Bull. Korean Math. Soc. **23**(1986), No. 2, pp. 141-148

A PSEUDOCONVEX PROGRAMMING IN A HILBERT SPACE

BYUNG HO YOON AND IN SOO KIM

1. Introduction

In [1], M. Guignard considered a constraint set in a Banach space, which is similar to that in [2] and gave a first order necessary optimality condition which generalized the Kuhn-Tucker conditions [3]. Sufficiency is proved for objective functions which is either pseudoconcave [5] or quasi-concave [6] where the constraint sets are taken pseudoconvex.

In this note, we consider a pseudoconvex programming problem in a Hilbert space. Constraint set in a Hilbert space being pseudoconvex and the objective function is restrained by an operator equation. Then we use the methods similar to that in [1] and [6] to obtain a necessary and sufficient optimality condition.

2. Preliminaries

Let U be a real Hilbert space. For u, v in U, we use the notation $\langle u, v\rangle$ to denote the scalar product of u and v. Thus, for $u\in U$, $\|u\|=\langle u, u\rangle^{1/2}$ is the norm of u. If f is an element of U^*, topological dual of U, then the value of f at u will be denoted by (f, u).

Let M be a subset of U. The closure of M will be denoted by $\overline{M}$ and (M) denotes the convex hull of M; that is, the smallest convex set in U containing M.

DEFINITION 2.1. A subset C of U is called a cone if, for any $u\in C$, $\alpha u\in C$ for all $\alpha\geq 0$.

DEFINITION 2.2. Let C be a cone in U. Subsets of U^* defined by

$$C^-=\{f\in U^* \mid (f, u)\leq 0 \text{ for all } u\in C\}$$
$$C^+=\{f\in U^* \mid (f, u)\geq 0 \text{ for all } u\in C\}$$

are, respectively, called the negative and positive polar cone of C.

Received March 4, 1986.
This research was supported by MOE grant, 1985.

Note that C^- and C^+ are closed convex cones in U^*.

Let M be a nonempty subset of U and $u_0 \in M$. Cones tangent and pseudotangent to M are defined as follows:

DEFINITION 2.3. A vector $v \in U$ is said to be tangent to M at u_0 if there exists a sequence $\{u_k\}$ in M converging to u_o and a sequence $\{\lambda_k\}$ of nonnegative real numbers such that the sequence $\{\lambda_k(u_k - u_o)\}$ converges to v.

DEFINITION 2.4. The set $T(M, u_o)$ of all the vectors tangent to M at u_o is called the cone tangent to M at u_o.

DEFINITION 2.5. The closure of the convex hull of $T(M, u_o)$ is called the cone pseudotangent to M at u_o and is denoted by $P(M, u_o)$.

Note that $T(M, u_o)$ is a cone and $P(M, u_o)$ is a closed convex cone.

Let $\{M_i \mid i \in I\}$ be any family of subsets of U such that $\bigcap_{i \in I} M_i \neq \phi$ and let

$$M = \bigcap_{i \in I} M_i, \quad \tilde{M} = \bigcup_{i \in I} M_i$$

For $u_o \in M$, defined cones have the following properties.

PROPOSITION 2.1. $T(M, u_o) \subset \bigcap_{i \in I} T(M_i, u_o)$, $P(M, u_o) \subset \bigcap_{i \in I} P(M_i, u_o)$

Proof. It is clear that $T(M, u_o) \subset T(M_i, u_o)$ for all $i \in I$. Thus

$$T(M, u_o) \subset \bigcap_{i \in I} T(M_i, u_o) \subset \bigcap_{i \in I} P(M_i, u_0)$$

But the intersection of any number of closed convex cones is a closed convex cone. Thus $\bigcap_{i \in I} P(M_i, u_0)$ is a closed convex cone containing $T(M, u_o)$. Hence, by Definition 2.5, we have

$$P(M, u_o) \subset \bigcap_{i \in I} P(M_i, u_o).$$

PROPOSITION 2.2. $\bigcup_{i \in I} T(M_i, u_0) \subset T(\tilde{M}, u_o)$, $\bigcup_{i \in I} P(M_i, u_o) \subset P(\tilde{M}, u_o)$

Proof. It is clear that $T(M_i, u_o) \subset T(\tilde{M}, u_o)$ for all $i \in I$. Thus

$$\bigcup_{i \in I} T(M_i, u_o) \subset T(\tilde{M}, u_o)$$

Moreover,

$$T(M_i, u_0) \subset T(\tilde{M}, u_o) \subset P(\tilde{M}, u_0)$$

for all $i \in I$. Since $P(M_i, u_o)$ is the smallest closed convex cone containing

$T(M_i, u_o)$, it follows that $P(M_i, u_0) \subset P(\tilde{M}, u_o)$ for all $i \in I$.

DEFINITION 2.6. Let M be a subset of U and $u_o \in M$. If, for every $u \in M$, $u - u_o$ belongs to $P(M, u_o)$, then we say M is pseudoconvex at u_o. If M is pseudoconvex at all of its points, then we simply say that M is pseudoconvex.

PROPOSITION 2.3. *If all M_i, $i \in I$, are pseudoconvex at u_o, then $\tilde{M} = \bigcup_{i \in I} M_i$ is pseudoconvex at u_o.*

Proof. Let u be any element of $\tilde{M}$. Then $u \in M_i$ for some $i \in I$. Since M_i is pseudoconvex at u_o, $u - u_o \in P(M_i, u_o)$. But then, by Proposition 2.5, $u - u_o$ belongs to $P(M, u_o)$.

REMARK. The intersection of pseudoconvex sets need not be pseudoconvex. For an example, we refer to [1].

PROPOSITION 2.4. If a subset M of U is convex, then M is pseudoconvex at every u_o in M.

Proof. Let $u \in M$ and $\{\lambda_k\}$ be a sequence of numbers such that $0 < \lambda_k < 1$ and converges to 0. For each k, let $u_k = u_o + \lambda_k(u - u_o)$. Then $\{u_k\}$ is a sequence in M converging to u_o. Let $\mu_k = 1/\lambda_k$ for each k. Then, for every k, $\mu_k(u_k - u_o) = u - u_o$ so that $u - u_o \in T(M, u_o) \subset P(M, u_o)$ which proves that M is pseudoconvex at u_o.

REMARK. The converse of Proposition 2.4 is not true as the following example shows.

In 1-dimensional Euclidean space E^1, let Q be the set of rational numbers and r_o be any element of Q. We will show that Q is pseudoconvex at r_o. Let $r \in Q$ and first consider the case $r > r_0$. Let $\{r_k\}$ be a sequence in Q such that $r_k > r_o$ and converges to r_o. For each k, let $\lambda_k = (r - r_o)/(r_k - r_o)$. Then $\{\lambda_k\}$ is a sequence of positive numbers and, clearly, $\lambda_k(r_k - r_o)$ converges to $r - r_o$. Thus $r - r_o \in T(Q, r_o)$. For the cases $r < r_o$ and $r = r_o$, we can similarly prove that $r - r_o$ belongs to $T(Q, r_o)$. Hence Q is pseudoconvex which is not convex.

DEFINITION 2.7. Let $\phi(u)$ be a real function of $u \in U$. ϕ is said to be quasi-convex if, for any real number λ, the set $\{u \in U \mid \phi(u) \leq \lambda\}$ is convex.

It is clear that any convex function is quasi-convex. Quasi-convex

Fréchet differentiable functions have the following properties. ∇ will denote the differential operator.

PROPOSITION 2.5. *If ϕ is quasi-convex and Fréchet-differentiable at $\bar{u}\in U$, then $(\nabla\phi(\bar{u}),\ u-\bar{u})>0$ implies $\phi(u)-\phi(\bar{u})>0$.*

The Proof of this proposition can be found in [7].

DEFINITION 2.8. Let $\phi(u)$ be a real function of $u\in U$, $M\subset U$ and $u_0\in M$. If ϕ is Fréchet-differentiable at u_0 and, for $u\in M$, $(\nabla\phi(u_0), u-u_0)\geq 0$ implies $\phi(u)-\phi(u_0)\geq 0$, then ϕ is said to be pseudoconvex over M at u_0. If ϕ is pseudoconvex over M at every $u_0\in M$, then we say that ϕ is pseudoconvex over M. If, in particular, M is equal to the whole space U, then we simply say that ϕ is pseudoconvex.

It can be proved as in [7] that convex Fréchet-differentiable functions are pseudoconvex and pseudoconvex functions are quasi-convex.

3. Pseudoconvex Programming

Suppose X is another real Hilbert space. For x, y in X and $g\in X^*$, we use the same notations, as in section 2 (without possible confusion), $\langle x, y\rangle$ and (g, x) to denote the scalar product and the value of g at x, respectively. We also use the notation $L(X, U)$ to denote the set of all bounded linear operators from X into U.

Let us now describe the problem that we wish to consider in this note.

Suppose we are given (i) a surjective linear operator $A\in L(X, U)$, (ii) a subset V of U which is pseudoconvex and (iii) a function $\phi : U\rightarrow E^1$ which is pseudoconvex over V. And we wish to consider the problem of minimizing the function

$$F(x, u)=\frac{1}{2}\|x\|^2+\phi(u) \tag{3.1}$$

over $u\in V$ and $Ax=u$. This problem is a pseudoconvex programming problem and will be called Prob. (PC).

Following well known facts which can be found in [8] are needed in the discussion of our problem.

DEFINITION 3.1. Suppose $S\in L(X,\ U)$. We say S is right invertible if there exists a bounded linear operator $T : U\rightarrow X$ such that $ST=I$ where I is the identity map on U.

Thus if $S \in L(X, U)$ is right invertible, then S is surjective. The converse is also true from the following proposition.

PROPOSITION 3.1. *If* $S \in L(X, U)$ *is surjective, then* S *is right invertible.*

Though this proposition is well known, we give a proof for the purpose of later use.

Proof. Let $M = \text{Ker } S$ be the kernel of S and N be the orthogonal complement of M. Let $P : X \to M$ and $Q : X \to N$ be the orthogonal projections. If $x \in X$, then $x = Px + Qx$ and thus $Sx = SPx + SQx = SQx$. Let $S_1 = SQ$. Then S_1 considered as a map from N into U is a bounded linear operator. Moreover, $S_1 : N \to U$ is a bijection. In fact, suppose $u \in U$. Since S is surjective, there exists $x \in X$ such that $Sx = u$. But then $Qx \in N$ and

$$S_1Qx = SQ^2x = SQx = Sx = u.$$

Thus $S_1 : N \to U$ is surjective. In order to show that S_1 is injective, suppose $S_1x = 0$ with $x \in N$. Then $Qx = x$ and $Sx = SQx = 0$. Thus $x \in M$. Since $x \in N$, it follows that $x = 0$.

Now let $S_+ = S_1^{-1}$. Then $S_+ \in L(U, N)$ and, for any $u \in U$,

$$SS_+u = SS_1^{-1}u = SQS_1^{-1}u = S_1S_1^{-1}u = u.$$

This completes the proof of the proposition.

Let $J : X \to X^*$ be the duality operator from X onto its dual X^*: that is, J is an isometry from X onto X^* such that $(Jx, y) = \langle x, y \rangle$ for all $x, y \in X$. Then the right inverse S_+ of S in Proposition 3.1 satisfies the following.

PROPOSITION 3.2. *Let* $S \in L(X, U)$ *be a surjective linear operator. Then* $SJ^{-1}S^*$ *is invertible, where* S^* *is the transpose of* S, *and the right inverse* S_+ *of* S *is given by*

$$S_+ = J^{-1}S^*(SJ^{-1}S^*)^{-1}.$$

Moreover,

$$S_+^*JS_+ = (SJ^{-1}S^*)^{-1}.$$

The Proof of this proposition can be found in [8].

Let us now return to our problem Prob. (PC). Operator $A \in L(X, U)$ given in Prob. (PC) is surjective. Thus A has the right inverse $A_+ \in L(U, X)$ and, for any $u \in U$, A_+u belongs to the orthogonal complement of Ker A. For each $u\in V$, let

$$K_u=\{x\in X\mid Ax=u\}.$$

Then K_u is closed and convex. Moreover, if $x\in K_u$, then

$$A(A_+u-x)=u-Ax=u-u=0$$

so that $A_+u-x\in\text{Ker } A$. Thus $\langle A_+u, A_+u-x\rangle=0$ for all $x\in K_u$. This implies that A_+u is the best approximate projector of zero onto K_u and thus

$$\|A_+u\|=\inf_{x\in K_u}\|x\|.$$

Therefore, the problem of minimizing the function $F(x,u)$ of (3.1) over $u\in V$ and $Ax=u$ is equivalent to the problem of minimizing

$$f(u)=\frac{1}{2}\|A_+u\|^2+\phi(u) \tag{3.2}$$

over $u\in V$.

Following lemma can be proved easily by a straightforward computation.

LEMMA 3.1. *The function* $\phi(u)=\frac{1}{2}\|A_+u\|^2$ *of* $u\in U$ *is convex and Fréchet-differentiable with the derivative*

$$\nabla\phi(u)=A_+^*JA_+u=(AJ^{-1}A^*)^{-1}u.$$

Thus the objective function $f(u)$ of (3.2) is pseudoconvex over V and the derivative is given by

$$\nabla f(u)=(AJ^{-1}A^*)^{-1}u+\nabla\phi(u).$$

THEOREM 3.1. *A necessary and sufficient condition for* $u_o\in V$ *to minimize* $f(u)$ *of* (3.2) *over* V *is that* $\nabla f(u_o)\in P^+(V,u_o)$.

Proof. Suppose $u_o\in V$ minimizes $f(u)$ over V. Let $u\in T(V,u_o)$. Then there exists a sequence $\{u_k\}$ in V converging to u_o and a sequence $\{\lambda_k\}$, $\lambda_k>0$ for all k, such that

$$\lim_k \lambda_k(u_k-u_o)=u.$$

Since u_o minimizes $f(u)$ over V, $f(u_k)-f(u_o)\geq 0$ for all k. Moreover,

$$f(u_k)-f(u_o)=(\nabla f(u_o), u_k-u_o)+o(\|u_k-u_o\|).$$

Thus

$$(\nabla f(u_o),\ \lambda_k(u_k-u_o))\geq\frac{-o(\|u_k-u_0\|)}{\|u_k-u_0\|}\cdot\lambda_k\|u_k-u_0\|.$$

Letting k go to infinite, we have

$$(\nabla f(u_0), u) \geq o\|u\| = 0.$$

Therefore,

$$\nabla f(u_0) \in T^+(V, u_0) = P^+(V, u_0).$$

This proves the necessity.

Conversely, let $u \in V$. Since V is pseudoconvex at u_0, we have $u - u_0 \in P(V, u_0)$. Thus if $\nabla f(u_0) \in P^+(V, u_0)$, then $(\nabla f(u_0), u - u_0) \geq 0$. But f is pseudoconvex over V at u_0. Hence $f(u_0) \leq f(u)$ and u_0 minimizes $f(u)$ over V.

In order to illustrate how we could apply Theorem 3.1, we consider the following example.

EXAMPLE. Suppose, in Prob. (PC), $\phi(u) = \frac{1}{2}\|u\|^2$; that is,

$$f(u) = \frac{1}{2}\|A_+ u\|^2 + \frac{1}{2}\|u\|^2 \tag{3.3}$$

and V is convex. If $u_0 \in V$ minimizes $f(u)$ over V, then

$$\langle A_+ u_0, A_+ u \rangle + \langle u_0, u \rangle \geq \|A_+ u_0\|^2 + \|u_0\|^2$$

for all $u \in V$.

Proof. Let $K : U \to U^*$ be the duality operator. Then

$$\nabla f(u_0) = A_+{}^* J A_+ u_0 + K u_0.$$

Thus, by Theorem 3.1, if u_0 minimizes $f(u)$ of (3.3) over V, then

$$A_+{}^* J A_+ u_0 + K u_0 \in P^+(V, u_0).$$

Since V is convex, $u - u_0 \in P(V, u_0)$ for all $u \in V$. Thus if u_0 minimizes $f(u)$ of (3.3), then

$$\begin{aligned} 0 &\leq (A_+{}^* J A_+ u_0, u - u_0) + (K u_0, u - u_0) \\ &= (J A_+ u_0, A_+(u - u_0)) + (K u_0, u - u_0) \\ &= \langle A_+ u_0, A_+(u - u_0) \rangle + \langle u_0, u - u_0 \rangle \end{aligned}$$

for all $u \in V$.

References

1. M. Guignard, *Generalized Kuhn-Tucker Conditions for Mathematical Programming Problems in a Banach Space*, SIAM J. on Control, **7**(1969), 232–241.
2. P.P. Varaiya, *Nonlinear Programming in a Banach Space*, SIAM J. on

Appl. Math., **15**(1967), 284–293.

3. H.W. Kuhn and A.W. Tucker, *Nonlinear Programming*, Proc. Second Berkeley Symposium on Mathematical Statistics and Probability, J. Neyman, ed., University of California Press, Berkeley (1951), 481–492.
4. K.J. Arrow, L. Hurwicz and H. Uzawa, *Constraint Qualification in Maximization Problems*, Naval Res. Logist. Quart., **8**(1961), 175–191.
5. O.L., Mangasarian, *Pseudo-convex functions*, SIAM J. on Control, **3**(1965), 281–290.
6. K.J. Arrow and A.C. Enthoven, *Quasi-concave Programming*, Econometrica, **29**(1961), 779–800.
7. J. Ponstein, *Seven Kinds of Convexity*, SIAM Review, **9**(1967), 115–119.
8. J.P. Aubin, *Applied Functional Analysis*, A Wiley-Interscience Pub., John Wiley & Sons, New York, 1979.

Sogang University
Seoul 121, Korea