INSTITUTUL
DE
MATEMATICĂ

INSTITUTUL NAȚIONAL
PENTRU CREAȚIE
ȘTIINȚIFICĂ ȘI TEHNICĂ

ISSN 0250 3638

THE RELATION BETWEEN MINIMAX AND MINIMUM; CONVEX PROGRAMMING VIA MINIMAX

by

Dan TIBA

PREPRINT SERIES IN MATHEMATICS
No.8/1985

BUCUREȘTI

THE RELATION BETWEEN MINIMAX AND MINIMUM; CONVEX PROGRAMMING VIA MINIMAX

by

Dan TIBA*)

January 1985

*) Department of Mathematics, The National Institute for Scientific and Technical Creation, Bd.Pacii 220, 79622 Bucharest, Romania.

THE RELATION BETWEEN MINIMAX AND MINIMUM; CONVEX PROGRAMMING VIA MINIMAX

Dan TIBA

1. INTRODUCTION

Let X and Y be Banach spaces with duals X^*, Y^* and $\mathcal{K}: X\times Y \to [-\infty, +\infty]$ be a closed, proper, saddle (concave-convex) function. By means of a partial Fenchel-Legendre transform, Moreau [4], Rockafellar [5], established a one-to-one correspondence between closed saddle functions on XxY and convex, lower semicontinuous functions on $X\times Y^*$ which preserves, in a certain sense, the subdifferentials. For more details and for a general background in convex analysis we quote the monographs of R.T. Rockafellar [6], I.Ekeland and R.Temam [3], V.Barbu and Th. Precupanu [1].

We associate with each saddle function $\mathcal{K}$ a convex, lower semicontinuous function L on XxY such that any saddle point of $\mathcal{K}$ is a minimum point of L and conversely. A similar idea, for the case of quadratic forms, appears in Ekeland [2], but our approach is different.

In section 2 we define and study this transform. As a Corollary we prove that the set of saddle points of a given concave-convex function is rectangular, i.e. the product of two convex, closed sets from X, respectively Y.

In section 3 we apply the results to the classical convex programming problem. We obtain necessary and sufficient conditions in terms of the Lagrangian for the validity of the Kuhn--Tucker property. We point out, via minimax, a sufficient condition for the Kuhn-Tucker characterization, weaker than the Slater condition.

Other conclusions, which also seem to be new, are that the Slater "interiority" assumption can be viewed as a coercivity assumption on the Lagrangian and that if λ_o is a Lagrange multiplier corresponding to a certain solution x_o of a given convex programming problem, then λ_o is Lagrange multiplier for any other solution of the problem.

Throughout this paper, we write the symbol $|.|$ for all the norms we use.

2. MINIMAX AND MINIMUM

We define the convex function $L: X\times Y \to]-\infty,+\infty]$ associated with $\mathcal{K}$ by:

$$L(x,y)=\sup_{t,z}\left\{ cl_2\mathcal{K}(t,y)-cl_1\mathcal{K}(x,z)\right\} \tag{2.1}$$

where cl_i , $i=1,2$, denote the closure with respect to the corresponding argument of $\mathcal{K}$.

Proposition 1. *L is a convex, lower semicontinuous function on XxY which never assumes the value* $-\infty$.

Remark 2. L may be identically $+\infty$, as the following example shows

$$\mathcal{K}(x,y)=\begin{cases} x(y^2+1) & x,y\in R,\ x\geqslant 0\ , \\ -\infty & \text{otherwise.} \end{cases}$$

Proposition 3. Assume $\mathcal{K}$ to be an upper-lower semicontinuous function, then dom $L\subset \mathrm{dom}\,\mathcal{K}$.

Proof

$$(2.2)\qquad L(x,y)=\sup_t \mathcal{K}(t,y)-\inf_z \mathcal{K}(x,z).$$

Let $[x,y]\in$ dom L, then $\sup_t \mathcal{K}(t,y)<+\infty$ and $\inf_z \mathcal{K}(x,z)>-\infty$ therefore $x\in \mathrm{dom}_1\mathcal{K}$ and $y\in \mathrm{dom}_2\mathcal{K}$, that is $[x,y]\in \mathrm{dom}\,\mathcal{K}$.

Now,we state the correspondence theorem:

Theorem 4. If the closed, proper, saddle function $\mathcal{K}$ has a saddle point, then the minimum value of L is zero. Any saddle point of $\mathcal{K}$ is a global minimum of L and conversely any minimum of L is a saddle point of $\mathcal{K}$.

Proof

Assume that

$$(2.3)\qquad \mathcal{K}(x,\tilde{y})\leqslant \mathcal{K}(\tilde{x},\tilde{y})\leqslant \mathcal{K}(\tilde{x},y)\ , \qquad \forall\ x,y\in X\times Y\ .$$

It is known that $[\tilde{x},\tilde{y}]$ is a saddle point with the same saddle value for $cl_1\mathcal{K}$, $cl_2\mathcal{K}$ too. We have:

$$L(x,y)=\sup\{ cl_2\mathcal{K}(t,y)-cl_1\mathcal{K}(x,z)\}\geqslant cl_2\mathcal{K}(\tilde{x},y)-cl_1\mathcal{K}(x,\tilde{y})=$$

$$=cl_2\mathcal{K}(\tilde{x},y)-cl_2\mathcal{K}(\tilde{x},\tilde{y})+cl_1\mathcal{K}(\tilde{x},\tilde{y})-cl_1\mathcal{K}(x,\tilde{y})\geqslant 0.$$

On the other hand, by the minimax inequalities (2.3), we get

$$L(\tilde{x},\tilde{y})=cl_2\mathcal{K}(\tilde{x},\tilde{y})-cl_1\mathcal{K}(\tilde{x},\tilde{y})=0.$$

Conversely, let $[a,b]$ be a global minimum for L, that is $L(a,b)=0$. Then

(2.4) $$cl_2\mathcal{K}(t,b)\leq cl_1\mathcal{K}(a,z) \qquad \forall\ t,z\in X\times Y$$

Take the closure with respect to z in (2.4). Since $\mathcal{K}$ is closed $cl_2 cl_1\mathcal{K}=cl_2\mathcal{K}$ and we obtain

$$cl_2\mathcal{K}(t,b)\leq cl_2\mathcal{K}(a,z) \qquad \forall\ t,z\in X\times Y$$

and the proof is finished.

Remark 5. Without assuming that $\mathcal{K}$ has a saddle point, the theorem is not true since L may be identically $+\infty$.

Moreover, from the proof, it is obvious that the only essential hypotheses are the existence of saddle points, respectively minimum points, therefore the result may be extended to more general classes of functions.

Remark 6. If the transform L is proper and $\min L=0$ the converse part of Theorem 4 follows without assuming that $\mathcal{K}$ has saddle points.

Corollary 7. If $\mathcal{K}$ satisfies the coercivity condition

(2.5) there are $\bar{x},\bar{y}\in X\times Y$ such that

$$\lim_{|x|+|y|\to\infty}\{\mathcal{K}(x,\bar{y})-\mathcal{K}(\bar{x},y)\}=-\infty$$

then L is proper in reflexive spaces X,Y.

Proof

By Corollary 3.4, Barbu-Precupanu [1], p.138, $\mathcal{K}$ has a saddle point on XxY and Theorem 4 may be applied.

Corollary 8. The set of all saddle points of a closed function $\mathcal{K}$ in Banach space XxY is rectangular.

Proof.

Let $\varphi(y)=\sup_t \mathrm{cl}_2\mathcal{K}(t,y)$, $\psi(x)=\sup_z\{-\mathrm{cl}_1\mathcal{K}(x,z)\}$. Then $L(x,y)= =\psi(x)+\varphi(y)$ and, of course, the set of minimum points of L is rectangular. By Theorem 4 the proof is finished.

Now, we turn to the reciprocal of the correspondence (2.1).

Let $L: X\times Y \to]-\infty,+\infty]$ be a convex, lower semicontinuous, proper function. We define

$$(2.6)\quad \mathcal{K}(x,y)=\begin{cases}\inf_t L(t,y)-\inf_z L(x,z) & \text{if one is finite,}\\ -\infty & \text{otherwise.}\end{cases}$$

Proposition 9. $\mathcal{K}$ is a concave - convex function.

Proof

We show that $l(y)=\inf_t L(t,y)$ is convex. For any $\varepsilon>0$ there are points t_1^ε, t_2^ε:

$$L(t_1^\varepsilon,y_1)-\varepsilon\leq l(y_1)\ ,$$

$$L(t_2^\varepsilon,y_2)-\varepsilon\leq l(y_2)\ .$$

Then, for all $\lambda\in[0,1]$

$$\lambda l(y_1)+(1-\lambda)l(y_2)\geq\lambda L(t_1^\varepsilon,y_1)+(1-\lambda)L(t_2^\varepsilon,y_2)-\varepsilon\geq$$

$$L(\lambda t_1^\varepsilon+(1-\lambda)t_2^\varepsilon,\lambda y_1+(1-\lambda)y_2)-\varepsilon\geq l(\lambda y_1+(1-\lambda)y_2)-\varepsilon\ .$$

Proposition 10. If L satisfies in reflexive spaces

$$(2.7) \qquad \lim_{|x|+|y|\to\infty} L(x,y)=+\infty$$

then $\mathcal{K}$ is upper-lower semicontinuous.

Proof

We show that l is lower semicontinuous.

Let $y_n \to y$ in Y and $l(y_n)\leq M$. For any $\varepsilon>0$ there is t_n^{ε} such that $L(t_n^{\varepsilon},y_n)\leq M+\varepsilon$. Condition (2.7) gives $\{t_n^{\varepsilon}\}_n$ bounded in X.

On a subsequence, $t_n^{\varepsilon} \rightharpoonup t^{\varepsilon}$ weakly in X and as L is weakly lower semicontinuous, we have

$$L(t^{\varepsilon},y)\leq \liminf L(t_n^{\varepsilon},y_n)\leq M+\varepsilon$$

therefore $l(y)\leq M+\varepsilon$ and the proof is finished.

Proposition 11. If $[\bar{x},\bar{y}]$ is a minimum point for L then it is a saddle point for $\mathcal{K}$ and $\mathcal{K}(\bar{x},\bar{y})=0$.

Proof

$$\mathcal{K}(\bar{x},\bar{y})=\inf_t \sup_z \{L(t,\bar{y})-L(\bar{x},z)\}=\inf_t L(t,\bar{y})-\inf_z L(\bar{x},\bar{z})=$$

$$=L(\bar{x},\bar{y})-L(\bar{x},\bar{y})=0.$$

$$\mathcal{K}(x,\bar{y})=\inf_t \sup_z \{L(t,\bar{y})-L(x,z)\}=L(\bar{x},\bar{y})-\inf_z L(x,z)\leq 0.$$

$$\mathcal{K}(\bar{x},y)=\inf_t \sup_z \{L(t,y)-L(\bar{x},z)\}=\inf_t L(t,y)-L(\bar{x},\bar{y})\geq 0.$$

Remark 12. Generally, dom $L\subset$ dom$\mathcal{K}$ and also the set of minimum points of L is included in the set of saddle points of $\mathcal{K}$, as the following example shaws; $L:R\times R\to]-\infty,+\infty]$

$$L(x,y)=\begin{cases} 0 & \text{on the unit disc,} \\ + & \text{otherwise .} \end{cases}$$

Then, the associated saddle function is

$$\mathcal{K}(x,y)=\begin{cases} 0 & \text{for } |x|\leqslant 1,\ |y|\leqslant 1 \\ +\infty & \text{for } |y|>1,\ |x|\leqslant 1 \\ -\infty & \text{otherwise.} \end{cases}$$

3. CONVEX PROGRAMMING

We consider the standard problem

(P) Minimize $f(x)$

subject to

(3.1) $g_i(x)\leqslant 0 \quad i=\overline{1,n}$,

(3.2) $r_j(x)=0 \quad j=\overline{1,m}$.

Above $f,g_i:X \to]-\infty,+\infty]$ are convex, lower semicontinuous functions, $r_j:X\to R$ are affine, continuous functions.

We define the Lagrangian associated with (P) by $\mathcal{K}:R^{n+m}xX \to [-\infty,+\infty]$

$$(3.3)\ \mathcal{K}(\lambda,\mu,x)=\begin{cases} f(x)+\sum_{i=1}^{n}\lambda_i g_i(x)+\sum_{j=1}^{m}\mu_j r_j(x) & \text{if } \lambda_i\geqslant 0,\ i=\overline{1,n} \\ -\infty & \text{otherwise .} \end{cases}$$

The question is to characterize the solutions of the constrained optimization problem (P) as saddle points of $\mathcal{K}$.

<u>Theorem 13. If $\mathcal{K}$ has saddle points then</u> (P) <u>has solutions and a point x_o is an optimal solution for</u> (P) <u>iff there exist</u> $\lambda^o=(\lambda_1^o,\dots,\lambda_n^o)$ <u>positive and</u> $\mu^o=(\mu_1^o,\dots,\mu_m^o)$ <u>such that</u>

$[\lambda^o, \mu^o, x_o]$ is a saddle point for $\mathcal{K}$.

Proof

If $\mathcal{K}$ has saddle points, Uzawa [8] proved that the projection on the X coordinate is a solution of (P).

For the second part of the theorem we compute the convex function $L(\lambda,\mu,x)=\Psi(\lambda,\mu)+\varphi(x)$ associated with $\mathcal{K}$ by (2.1):

(3.4) $$\varphi(x)=\begin{cases} f(x) & \text{if } g_i(x)\leq 0,\ r_j(x)=0, \quad \forall\ i,j \\ +\infty & \text{otherwise .}\end{cases}$$

(3.5) $$\Psi(\lambda,\mu)=-\inf_x \mathcal{K}(\lambda,\mu,x) \ .$$

Since $\mathcal{K}$ has saddle points one may use Theorem 4 and the conclusion follows.

Corollary 14. If $\mathcal{K}$ has saddle points, we have

(3.6) $$\min(P)=\min_x \max_{\lambda,\mu} \mathcal{K}(\lambda,\mu,x)$$

Proof

This follows by $\min L(\lambda,\mu,x)=0$.

Corollary 15. If the solution x_o of (P) has the multiplier $[\lambda^o,\mu^o]$, $\lambda^o\geq 0$, then $[\lambda^o,\mu^o]$ is a multiplier for any other solution of (P).

Proof

By Corollary 8

Remark 16. If the Kuhn-Tucker property is valid and (P) has solutions, obviously $\mathcal{K}$ has saddle points. Therefore, Theorem 13 shows that the statement i) is equivalent with ii) plus iii):

i) $\mathcal{K}$ has saddle points

ii) (P) has solutions

iii) for every solution x_o of (P) there are $\lambda^o \geq 0$, μ^o such that $[\lambda^o, \mu^o, x_o]$ is a saddle point for $\mathcal{K}$.

Remark 17. By the standard condition for minimax, the search of the solutions and of the multipliers is reduced to solving the equation

$$[0,0] \in \partial\mathcal{K}(\lambda,\mu,x) .$$

Any minimax condition for $\mathcal{K}$ may be viewed as a "constrained qualification" for problem (P). For instance

Corollary 18. If there are $\bar{x}, \bar{\mu}, \bar{\lambda} \geq 0$, $M \geq 0$ such that

(3.7) $$\mathcal{K}(\lambda,\mu,\bar{x}) - \mathcal{K}(\bar{\lambda},\bar{\mu},x) \leq 0$$

for $|\lambda|+|\mu|+|x| \geq M$, then (P) has solutions and for every solution x_o of (P) there are $\lambda^o \geq 0, \mu^o$ such that $[\lambda^o, \mu^o, x_o]$ is a saddle point for $\mathcal{K}$.

Proof

(3.7) is also a minimax condition weaker than (2.5).

Now, we give some brief considerations on the Slater condition. We take m=0 (no equality constraints) and we assume

(3.8) $$\lim_{|x| \to \infty} f(x) = +\infty \quad .$$

The Slater condition is

(3.9) there is $\bar{x}$ in the domain of f such that

$$g_i(\bar{x}) < 0 \quad , \quad i=\overline{1,n}$$

Corollary 19. Under the above hypotheses $\mathcal{K}$ has saddle points.

Proof

For $\lambda \geq 0$, we have

$$\mathcal{K}(0,x)-\mathcal{K}(\lambda,\bar{x})=f(x)-f(\bar{x})+\sum_{i=1}^{n}(-g_i(\bar{x}))\lambda_i$$

and (2.5) is fulfilled.

Remark 20. Under the quite usual assumption (3.8), the Slater condition implies the coercivity condition (2.5) on the Lagrangian and it is stronger than (3.7).

Remark 21. The general problem (P), under assumption (3.8), (3.9) may be easily reduced to Corollary 19. Denote $Z=\{x \in X;\ r_j(x)=0,\ j=\overline{1,m}\}$ nad by $r_j^* \in X^*$, $j=\overline{1,m}$, the elements defining r_j.

Let $i(x)$ be the indicator function of Z and $f_1=f+i$. Then (P) reduces formally to the case m=0 and we obtain the multipliers $\lambda_i \geq 0$, $i=\overline{1,n}$ corresponding to g_i. By the remark that

$$\partial i(x)=\{\ x^* \in X^*;\ x^*=\sum_{j=1}^{m}\mu_j r_j^*\ ,\quad \forall \mu_j \in R\}$$

and some computations, we finish the proof.

Remark 22. Without assuming (3.8), direct comparisons of (3.7) with the Slater conditions are not simple. We quote the paper of J. Stoer [7], where the proof of Theorem 2.16 may be interpreted in this sense, in the case m=0.

Remark 23. In the proof of Corollary 19 we use $\bar{\lambda}=0$, which generally is not a Lagrange multiplier for (P). Therefore the search of $\bar{\lambda}$, $\bar{\mu}$ in (3.7) doesn't reduce to the search of the

Lagrange multipliers.

Remark 24. Important features of condition (3.7) are that it involves all the elements defining (P) and that it is also intimately related to the existence theory for (P).

Finally, we give a general example where the Slater condition is not fulfilled.

Let f be a convex, lower semicontinuous, proper function and g be the indicator of a convex, closed set in X. Consider the problem

$$(P_1) \qquad \min f(x) \ , \qquad g(x) \leq 0 .$$

Corollary 25. Assume that (P_1) has solutions. Then, for every solution x_o of (P_1) there is $\lambda_o \geq 0$ such that $[\lambda_o, x_o]$ is a saddle point for $\mathcal{K}$.

Proof

$$\mathcal{K}(\lambda, x_o) - \mathcal{K}(0,x) = \begin{cases} f(x_o) - f(x) & \text{if } g(x) \leq 0, \\ -\infty & \text{otherwise} \end{cases}$$

and (3.7) is satisfied for any $\lambda \geq 0$, $x \in X$.

REFERENCES

1. V.Barbu; Th.Precupanu - "Convexity and optimization in Banach spaces", Ed.Acad. - Noordhoff, Leyden, 1978.

2. I.Ekeland - "Dual variational methods in non-convex optimization and differential equations", Lecture Notes in Mathematics 979, Springer Verlag 1983, 108-121.

3. I.Ekeland; R.Temam - "Analyse convexe et problèmes variationels", Dunod, Paris, 1974.

4. J.J.Moreau - "Théorèmes inf-sup", C.R.A.S.Paris 258, 1964, 2720-2722.

5. R.T.Rockafellar - "A general correspondence between dual minimax problems and convex programs", Pacific J. Math., 25, 1968, 597-611.

6. R.T.Rockafellar - "Convex analysis", Princeton Univ. Press, Princeton, N.J., 1969.

7. J.Stoer - "Duality in nonlinear programming and the minimax theorem", Numer. Math. 5, 1963, 371-379.

8. H.Uzawa - "The Kuhn-Tucker theorem in concave programming" in K.J. Arrow, L.Hurwicz and H.Uzawa, "Studies in linear and nonlinear programming", Stanford Univ. Press, 1958, 32-37.

Department of Mathematics, INCREST
Bd.Păcii 220, 79622 Bucharest
Romania