an equilibrium point if and only if $\Phi(z, \varsigma^*) \geq 0$ for all $z \in Z$. This is again problem (8), and we can apply the general decomposition scheme (9)–(10) given above. A simple realization with $H = 0$ is as follows:

At the beginning of the n-th iteration we are given finite subsets $Z_i^n \subset Z_i$ $(i \in I)$.

We determine $\varsigma^n \in Z$ with $\varsigma^n = (\varsigma_i^n)_{i\in I}$ such that for all $i \in I$

$$f_i(\varsigma_i^n, \varsigma_{\sim i}^n) \leq f_i(z_i, \varsigma_{\sim i}^n) \quad \forall z_i \in Z_i^n . \tag{35}$$

We determine $z^n \in Z$ with $z^n = (z_i^n)_{i\in I}$ such that for all $i \in I$

$$f_i(z_i^n, \varsigma_{\sim i}^n) \leq f_i(z_i, \varsigma_{\sim i}^n) \quad \forall z_i \in Z_i .$$

We set $Z_i^{n+1} := Z_i^n \cup \{z_i^n\}$, and start the next iteration.

Recall that (35) is solvable because of the existence of an equilibrium point on $\prod_{i\in I} conv\ Z_i^n$. Every limit point of the sequence $\{\varsigma^n\}$ is an equilibrium point.

Remark. For algorithmic reasons we needed the theorem of Sion [15, 9] only in the situation where the functions occuring are continuous (whereas the original formulation of this theorem needs only appropriate semicontinuity requirements). Under the stronger assumption of continuity Sion's theorem as well as Nash's result [10, 11] follow readily from Fan's fixed point theorem. Indeed, to obtain Nash's result assume that (32), (33), (34) are satisfied. Define multivalued mappings $A_i : Z \rightrightarrows Z_i \quad (i \in I)$ by
$A_i(z) := \{\varsigma_i^* \in Z_i | f_i(\varsigma_i^*, z_{\sim i}) \leq f_i(\varsigma_i, z_{\sim i}) \quad \forall \varsigma_i \in Z_i\}$. Let

$$A(z) := \prod_{i \in I} A_i(z) : Z \rightrightarrows Z.$$

Then $A(z)$ is convex, compact and nonempty for all $z \in Z$, and by the result of [2, p.123] $A(\cdot)$ is upper semicontinuous. Hence by Fan's fixed point theorem [8] A has a fixed point $\varsigma^* \in A(\varsigma^*)$. With $\varsigma^* = (\varsigma_i^*)_{i \in I}$ this means that $\varsigma_i^* \in Z_i$ minimizes $f_i(\cdot, \varsigma_{\sim i}^*)$ over Z_i, hence ς^* is an equilibrium: Nash's result. Sion's result becomes a special case of Nash's result: choose $\varphi = f_1 = -f_2$ in the latter to obtain

$$\varphi(\xi^*, y) \leq \varphi(\xi^*, \eta^*) \leq \varphi(x, \eta^*) \quad \forall x \in X, \ \forall y \in Y,$$

which is Sion's result.

References

[1] A.Auslender: Two General Methods for Computing Saddle Points with Application for Decomposing Convex Programming Problems. Applied Mathematics and Optimization 13 (1985), 79-95.

[2] C.Berge: Espaces topologiques: fonctions multivoques. 2.éd. Paris, Dunod, 1966.

[3] E.Blum, W.Oettli: Mathematische Optimierung: Grundlagen und Verfahren. Berlin, Springer, 1975.

[4] E.Blum, W.Oettli: The Principle of Feasible Directions for Nonlinear Approximants and Infinitely Many Constraints. Symposia Mathematica 19 (1976), 91-101.

[5] P.Broise, P.Huard, J.Sentenac: Décomposition des programmes mathématiques. Paris, Dunod, 1968.

[6] G.Cohen: Auxiliary Problem Principle and Decomposition of Optimization Problems. Journal of Optimization Theory and Applications 32 (1980), 277-305.

[7] G.B.Dantzig: Linear Programming and Extensions. Princeton, Princeton University Press, 1963.

[8] K.Fan: A Minimax Inequality and Applications. Proceedings of the Third Symposium on Inequalities. Ed. by O.Shisha. New York, Academic Press, 1972, 103-113.

[9] K.Fan: Sur un théorème minimax. Comptes Rendus de l' Académie des Sciences de Paris, Série A 259 (1964), 3925-3928.

[10] K.Fan: Applications of a Theorem Concerning Sets with Convex Sections. Mathematische Annalen 163 (1966), 189-203.

[11] J.Nash: Non-Cooperative Games. Annals of Mathematics 54 (1951), 286-295.

[12] W.Oettli: Eine allgemeine, symmetrische Formulierung des Dekompositionsprinzips für duale Paare nichtlinearer Minmax- und Maxmin-Probleme. Zeitschrift für Operations Research 18 (1974), 1-18.

[13] R.T.Rockafellar: Convex Analysis. Princeton, Princeton University Press, 1970.

[14] R.T.Rockafellar, R.J.-B.Wets: A Lagrangian Finite Generation Technique for Solving Linear-Quadratic Problems in Stochastic Programming. Mathematical Programming Studies 28 (1986), 63-93.

[15] M.Sion: On General Minimax Theorems. Pacific Journal of Mathematics 8 (1958), 171-176.

[16] W.I.Zangwill: Nonlinear Programming: A Unified Approach. Englewood Cliffs, Prentice Hall, 1969.