

A PSEUDOCONVEX MINIMIZATION PROBLEM

BYUNG HO YOON

1. Preliminaries

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 of [1] and [4] to obtain a necessary and sufficient optimality condition.

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 \bar{M} and (M) denotes the convex hull of M ; that is, the smallest convex set in U containing M .

DEFINITION 1.1. 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 1.2. Let C be a cone in U . Subsets of U^* defined by

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

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

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 1.3. 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_0 and a sequence $\{\lambda_k\}$

of non-negative real numbers such that the sequence $\{\lambda_k(u_k - u_0)\}$ converges to v .

DEFINITION 1.4. The set $T(M, u_0)$ of all the vectors tangent to M at u_0 is called the cone tangent to M at u_0 .

DEFINITION 1.5. The closure of the convex hull of $T(M, u_0)$ is called the cone pseudotangent to M at u_0 and is denoted by $P(M, u_0)$.

Note that $T(M, u_0)$ is a cone and $P(M, u_0)$ is a closed convex cone.

DEFINITION 1.6. Let M be a subset of U and $u_0 \in M$. If, for every $u \in M$, $u - u_0$ belongs to $P(M, u_0)$, then we say M is pseudoconvex at u_0 . If M is pseudoconvex at all of its points, then we simply say that M is pseudoconvex.

Following propositions are well known [1].

PROPOSITION 1.1. If all $M_i, i \in I$, are pseudoconvex at u_0 , then $\tilde{M} = \bigcup_{i \in I} M_i$ is pseudoconvex at u_0 .

PROPOSITION 1.2. If a subset M of U is convex, then M is pseudoconvex at every u_0 in M .

DEFINITION 1.7. 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 $\{u \in U | \phi(u) \leq \lambda\}$ 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 1.3. If ϕ is quasi-convex and Fréchet-differentiable at $\bar{u} \in U$, then $(\Delta\phi(\bar{u}), u - \bar{u}) > 0$ implies $\phi(u) - \phi(\bar{u}) > 0$.

The Proof of this proposition can be found in [5].

DEFINITION 1.8. 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 [5] that convex Fréchet-differentiable functions

are pseudoconvex and pseudoconvex functions are quasi-convex.

2. 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 1 (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
- (iii) a function $\phi : U \rightarrow E^1$ which is pseudoconvex over V
- (iv) a function $\phi : X \rightarrow E^1$ which is convex differentiable and increasing in $\|x\|$.

And we wish to consider the problem of minimizing the function

$$(2.1) \quad F(x, u) = \phi(x) + \phi(u)$$

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

Following well known facts which can be found in [6] are needed in the discussion of our problem.

DEFINITION 2.1. 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 2.1. *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 $P : X \rightarrow M$ and $Q : X \rightarrow N$ 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, $S_1 : N \rightarrow U$ 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

$$S_1 Qx = SQ^2x = SQx = Sx = u.$$

Thus $S_1 : N \rightarrow U$ 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 $J : X \rightarrow X^*$ 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 2.1 satisfies the following.

PROPOSITION 2.2. *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 [6].

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 $\text{Ker } A$. For each $u \in U$, 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 (2.1) over $u \in V$ and $Ax = u$ is equivalent to the problem of minimizing

$$(2.2) \quad f(u) = \phi(A_+u) + \psi(u)$$

over $u \in V$.

Notice that the objective function $f(u)$ of (2.2) is pseudoconvex over V and the derivative is given by

$$\nabla f(u) = A_+^* \nabla \phi(A_+u) + \nabla \psi(u).$$

THEOREM 2.1. *A necessary and sufficient condition for $u_0 \in V$ to minimize $f(u)$ of (2.2) over V is that $\nabla f(u_0) \in P^+(V, u_0)$.*

Proof. Suppose $u_0 \in V$ minimizes $f(u)$ over V . Let $u \in T(V, u_0)$. Then there exists a sequence $\{u_k\}$ in V converging to u_0 and a sequence $\{\lambda_k\}$, $\lambda_k > 0$ for all k , such that

$$\lim_k \lambda_k (u_k - u_0) = u.$$

Since u_0 minimizes $f(u)$ over V , $f(u_k) - f(u_0) \geq 0$ for all k . Moreover,

$$f(u_k) - f(u_0) = (\nabla f(u_0), u_k - u_0) + o(\|u_k - u_0\|).$$

Thus

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

Letting k go infinite, we have

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

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 $\langle \nabla f(u_0), u - u_0 \rangle \geq 0$. 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 2.1, we consider the following example.

EXAMPLE. Suppose' in Prob. (PC),

$$(2.3) \quad f(u) = \frac{1}{2} \|A_+ u\|^2 + \frac{1}{2} \|u\|^2$$

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 $K : U \rightarrow U^*$ be the duality operator. Then

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

Thus, by Theorem 2.1, if u_0 minimizes $f(u)$ of (2.3) 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 (2.3), then

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

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. O.L. Mangasarian, *Pseudo-convex functions*, SIAM J. on Control, **3** (1965), 281-290.
4. K.J. Arrow and A.C. Enthoven, *Quasi-concave Programming*, Econometrica, **29** (1961), 779-800.
5. J. Ponstein, *Seven Kinds of Convexity*, SIAM Review, **9** (1967), 115-119.
6. J.P. Aubin, *Applied Functional Analysis*, A Wiley-Interscience Pub., John Wiley & Sons, New York, 1979.

Sogang University
Seoul 121-742, Korea