PACIFIC JOURNAL OF MATHEMATICS
Vol. 102, No. 2, 1982

ON TWO-STAGE MINIMAX PROBLEMS

JOACHIM HARTUNG

Minimax problems are considered whose admissable sets are given implicitly as the solution sets of another minimax problem. For the solution a parametric method is proposed. Special cases of it are extensions of Courant's exterior penalty method and Tihonov's regularization method of Nonlinear Programming to minimax problems.

In solving quadratic problems explicitly, a representation of modified best approximate solutions of linear equations in Hilbert spaces is given that extends results for the usual case.

1. Introduction. Let X and Y be not empty subsets of real linear topological Hausdorff spaces $\mathscr{X}$ and $\mathscr{Y}$, respectively,

$$f: X \times Y \longrightarrow \boldsymbol{R}, \text{ and } g: X \times Y \longrightarrow \boldsymbol{R}$$

be two real valued functions on $X \times Y$, and denote $X_f \times Y_f$ the solution set of the minimax problem (X, Y, f), i.e.,

$$(x_0, y_0) \in X_f \times Y_f: \Longleftrightarrow \bigwedge_{x \in X} \bigwedge_{y \in Y} f(x, y_0) \leqq f(x_0, y_0) \leqq f(x_0, y) .$$

Note that if (x_1, y_1) and (x_2, y_2) are in $X_f \times Y_f$ then also $(x_1, y_2) \in X_f \times Y_f$, being thus a product set.

Under the assumption that X_f and Y_f are not empty, we give the following

DEFINITION 1. A two-stage minimax problem, in the notation $\mathscr{M}_{g/f}$, is the minimax problem

$$\mathscr{M}_{g/f}: = (X_f, Y_f, g/X_f \times Y_f) .$$

Considering $\mathscr{M}_{g/f}$ as a two-person zero-sum game, it describes the following conflict situation: Two antagonists choose independently from each other $x \in X$, resp. $y \in Y$, and the first one gets from the second one the vector-payoff $(f(x, y), g(x, y)) \in \boldsymbol{R}^2$. The preference relation may be induced by the lexicographic order of $\boldsymbol{R}^2$:

(x_1, y_1) is better than (x_2, y_2) for the first (second) player, if $(f(x_1, y_1), g(x_1, y_1))$ is lexicographically greater (smaller) than $(f(x_2, y_2), g(x_2, y_2))$. If the players are cautious, they have to take as optimal strategies the components of a solution of $\mathscr{M}_{g/f}$, provided there exists one.

Many games are of this nature; for example (see §§ 3, 4 and 5 below) constrained games, where on the first stage the constraints have to be satisfied, or games, in which you are interested in optimal strategies of minimum (semi-) norm, like for instance in certain differential games, where the (semi-) norm represents the consumption of energy, which of course should be minimal among all optimal strategies.

A method for solving $\mathscr{M}_{g/f}$ that first produces the whole sets X_f and Y_f, meets with great numerical difficulties. Therefore the following *algorithm* is of interest that solves $\mathscr{M}_{g/f}$ without computing X_f and Y_f: Take an arbitrary real positive nullsequence $\{r_n\}_{n\in N}\subset \boldsymbol{R}$ and find a solution (x_n, y_n) of the problem $(X, Y, f + r_n g)$, $(n \in \boldsymbol{N})$. Under certain conditions the accumulation points of $\{x_n\}_{n\in N}$, $\{y_n\}_{n\in N}$ (unique in some cases) build a solution of $\mathscr{M}_{g/f}$, as is shown below.

2. A solution algorithm for the general problem $\mathscr{M}_{g/f}$.

DEFINITION 2.

(a) A function $f: X \to \boldsymbol{R}$ is called

(i) *inf-compact*, if $\{x \mid x \in X,\ f(x) \leqq c\}$, $c \in \boldsymbol{R}$, is compact.

(ii) *sup-compact*, if $(-f)$ is inf-compact.

(b) A function h: $X \times Y \to \boldsymbol{R}$ is called (x_1, y_1)-*supinf-compact*, for a fixed $(x_1, y_1) \in X \times Y$, if $h(x_1, \cdot)$ is inf-compact and $h(\cdot, y_1)$ is sup-compact.

We say that a real function $h(x, y)$ on $X \times Y$ is u.s.c.-l.s.c., if $h(x, y)$ is upper semi-continuous in x for each $y \in Y$ and lower semi-continuous in y for each $x \in X$.

For a real positive sequence

$$\{r_n\}_{n\in N} \subset \boldsymbol{R},\ \text{with}\ r_n \longrightarrow +0\ \text{for}\ n \longrightarrow \infty\ ,$$

let p_n be defined by

$$p_n\colon \begin{matrix} X \times Y \longrightarrow \boldsymbol{R} \\ (x, y) \longmapsto f(x, y) + r_n g(x, y) \end{matrix}\ ,\quad (n \in \boldsymbol{N})\ .$$

THEOREM 1. *Under the conditions*

(i) *X and Y are convex and closed.*

(ii) *f and g are u.s.c.-l.s.c., and g is bounded above in x for each $y \in Y$ and bounded below in y for each $x \in X$.*

(iii) *There exists a (fixed) $(x_0, y_0) \in X_f \times Y_f$ such that g is (x_0, y_0)-supinf-compact.*

(iv) *p_n is quasi-concave-convex, $(n \in \boldsymbol{N})$.*

we have

(v) *(X, Y, p_n) has a solution (x_n, y_n), $(n \in \boldsymbol{N})$.*

(vi) $\{x_n\}_{n\in N}$ *and* $\{y_n\}_{n\in N}$ *have cluster points* $\hat{x}$ *and* $\hat{y}$*, respectively, and each* $(\hat{x}, \hat{y})$ *solves* $\mathscr{M}_{g/f}$.

(vii) $\lim_{n\to\infty} p_n(x_n, y_n) = f(\hat{x}, \hat{y})$.

(viii) $\lim_{n\to\infty} (p_n(x_n, y_n) - f(x_0, y_0))/r_n = g(\hat{x}, \hat{y})$.

Proof. The sum of two u.s.c., (l.s.c.), functions on a closed set is u.s.c., (l.s.c.), and so by (ii) $p_n = f + r_n g$, $(n \in \boldsymbol{N})$, is u.s.c-l.s.c.

For $n \in \boldsymbol{N}$ and $c \in \boldsymbol{R}$ we have

$$\begin{aligned}&\{y \mid y \in Y,\ p_n(x_0, y) \leqq c\}\\ &\quad\subset \{y \mid y \in Y,\ r_n g(x_0, y) \leqq c - \inf_{y\in Y} f(x_0, y)\}\\ &\quad\subset \left\{y \;\middle|\; y \in Y,\ g(x_0, y) \leqq \frac{1}{r_n}(c - f(x_0, y_0))\right\},\end{aligned}$$

the last set is compact by (iii), and so $p_n(x_0, \cdot)$ is inf-compact. Similarly, $p_n(\cdot, y_0)$ is sup-compact. Applying now Theorem 1 of Hartung [5], we get the existence of a saddle point (x_n, y_n) of p_n over $X \times Y$, $(n \in \boldsymbol{N})$. For all $x_f \in X_f$ and $y_f \in Y_f$ we then get, with $n \in \boldsymbol{N}$,

$$\begin{aligned}[f(x_f, y_n) + r_n g(x_f, y_n)] - f(x_f, y_n) &\leqq p_n(x_n, y_n) - f(x_f, y_f)\\ &\leqq [f(x_n, y_f) + r_n g(x_n, y_f)] - f(x_n, y_f),\end{aligned} \tag{1}$$

or

$$r_n g(x_f, y_n) \leqq p_n(x_n, y_n) - f(x_f, y_f) \leqq r_n g(x_n, y_f). \tag{2}$$

Putting $x_f = x_0$, $y_f = y_0$, (2) gives because of (ii)

$$-\infty < r_n \inf_{y\in Y} g(x_0, y) \leqq p_n(x_n, y_n) - f(x_0, y_0) \leqq r_n \sup_{x\in X} g(x, y_0) < +\infty, \tag{3}$$

and so

$$p_n(x_n, y_n) \longrightarrow f(x_0, y_0),\ \text{as } r_n \longrightarrow +0 \text{ for } n \longrightarrow \infty. \tag{4}$$

Dividing in (2) by r_n, we get

$$g(x_0, y_n) \leqq \sup_{x\in X} g(x, y_0),\ \inf_{y\in Y} g(x_0, y) \leqq g(x_n, y_0), \tag{5}$$

which by (iii) means that x_n, y_n are elements of compact sets independent of n. Therefore $\{x_n\}_{n\in N}$, $\{y_n\}_{n\in N}$ have cluster points $\hat{x} \in X$, $\hat{y} \in Y$. Let $\{x_{n_k}\}$ be a subnet of $\{x_n\}_{n\in N}$ converging to $\hat{x}$. By (ii) and (4) it follows that

$$\begin{aligned}f(\hat{x}, y) &\geqq \limsup_{x_{n_k}\to\hat{x}} f(x_{n_k}, y)\\ &\geqq \limsup\, (p_{n_k}(x_{n_k}, y_{n_k}) - r_{n_k} g(x_{n_k}, y))\\ &\geqq \limsup\, (p_{n_k}(x_{n_k}, y_{n_k}) - r_{n_k} \sup_{x\in X} g(x, y))\\ &\geqq f(x_0, y_0),\ \text{for all } y \in Y,\ \text{i.e., } \hat{x} \in X_f,\end{aligned} \tag{6}$$

and analogously, $\hat{y} \in Y_f$. Let now $\tilde{y}$ be a cluster point of the subnet $\{y_{n_k}\}$ of $\{y_n\}_{n \in N}$, existing by (5), and $\{y_{n_{k_i}}\}$ a subnet of it converging to $\tilde{y}$. Then of course $x_{n_{k_i}} \to \hat{x}$, and

$$(\hat{x}, \tilde{y}) \in X_f \times Y_f \ . \tag{7}$$

From (2) we get, since $f(x_f, y_f) = \text{const} = f(x_0, y_0)$ for $(x_f, y_f) \in X_f \times Y_f$,

$$\sup_{x \in X_f} g(x, y_n) \leqq \frac{p_n(x_n, y_n) - f(x_0, y_0)}{r_n} \leqq \inf_{y \in Y_f} g(x_n, y) \ . \tag{8}$$

The functions $x \mapsto \inf_{y \in Y_f} g(x, y)$ and $y \mapsto \sup_{x \in X_f} g(x, y)$ are u.s.c., resp. l.s.c., and thus (8) yields

$$\begin{aligned} \sup_{x \in X_f} g(x, \tilde{y}) &\leqq \liminf_{y_{n_{k_i}} \to \tilde{y}} \sup_{x \in X_f} g(x, y_{n_{k_i}}) \\ &\leqq \limsup_{x_{n_{k_i}} \to \hat{x}} \inf_{y \in Y_f} g(x_{n_{k_i}}, y) \\ &\leqq \inf_{y \in Y_f} g(\hat{x}, y) \ , \end{aligned} \tag{9}$$

which gives

$$g(\hat{x}, \tilde{y}) \leqq \sup_{x \in X_f} g(x, \tilde{y}) \leqq \inf_{y \in Y_f} g(\hat{x}, y) \leqq g(\hat{x}, \tilde{y}) \ , \tag{10}$$

i.e., $(\hat{x}, \tilde{y})$ is a saddle point of $g/X_f \times Y_f$. Similarly, $\hat{y}$ is a saddle point component of $g/X_f \times Y_f$, and so (vi) is shown. The statement (vii) now follows from (4). Let

$$\begin{aligned} b_n &:= \sup_{x \in X_f} g(x, y_n) \ , \qquad c_n := \inf_{y \in Y_f} g(x_n, y) \ , \\ b &:= \liminf_{n \to \infty} b_n \ , \quad \text{and} \quad c := \limsup_{n \to \infty} c_n \ , \end{aligned}$$

and $\{b_{n_s}\}_{s \in N}$, $\{c_{n_t}\}_{t \in N}$ be sequences converging to b and c, respectively. The corresponding y_{n_s} and x_{n_t} are contained in compact sets by (5), and thus there exist subnets $\{y_{n_{s_i}}\}$ and $\{x_{n_{t_j}}\}$ converging resp. to a $y^* \in Y_f$ and an $x^* \in X_f$. Then of course $b_{n_{s_i}}$ is converging to b and $c_{n_{t_j}}$ to c, and we get from (8)

$$\begin{aligned} \sup_{x \in X_f} g(x, y^*) &\leqq \liminf_{y_{n_{s_i}} \to y^*} \sup_{x \in X_f} g(x, y_{n_{s_i}}) \\ &\leqq \liminf_{n \to \infty} \sup_{x \in X_f} g(x, y_n) \\ &\leqq \liminf_{n \to \infty} \frac{p_n(x_n, y_n) - f(x_0, y_0)}{r_n} \\ &\leqq \limsup_{n \to \infty} \frac{p_n(x_n, y_n) - f(x_0, y_0)}{r_n} \end{aligned} \tag{11}$$

$$\leqq \limsup_{n\to\infty} \inf_{y\in Y_f} g(x_n, y)$$
$$\leqq \limsup_{x_{n_{i_j}}\to x^*} \inf_{y\in Y_f} g(x_{n_{i_j}}, y) \leqq \inf_{y\in Y_f} g(x^*, y) ,$$

which gives (viii).

COROLLARY 1. *If we have for some* $(x_i, y_i) \in X_f \times Y_f$, $(i = 1, 2)$, *and for* $c \in \boldsymbol{R}^2$ *that the level sets*

$$\{x \mid x \in X,\ f(x, y_1) \geqq c_1,\ g(x, y_2) \geqq c_2\} ,$$
$$\{y \mid y \in Y,\ f(x_1, y) \leqq c_1,\ g(x_2, y) \leqq c_2\}$$

are compact and g *satisfies the boundedness condition of* (ii), *we can take instead of* g *the function*

$$\tilde{g}(x, y) := f(x, y_1) + f(x_1, y) + g(x, y) ,$$

which is (x_2, y_2)-sup inf-*compact, and*

$$\tilde{g}/X_f \times Y_f = g/X_f \times Y_f + \text{const} .$$

Proof. We show that $\tilde{g}(\cdot, y_2)$ is sup-compact. For $c \in \boldsymbol{R}$ and $x \in X$ we have:

$$\tilde{g}(x, y_2) \geqq c \Longrightarrow (g(x, y_2) \geqq c - f(x_1, y_2) - \max_{x\in X} f(x, y_1) ,$$
$$\text{and}$$
$$f(x, y_1) \geqq c - f(x_1, y_2) - \sup_{x\in X} g(x, y_2)) .$$

DEFINITION 3. Let U be a convex subset of a real normed linear space, then a function $h: U \to \boldsymbol{R}$ is called *uniformly quasi-convex*, if there exists a continuous isotonic function $\delta: [0, \infty) \to [0, \infty)$ with $\delta(0) = 0$, $\delta(t) > 0$ for $t > 0$, such that for all $u_1, u_2 \in U$

$$h\Big(\frac{1}{2}(u_1 + u_2)\Big) \leqq \max\{h(u_1), h(u_2)\} - \delta(\|u_1 - u_2\|) .$$

Similarly, h is *uniformly quasi-concave*, if $(-h)$ is uniformly quasi-convex.

THEOREM 2. *If in addition to* (i), (ii), (iv) *of Theorem* 1, $\mathscr{X}$ *and* $\mathscr{Y}$ *are reflexive Banach spaces,* X_f *and* Y_f *are not empty, and* g *is uniformly quasi-concave-convex, then*

> (X, Y, p_n) *has a solution* (x_n, y_n), $(n \in \boldsymbol{N})$, $\{x_n\}_{n\in \boldsymbol{N}}$ *and* $\{y_n\}_{n\in \boldsymbol{N}}$ *converge (strongly) to an* $\hat{x} \in X$ *and a* $\hat{y} \in Y$, *resp., and* $(\hat{x}, \hat{y})$ *is the solution of* $\mathscr{M}_{g/f}$.

Proof. Let $x_f \in X_f$ be fixed, then by Definition 3 there exists a continuous isotonic function $\delta_{x_f}: [0, \infty) \to [0, \infty)$ with $\delta_{x_f}(t) = 0 \Leftrightarrow t = 0$, such that for all $y \in Y$ and $y_f \in Y_f$

$$\delta_{x_f}(\| y - y_f \|) \leqq \max \{g(x_f, y), g(x_f, y_f)\} - g\Big(x_f, \frac{1}{2}(y + y_f)\Big) . \tag{12}$$

For $c \in \boldsymbol{R}$ we have

$$g(x_f, y) \leqq c \Longrightarrow \| y - y_f \| \leqq \delta_{x_f}^{-1}\Big(\max \{c, g(x_f, y_f)\} - \inf_{y \in Y} g\Big(x_f, \frac{1}{2}\Big(y + y_f\Big)\Big)\Big) ,$$

and so the level set

$$T_{x_f}^c := \{y \mid y \in Y,\ g(x_f, y) \leqq c\} \text{ is bounded} .$$

$g(x_f, \cdot)$ is l.s.c. and quasi-convex, and thus $T_{x_f}^a$ is convex and closed, hence weakly compact, and so $g(x_f, \cdot)$ is weakly inf-compact, for all $x_f \in X_f$. Similarly, $g(\cdot, y)$ is weakly sup-compact, for all $y \in Y_f$. Herewith all conditions of Theorem 1 are fulfilled in the weak topology, and we get the existence of a solution (x_n, y_n) of (X, Y, p_n), $(n \in \boldsymbol{N})$. Since g is uniformly quasi-concave-convex, there exists a unique solution $(\hat{x}, \hat{y})$ of $\mathscr{M}_{g/f}$, and so the whole sequences $\{x_n\}_{n \in \boldsymbol{N}}$, $\{y_n\}_{n \in \boldsymbol{N}}$ are converging weakly to $\hat{x}$ and $\hat{y}$, respectively.

Putting in (12) $x_f = \hat{x}$, $y = y_n$ and $y_f = \hat{y}$, we get with (8)

$$\delta_{\hat{x}}(\| y_n - \hat{y} \|) \leqq \max \Big\{\frac{p_n(x_n, y_n) - f(x_f, y_f)}{r_n},\ g(\hat{x}, \hat{y})\Big\} - g\Big(\hat{x}, \frac{1}{2}(y_n + \hat{y})\Big) . \tag{13}$$

$1/2(y_n + \hat{y}) \to \hat{y}$, for $n \to \infty$, $g(x, \cdot)$ is weakly l.s.c., and so (13) yields by using (viii) of Theorem 1

$$\limsup_{n \to \infty} \delta_{\hat{x}}(\| y_n - \hat{y} \|) \leqq g(\hat{x}, \hat{y}) - g(\hat{x}, \hat{y}) , \tag{14}$$

which gives the strong convergence of $\{y_n\}_{n \in \boldsymbol{N}}$ to $\hat{y}$. Analogously the strong convergence of $\{x_n\}_{n \in \boldsymbol{N}}$ to $\hat{x}$ follows.

3. The exterior penalty method for constrained minimax problems. Let A and B be subsets of X and Y, resp., then we consider the constrained minimax problem

$$(A, B, g) .$$

In [5] we give for this problem an interior penalty method, which works only if A and B have interior points, but if this is the case, it needs for convergence some sup inf-compactness of g only over the sets A and B, which especially is given, if A and B are compact.

If A and B have no interior points, we propose a sequential method approximating a solution of (A, B, g) from the exterior in X and Y of the admissable sets, which is profitable, if the boundaries of X and Y are numerically less complicated than the boundaries of A and B, which is especially the case, when X and Y are the whole spaces.

The *penalty functions*

$$P_A: X \longrightarrow \boldsymbol{R},\ P_B: Y \longrightarrow \boldsymbol{R}$$

are assumed to have the properties

$$P_A(x) = \begin{cases} 0 & \text{for } x \in A \\ > 0 & x \in X \backslash A \end{cases},\quad P_B(y) = \begin{cases} 0 & \text{for } y \in B \\ > 0 & y \in Y \backslash B \end{cases}.$$

Putting

$$f: = P_B - P_A\ ,$$

we get

$$X_f = A,\ Y_f = B,\ f/X_f \times Y_f = 0\ ,$$

and

$$p_n = P_B - P_A + r_n g,\ \text{with } r_n \longrightarrow +0,\ \text{for } n \longrightarrow \infty,\ (n \in \boldsymbol{N})\ .$$

THEOREM 3. *If A and B are convex and closed, and the conditions* (i), (ii), (iii), (iv) *of Theorem* 1 *are fulfilled, then (X, Y, p_n) has a solution (x_n, y_n), $(n \in \boldsymbol{N})$, $\{x_n\}_{n \in \boldsymbol{N}}$, $\{y_n\}_{n \in \boldsymbol{N}}$ have cluster points $\hat{x}$, $\hat{y}$, resp., solving (A, B, g),*

$$\lim_{n\to\infty} P_A(x_n) = 0,\ \lim_{n\to\infty} P_B(y_n) = 0\ ,$$

and

$$\lim_{n\to\infty} g(x_n, y_n) + \frac{1}{r_n}(P_B(y_n) - P_A(x_n)) = g(\hat{x}, \hat{y})\ .$$

Proof. By Theorem 1 we get the existence of a solution (x_n, y_n) of (X, Y, p_n), $(n \in \boldsymbol{N})$, and for $x \in A$, $y \in B$

$$r_n g(x, y_n) + P_B(y_n) \leqq p_n(x_n, y_n) \leqq r_n g(x_n, y) - P_A(x_n)\ ,$$

or

$$\begin{aligned} -\infty < r_n \inf_{y \in Y} g(x, y) + P_B(y_n) &\leqq p_n(x_n, y_n) \\ &\leqq r_n \sup_{x \in X} g(x, y) - P_A(x_n) < +\infty\ , \end{aligned}$$

which yields with (4)

$$0 \leqq \limsup_{n\to\infty} P_B(y_n) \leqq \lim_{n\to\infty} p_n(x_n, y_n) = 0 \leqq \liminf_{n\to\infty} (-P_A(x_n)) \leqq -\limsup_{n\to\infty} P_A(x_n) \leqq 0 .$$

Since $P_B \geqq 0$, $P_A \geqq 0$, that gives

$$\lim_{n\to\infty} P_A(x_n) = 0, \ \lim_{n\to\infty} P_B(y_n) = 0 .$$

The remaining assertions follow from Theorem 1.

Corollary 1 and Theorem 2 then give a refined method. If for example A is given by

$$A = \{x \mid x \in X, \ G_i(x) = 0, \ (i = 1, \cdots, m_1), \ G_j(x) \leqq 0, \ (j = m_1 + 1, \cdots, m)\}$$

for some real valued functions G_i on X, $(i = 1, \cdots, m)$, we can take as a penalty function for instance

$$P_A(x) := \sum_{i=1}^{m_1} (G_i(x))^2 + \sum_{i=m_1+1}^{m} \max\,[0, G_i(x)]^2 ,$$

which is differentiable, when the G_i are.

4. A regularization algorithm for finding saddle points. To solve a minimax problem (X, Y, f) you often have to take algorithms which need for convergency the solution to be unique, as for example the Arrow-Hurwicz-Uzawa gradient methods [1] (like the Lagrangeian method for convex programming) or the successive approximation method of Dem'janov [3]. Therefore, if this is not the case, we approximate f by a sequence of regularized functions, which have this missing property. Theorem 2 offers many possibilities for doing this. In the method we choose, the unique saddle points of the sequential functions are converging to the saddle point of f with minimum norm, which is of particular interest in certain problems. We don't need compactness conditions and thus f can be a Lagrange function of an ordinary convex program. Let $\mathscr{X}$ and $\mathscr{Y}$ be real Hilbert spaces, $\langle \cdot, \cdot \rangle$ denoting the inner product define the norm, $\|\cdot\| := \langle \cdot, \cdot \rangle^{1/2}$, resp., and $\mathscr{X} \times \mathscr{Y}$ may be provided with the induced norm.

Then we define for a real positive nullsequence $\{r_n\}_{n \in N}$ the regularized functionals

$$p_n(x, y) := f(x, y) + r_n(\langle y, y \rangle - \langle x, x \rangle), \ (n \in N) .$$

THEOREM 4. *Let X and Y be convex and closed, (X, Y, f) solv-*

able, and f be u.s.c.-l.s.c. and concave-convex, then

> (X, Y, p_n) *has a unique solution* (x_n, y_n), $(n \in N)$, $\hat{x} := \lim_{n\to\infty} x_n$ *and* $\hat{y} := \lim_{n\to\infty} y_n$ *exist, and* $(\hat{x}, \hat{y})$ *is the solution of* (X, Y, f) *with minimum norm.*

Proof. By the parallelogram law the function

$$g(x, y) := \langle y, y\rangle - \langle x, x\rangle$$

is strictly concave-convex and uniformly quasi-concave-convex. Then $p_n(x, y)$ has these properties, too, and the saddle points of p_n are uniquely determined. The rest of the assertions follow from Theorem 2.

5. An explicit solution of quadratic minimax problems. Let $\mathscr{X}$ and $\mathscr{Y}$ be real Hilbert spaces as in §4, and $X = \mathscr{X}$, $Y = \mathscr{Y}$. Then we consider the quadratic functionals

$$F(x, y) := \langle x, Px\rangle - 2\langle x, c\rangle + 2\langle x, Ly\rangle + \langle y, Qy\rangle - 2\langle d, y\rangle ,$$
$$G(x, y) := \langle x, Sx\rangle + \langle y, Ty\rangle ,$$

where $c \in X$, $d \in Y$; P and S are self-adjoint negative semidefinite linear operators on X, Q and T are self-adjoint positive semidefinite linear operators on Y, L is a linear operator of Y into X and all operators are bounded, and the two stage minimax problem

$$(1) \qquad \mathscr{M}_{G/F} = \mathscr{M}_{G/F}(c, d) .$$

$\langle x, -Sx\rangle$ and $\langle y, Ty\rangle$ are seminorms to the power two, representing for instance in differential games often the consumption of energy, which should be minimal among the optimal strategies of (X, Y, F).

Defining now a linear and bounded operator $\begin{pmatrix} P & L \\ L^* & Q \end{pmatrix} =: A$ by

$$A\colon \begin{array}{l} X \times Y \longrightarrow X \times Y \\ (x, y) \longmapsto (Px + Ly, L^*x + Qy), \ (L^* \text{ denotes the adjoint}) , \end{array}$$

we assume that $(c, d) \in R(A)$, and (as it can be seen by putting the derivatives of $F(x, y)$ with respect to x and y equal to zero) this is a necessary and sufficient condition for the solution set of (X, Y, F) to be not empty, which then is given by

$$X_F \times Y_F = \{(x, y) \mid (x, y) \in X \times Y,\ A(x, y) = (c, d)\} .$$

Let A be normally solvable ($R(A)$ is closed), then the element of $X_F \times Y_F$ with minimum norm is

$$(x', y') := A^{+}(c, d)$$

where A^{+} denotes the pseudoinverse (e.g., Holmes [6], p. 220). Note that $A^{+}w = A^{+}\,\mathrm{Proj}_{R(A)}w$, for $w \in X \times Y$, and $R(A^{+}) \perp N(A)$. With

$$p_n(x, y) := F(x, y) + r_n G(x, y),\ r_n \in \boldsymbol{R},\ r_n \longrightarrow +0,\ \text{for } n \longrightarrow \infty\ , \quad (n \in \boldsymbol{N})\ ,$$

$$B := \begin{pmatrix} S & 0 \\ 0 & T \end{pmatrix}, \quad \text{and} \quad A_n := A + r_n B$$

the solution set of (X, Y, p_n) is

$$\{(x, y) \in X \times Y \mid A_n(x, y) = (c, d)\},\ (n \in \boldsymbol{N})\ .$$

If S and T are definite and normally solvable, then $\langle y, Ty\rangle^{1/2}$ and $\langle x, -Sx\rangle^{1/2}$ are representing norms equivalent to the given ones on Y and X, respectively. So by Theorem 4 (X, Y, p_n) has a unique solution

$$(x_n, y_n) = A_n^{-1}(c, d)\ ,$$

and

$$A^{+(S,T)}(c, d) := \lim_{n\to\infty} A_n^{-1}(c, d) \tag{2}$$

exists and is the solution of $\mathscr{M}_{G/F}$. Since (2) holds for all $(c, d) \in R(A)$, we have

$$A_n^{-1} \longrightarrow A^{+(S,T)}\ (\text{strongly}),\quad \text{as} \quad n \longrightarrow \infty\ , \tag{3}$$

where $A^{+(S,T)}$, the *solution operator* of $\mathscr{M}_{G/F}$, is a linear and bounded operator, because of Banach's inverse mapping theorem.

If I denotes the identity on the spaces, resp., then $A^{+(I,I)} = A^{+}$.

If S and T are not invertible, then (X, Y, p_n) and $\mathscr{M}_{G/F}$ are not uniquely solvable, in general. Then we are interested in the solutions of minimum norm.

The solution set $X_F \times Y_F$ of $A(x, y) = (c, d)$ is given by

$$A^{+}(c, d) + N(A), \quad \text{with} \quad A^{+}(c, d) \perp N(A)\ .$$

Now if $(x, y) \in N(A)$, then

$$\langle x, Px\rangle + \langle x, Ly\rangle = 0$$
$$\langle y, Qy\rangle + \langle x, Ly\rangle = 0\ ,$$

$$\langle x, Px\rangle \leqq 0 \Longrightarrow \langle x, Ly\rangle \geqq 0,\ \langle y, Qy\rangle \geqq 0 \Longrightarrow \langle x, Ly\rangle \leqq 0\ ,$$

and so $\langle x, Ly\rangle = 0$ and $x \in N(P)$, $y \in N(Q)$. Thus

$$X_F \times Y_F = (x, y) + N(P) \times N(Q),\ \text{for any } (x, y) \in X_F \times Y_F\ , \tag{4}$$

and

(5) $$P/X_F = \text{const} \quad \text{and} \quad Q/Y_F = \text{const} .$$

Let $\tilde{P}$ be a self-adjoint negative semidefinite bounded linear operator on X with

(6) $$N(\tilde{P}) \cap N(S) = N(P) \cap N(S),\ \tilde{P}/X_F = \text{const};\ (\text{e.g., } \tilde{P} = P) ,$$

and $\tilde{Q}$ be a self-adjoint positive semidefinite bounded linear operator on Y with

(7) $$N(\tilde{Q}) \cap N(S) = N(Q) \cap N(S),\ \tilde{Q}/Y_F = \text{const};\ (\text{e.g., } \tilde{Q} = Q) .$$

Putting

$$\tilde{S}: = \tilde{P} + S,\ \tilde{T}: = \tilde{Q} + T ,$$

we have

(8) $$N(\tilde{S}) = N(P) \cap N(S),\ N(\tilde{T}) = N(Q) \cap N(T) .$$

Let $\tilde{S}$ and $\tilde{T}$ be normally solvable, then (cf. Petryshyn [8])

$$\inf \{\|\tilde{S}x\| \,|\, x \in N(\tilde{S})^\perp,\ \|x\| = 1\} > 0 ,$$
$$\inf \{\|\tilde{T}y\| \,|\, y \in N(\tilde{T})^\perp,\ \|y\| = 1\} > 0 ,$$

and so $\langle x, -\tilde{S}x\rangle^{1/2}/N(\tilde{S})^\perp$, $\langle y, \tilde{T}y\rangle^{1/2}/N(\tilde{T})^\perp$ are equivalent norms to the given ones, resp., restricted correspondingly. With $\tilde{G}(x, y): = \langle x, \tilde{S}x\rangle + \langle y, \tilde{T}y\rangle$, $\tilde{p}_n: = F + r_n\tilde{G}$, $\tilde{B}: = \begin{pmatrix} \tilde{S} & 0 \\ 0 & \tilde{T} \end{pmatrix}$ and $\tilde{A}_n: = A + r_n\tilde{B}$, the solution set of $(X, Y, \tilde{p}_n)$ is

$$\tilde{A}_n^+(c, d) + N(\tilde{A}_n),\ (n \in N) .$$

Now $\tilde{A}_n^+(c, d) \perp N(\tilde{A}_n)$ and $N(\tilde{A}_n) = N(\tilde{S}) \times N(\tilde{T})$, thus $\tilde{A}_n^+(c, d)$ solves $(N(\tilde{S})^\perp, N(\tilde{T})^\perp, \tilde{p}_n)$, $(n \in N)$. Applying Theorem 4 to this problem, we get

(9) $$\tilde{A}^{+(S,T)}(c, d): = \lim_{n\to\infty} \tilde{A}_n^+(c, d)$$

solves uniquely

(10) $$\mathscr{M}_{\tilde{G}/\tilde{F}},\ \text{where } \tilde{F}: = F/N(\tilde{S})^\perp \times N(\tilde{T})^\perp .$$

Denote by Z the solution set of $\mathscr{M}_{\tilde{G}/F}$, and let (x_1, y_1), $(x_2, y_2) \in Z$; then we have by (4) for all $(u, v) \in N(P) \times N(Q)$:

$$\langle u, \tilde{S}^*\tilde{S}(x_1 - x_2)\rangle = 0,\ \langle v,\ \tilde{T}^*\tilde{T}(y_1 - y_2)\rangle = 0 .$$

With (4) again $(x_1 - x_2) \in N(P)$, $(y_1 - y_2) \in N(Q)$, and so $(x_1 - x_2) \in N(\tilde{S})$, $(y_1 - y_2) \in N(\tilde{T})$, hence we have, with (8), the representation

$$Z = (x, y) + N(\tilde{S}) \times N(\tilde{Q}),\ \text{for any } (x, y) \in Z .$$

Thus the element of Z with minimum norm is given by the solution of $\mathscr{M}_{\tilde{G}/\tilde{F}}$. Because of (6), (7) there are

$$\tilde{G}/X_F \times Y_F = G/X_F \times Y_F + \text{const} ,$$

and Z the solution set of $\mathscr{M}_{G/F}$, too. Then (9), (10) yield,

(11) $A^{+(S,T)}(c, d)$ is the solution of $\mathscr{M}_{G/F}$ with minimum norm.

Since (9) holds for all (c, d) in the range of A, we have just proved the

THEOREM 5. *Let with the definitions above $A, \tilde{S}, \tilde{T}$ be normally solvable, then there exists a linear and bounded operator*

$$A^{+(S,T)}\colon X \times Y \longrightarrow X \times Y$$

such that for all $(c, d) \in R(A)$

> *$A^{+(S,T)}(c, d)$ is the minimum norm solution of the two stage minimax problem* (1) *$\mathscr{M}_{G/F}(c, d)$, and permits the representation*

$$A^{+(S,T)}(c, d) = \lim_{r \to +0} \begin{pmatrix} P + r\tilde{S} & L \\ L^* & Q + r\tilde{T} \end{pmatrix}^+ (c, d) . \tag{12}$$

If $N(\tilde{S}) = \{0\}$, $N(\tilde{T}) = \{0\}$, then on the right hand side in (12) *we have ordinary inversion.*

Conveniently one takes

$$\tilde{S} = \begin{cases} S, & \text{if } N(S) = \{0\} \\ P + S, & \text{otherwise} \end{cases} , \quad \tilde{T} = \begin{cases} T, & \text{if } N(T) = \{0\} \\ Q + T, & \text{otherwise} \end{cases} .$$

6. A note on best approximate solutions of linear equations. Let W, X, Y be real Hilbert spaces as above and

$$C\colon X \longrightarrow Y, \ D\colon X \longrightarrow W$$

be continuous linear operators. We are given an element $y \in Y$ and the problem of finding an element $x \in X$ which solves the equation

$$Cx = y . \tag{1}$$

If $y \notin R(C)$, there exists no solution of (1). Then we consider the problem of finding an element $x(y) \in X$ of minimum seminorm $\|Dx\|$ which gives a minimum value for the discrepancy $\|Cx - y\|$, $x \in X$. An element $x(y)$ with this property may be called a '*D-best approximate solution*' of (1). In the case $D = I$ ($=$ identity) usually $x(y)$

is called a 'best approximate solution' (e.g., Holmes [6], p. 214) or 'pseudo-solution' (e.g., Morozov [7]) of (1). In order to find a D-best approximate solution of (1) we have to solve the problem

$$(2) \qquad \text{minimize } \{\langle x, D^*Dx\rangle \mid \langle x, C^*Cx\rangle - 2\langle x, C^*y\rangle = \min!,\ x \in X\}\ .$$

Applying now Theorem 5 to this special two stage problem (2) we get

THEOREM 6. *If C, $C^*C + D^*D$ are normally solvable, then there exists a continuous linear operator*

$$C^{+D}\colon \begin{array}{l} Y \longrightarrow X \\ y \longmapsto C^{+D}y\ , \end{array}$$

such that

for all $y \in Y$ $C^{+D}y$ is the D-best approximate solution to $Cx = y$ of minimum norm, $(x \in X)$,

and

$$(3) \qquad C^{+D} = \lim_{r \to +0} (C^*C + r\tilde{D})^+C^*\ ,$$

where

$$\tilde{D} = \begin{cases} D^*D, & \text{if } N(D) = \{0\} \\ C^*C + D^*D, & \text{otherwise} \end{cases}$$

If $N(\tilde{D}) = \{0\}$, then on the right hand side of (3) we have ordinary inversion, and especially for $D = I$ we get

$$(4) \qquad C^{+I} \equiv C^+ = \lim_{r \to +0} (C^*C + rI)^{-1}C^*\ ,$$

a representation given for instance by Morozov [7].

ACKNOWLEDGMENT. The author would like to thank the referee for helpful suggestions.

REFERENCES

1. K. J. Arrow, L. Hurwicz and H. Uzawa, *Studies in Linear and Non-linear Programming, part II*, Stanford University Press, Stanford (California), 1958.
2. R. Courant, *Variational Methods for the Solution of Problems of Equilibrium and Vibrations*, Bull. Amer. Math. Soc., **49** (1943), 1-23.
3. V. F. Dem'janov, *Successive approximations for finding saddle points*, Sov. Math. Dokl., **8** (1967), 6, 1350-1353.
4. A. V. Fiacco and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Wiley, New York, 1968.
5. J. Hartung, *An extension of Sion's minimax theorem and an application to a method for constrained games*, Pacific J. Math., to appear.

6. R. B. Holmes, *A course on optimization and best approximation*, Lect. Not. Math., 257, Springer, Berlin/Heidelberg/New York, 1972, 214–233.
7. V. A. Morozov, *Pseudo-solutions*, U.S.S.R. Comp. math. Phys., **9** (1969), 6, 196–203.
8. W. Petryshyn, *On generalized inverses and on the uniform convergence of* $(I-\beta K)^n$ *with application to iterative methods*, J. Math. Anal. Appl., **18** (1967), 417–439.
9. A. N. Tihonov, *Methods for the regularization of optimal control problems*, Sov. Math. Dokl., **6** (1965), 4, 761–763.

Received January 21, 1980 and in revised form July 28, 1981.

UNIVERSITÄ DORTMUND
ABTEILUNG STATISTIK
POSTFACH 50 05 00
D-4600 DORTMUND 50
WEST GERMANY