JOURNAL OF APPROXIMATION THEORY 73, 119–135 (1993)

Best Interpolation with Convex Constraints

KANG ZHAO*

Department of Mathematics,
University of Wisconsin-Madison, Madison, Wisconsin 53706, U.S.A.

Communicated by Frank Deutsch

Received April 12, 1991; accepted in revised form December 31, 1991

A characterization of any solution to the minimization problem

$$\min\{\|x-z\| : x \in K := C \cap A^{-1}d\}$$

is given, where A is a continuous linear map from a real Banach space X to a locally convex topological space Y, $z \in X$, $C \subset X$ is a closed convex set and $d \in AC$. The resulting characterization for the case that X is a Hilbert space is that the projection $P_K(z)$ of z to K is $P_C(z_0+z)$ for some $z_0 \in \operatorname{ran} A^*$ provided $d \in \operatorname{int} AC$. An analogous characterization is also obtained for the solution to the nonnegative best interpolation problem in the L_p norm.

1. INTRODUCTION

Let X be a real Banach space, Y a locally convex topological linear space, and $A: X \to Y$ a continuous linear map. For a closed convex set $C \subset X$, $z \in X$, and $d \in Y$, we are interested in characterizing the solutions to the following constrained best interpolation problem

$$\min\{\|x-z\| : x \in K := C \cap A^{-1}d\}. \tag{1.1}$$

Any solution x_0 to (1.1) is also called a projection of z to the closed convex set K and is denoted by $P_K(z)$ if it is unique.

In the particular case $X = L_p(\Omega)$ with $1 < p < \infty$, $Y = \mathbb{R}^n$, $z = 0$, and C the cone consisting of all nonnegative functions in $L_p(\Omega)$, [M85] has shown that the solution to (1.1) is of the form

$$x_0 = (z_0^*)_+^{1/(p-1)}, \tag{1.2}$$

for any $z_0^* \in \operatorname{ran} A^*$ such that $(z_0^*)_+^{1/(p-1)} \in A^{-1}d$, provided $d \in \operatorname{int} AC$. The

* Supported by a research assistantship from the National Science Foundation under Grant DMS-9000053.

0021-9045/93 $5.00

representation of the solution obtained in [M85] for the case that $d \notin \operatorname{int} AC$ is

$$x_0 = (z_0^*)_+^{1/(p-1)} \chi_{\Omega_0}, \tag{1.3}$$

for some $\Omega_0 \subset \Omega$, under the assumption that $L_p(\Omega)$ be separable. The importance of (1.2) or (1.3) is that the infinite-dimensional problem (1.1) is converted to a nonlinear problem of finite dimension in this case. Unfortunately, the subdomain Ω_0 in (1.3) is not specified explicitly. A further question is whether (1.2) and (1.3) hold when $\dim Y = \infty$.

In case X and Y are Hilbert spaces and C is a closed convex cone with vertex 0, (1.1) has been investigated intensively in [M88, C89, C90]. When $Y = \mathbb{R}^n$ and $z = 0$, it is proved in [M88] that, under the condition

$$\{A^*y : (y, d) \geq 0\} \cap C^0 = \{0\}, \tag{1.4}$$

the solution to (1.1), i.e., the projection $P_K(0)$ to K, can be represented by

$$P_K(0) = P_{C^\circ}(z_0), \tag{1.5}$$

for any $z_0 \in \operatorname{ran} A^*$ satisfying $P_{C^\circ}(z_0) \in A^{-1}d$. In this case, as proved in [C90], the condition (1.4) is equivalent to the following so-called Slater condition [B78, p. 159]

$$d \in \operatorname{int} AC. \tag{1.6}$$

We note that in general the condition $A^{-1}d \cap \operatorname{int} C \neq \emptyset$ used in [M88] is stronger than (1.6). In an effort to generalize the formula (1.5) to the case $z \neq 0$ and/or $\dim Y = \infty$, [C90] introduces the following condition, called CHIP (conical hull intersection property),

$$\overline{S(K; x_0)} = \overline{S(C; x_0)} \cap \ker A, \tag{1.7}$$

where $S(K; x_0)$ is the translation of the support cone for K at x_0, i.e.,

$$S(K; x_0) := \bigcup_{\lambda \geq 0} \lambda(K - x_0).$$

Under the CHIP condition (1.7), for the cone-constrained best interpolation problem (1.1) in Hilbert spaces, the characterization of the projection $P_K(z)$ to K obtained in [C90] is

$$z - P_K(z) \in \overline{C^0 \cap (P_K(z))^\perp + \operatorname{ran} A^*}. \tag{1.8}$$

Condition (1.8) is a generalization of (1.5) since, as proved in [C90], the CHIP condition (1.7) is satisfied as long as the Slater condition (1.6) holds.

As a continuation of [C90], [C89] proves that $C^0 \cap P_K(z))^{\perp} + \operatorname{ran} A^*$ is closed if the Slater condition (1.6) is satisfied, even for the case that C is not just a cone. Therefore, in this case, their characterization for the projection is

$$P_K(z) = P_C(z + z_0) \tag{1.9}$$

for any $z_0 \in \operatorname{ran} A^*$ such that $P_C(z + z_0) \in A^{-1}d$.

In this article, we tackle the problem (1.1) in a unified way, i.e., we consider the problem in the case that X is any real Banach space and Y is any locally convex linear topological space. In particular, the characterizations obtained in [M85, M88, C89, C90] are special cases of the results presented in this article. Moreover, we generalize the results in [M85] to the case $\dim Y = \infty$. The improved result for (1.3) obtained without the condition that $L_p(\Omega)$ be separable in this article is that

$$x_0 = (z_0^*)_+^{1/(p-1)} \tag{1.10}$$

for any $z_0^* \in \operatorname{ran} A_1^*$ such that $(z_0^*)_+^{1/(p-1)} \in A^{-1}d$, where A_1 is the restriction of A to the smallest closed subspace X_0 containing K. In contrast to the approaches taken in [M85, M88, C89, C90], we derive (1.2) and (1.9) from the Kuhn–Tucker minimization condition for (1.1). The unified treatment enables us to show that the Slater condition (1.6) is necessary for the stability of the minimization problem (1.1). This implies that any attempt to get rid of the Slater condition (1.6) is superfluous if the space X is appropriate for (1.1).

This article is organized as follows. In Section 2, we recall some basic notations concerning convex analysis. Section 3 is the main part of this article. It contains a characterization of the solutions to the general minimization problem

$$\min\{f(x) : x \in K := C \cap A^{-1}d\} \tag{1.11}$$

for a nonnegative convex function f. The fourth section is devoted to projection problems in Hilbert spaces. In Section 5, we apply the results obtained in the previous sections to the specific nonnegative best interpolation problems in the L_p norm.

Throughout this article we use the following conventions.

The kernel of a continuous linear map A is denoted by $\ker A$, $\operatorname{ran} A$ is the range of A, and $A^{-1}d$ is the preimage of $d \in Y$. We denote by A^* the adjoint of A. As usual, we denote by X^* the dual space of X and write (x^*, x) for the value $x^*(x)$ of $x^* \in X^*$ at $x \in X$, $x^{\perp} := \{x^* \in X^* : (x^*, x) = 0\}$. For $f: X \to \mathbb{R} \cup \{\infty\}$, the set $\operatorname{epi} f := \{[x, r] \in X \times \mathbb{R} : x \in X,$

$r \geqslant f(x)\}$ is the epigraph of f. The interior of $B \subset X$ and the closure of B are denoted by int B and $\bar{B}$, respectively. $\mathbb{R}_+$ stands for the half line of all nonnegative numbers.

2. Preliminaries from Convex Analysis

In this section we recall some related definitions from convex analysis.

Let X be a real locally convex linear topological space. The *subdifferential* of a convex function $f: X \to \mathbb{R} \cup \{\infty\}$ at x is defined by

$$\partial f(x) := \{x^* \in X^* : x^* - (x^*, x) \leqslant f - f(x)\}. \tag{2.1}$$

Hence, x solves the following unconstrained minimization problem

$$\min\{f(u) : u \in X\}$$

iff $0 \in \partial f(x)$. f is called *subdifferentiable* at x if $\partial f(x) \neq \emptyset$. A function f is called *Gateaux differentiable* at x if there exists $x^* \in X^*$ such that

$$\lim_{t \to 0+} (f(x+ty) - f(x))/t = (x^*, y), \qquad \forall y \in X. \tag{2.2}$$

The x^* in (2.2) is called the *Gateaux derivative* of f at x and is denoted by $\nabla f(x)$. In particular, when f is Gateaux differentiable at x, we have

$$\partial f(x) = \{\nabla f(x)\}.$$

For any $B \subset X$, denote by B^0 the *polar* of B, i.e., the set defined by

$$B^0 := \{x^* \in X^* : (x^*, x) \leqslant 1, \ \forall x \in B\}. \tag{2.3}$$

It can be veried that B^0 is a w^*-closed convex set in X^*. The *bipolar* of B is

$$B^{00} := \{x \in X : (x^*, x) \leqslant 1, \ \forall x^* \in B^0\}.$$

When $B \subset X$ is a cone with vertex 0, the polar B^0 is called the *dual cone* of B. It can also be characterized by

$$B^0 := \{x^* \in X^* : (x^*, x) \leqslant 0, \ \forall x \in B\}. \tag{2.4}$$

For any $B \subset X$, denote by $S(B; x)$ the translation of the *support* cone of B at x, that is,

$$S(B; x) := \bigcup_{\lambda \geqslant 0} \lambda(B - x). \tag{2.5}$$

In particular, $S(B; x) = B + \text{span}(x)$ if B is a cone with vertex 0 and $x \in B$. It can be verified that $S(B; x)$ is convex if B is convex. The dual cone of $S(B; x)$ is denoted by $S^0(B; x)$. Therefore, $S^0(B; x) = B^0 \cap x^\perp$ if B is a cone with vertex 0 and $x \in B$.

For any $B \subset X$, the *indicator function* I_B is defined by

$$I_B(x) := \begin{cases} 0, & \text{if } x \in B; \\ \infty, & \text{otherwise}. \end{cases} \tag{2.6}$$

Hence, I_B is convex and lower semicontinuous if B is a closed convex set.

For a convex function f, $\text{dom} f := \{x \in X : f(x) \text{ is finite}\}$. If $\text{dom} f$ is not empty, then f is called a *proper* convex function.

3. Kuhn–Tucker Minimization Condition

Suppose X and Y are two locally convex linear topological spaces and X is real. Let $A: X \to Y$ be a continuous linear map and $f: X \to \mathbb{R}_+ \cup \{\infty\}$ a proper lower semicontinuous convex function. For a fixed $d \in Y$ and a closed convex set $C \subset X$, we consider the minimization problem

$$m := \min\{f(x) : x \in K := C \cap A^{-1}d\}. \tag{3.1}$$

In this section, we give a necessary and sufficient condition for the existence of *Lagrange multipliers* for (3.1). Here we call $z^* \in Y^*$ a Lagrange multiplier for (3.1) if

$$(z^*, Ax - d) \leqslant f(x) - m, \qquad \forall x \in C \tag{3.2}$$

In case A is an open map or $Y = \mathbb{R}^n$, we show that one sufficient condition for the existence of Lagrange multipliers is the Slater condition

$$d \in \text{int}\, AC. \tag{3.3}$$

When C is a cone with vertex 0 and $d \neq 0$, it is proved later that, for the modified problem

$$\min\{f_|(x) : x \in C_0 \cap A_|^{-1}d\}, \tag{3.4}$$

the Slater condition (3.3) is satisfied automatically, where $C_0 := C \cap X_0$, X_0 is the smallest closed subspace containing $C \cap A^{-1}d$, and $A_|$, $f_|$ are the restrictions of A, f to X_0, respectively.

As usual, a map A is called open if it maps open sets to open sets.

PROPOSITION 3.1. *For the minimization problem* (3.1), *the function*

$$g_f : Y \to \mathbb{R}_+ \cup \{\infty\} : y \mapsto \inf f(C \cap A^{-1}y) \tag{3.5}$$

is convex.

Proof. For $y_1, y_2 \in \text{dom}(g_f)$, $\delta > 0$, there exist $x_1, x_2 \in \text{dom} f$ such that

$$g_f(y_i) \leqslant f(x_i) < g_f(y_i) + \delta, \qquad x_i \in C \cap A^{-1}y_i,\ i = 1, 2.$$

For any $0 < t < 1$, $x_t := tx_1 + (1-t)\,x_2 \in C \cap A^{-1}y_t$ with $y_t := Ax_t$. Since f is convex,

$$f(x_t) \leqslant tf(x_1) + (1-t)\,f(x_2) \leqslant tg_f(y_1) + (1-t)\,g_f(y_2) + \delta.$$

Therefore, $g_f(ty_1 + (1-t)\,y_2) \leqslant tg_f(y_1) + (1-t)\,g_f(y_2)$ and g_f is convex. ∎

As we see in the above proof, the convexity of g_f has nothing to do with the semicontinuity of f.

PROPOSITION 3.2. *If* $\text{int}\, A(\text{dom} f \cap C) \neq \emptyset$, *then the* g_f *defined by* (3.5) *is continuous* on $\text{int dom}\, g_f$ *under either one of the following conditions*:

(a) $Y = \mathbb{R}^n$;

(b) *A is an open map.*

In the proof of Proposition 3.2 we need the following lemmas.

LEMMA 3.1. *Suppose g is a proper convex function. If there exists a nonempty open set on which g is bounded above, then g is continuous on* int dom *g*.

Proof. See [H75, p. 82]. ∎

LEMMA 3.2. *Suppose* $g: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *is a convex function. Then g is continuous on* int dom *g*.

Proof. See [H75, p. 84]. ∎

Proof of Proposition 3.2. Since there exists $\bar{d} \in \text{int}\, A(\text{dom} f \cap C)$, we know from the definition of g_f that $\bar{d} \in \text{int dom}\, g_f$. Therefore, by Lemma 3.2, g_f is continuous on int dom g_f if $Y = \mathbb{R}^n$.

When A is open, as assumed, there exists $\bar{x} \in C \cap A^{-1}\bar{d} \cap \text{dom} f$, since f is lower semicontinuous,

$$N(\bar{x}) := (A^{-1}\, \text{int}\, AC) \cap \{x : f(x) < f(\bar{x}) + 1\}$$

is a nonempty open set on which f is bounded above.

Therefore g_f is bounded above on the open set $AN(\bar{x})$ containing $\bar{d}$. Hence, g_f is continuous on int dom g_f by Lemma 3.1. ∎

Remarks. (1) The continuity of g_f on int dom g_f does not depend on the semicontinuity of f if $Y = \mathbb{R}^n$.

(2) When X is a real Banach space and Y is of second category, the condition in Proposition 3.2 that A is an open map is satisfied automatically since int $AC \neq \varnothing$ implies ran $A = Y$. The assumption $f \geqslant 0$ is only used to prove that g_f is finite on AC. For the constrained best interpolation problem (1.1), $f := \|\cdot - z\|$, so the condition that $f \geqslant 0$ is always satisfied.

From the definition of g_f we know that the continuity of g_f indicates the stability of the minimization problem (3.1). So it is natural to request the continuity of g_f at d and $d \in \operatorname{int} AC$. Proposition 3.2 provides sufficient conditions for g_f to be continuous on int dom g_f.

THEOREM 3.1. *Suppose x_0 is a solution to* (3.1). *Then, there exists $z^* \in (AX)^*$ such that*

$$(z^*, A(x - x_0)) \leqslant f(x) - f(x_0), \qquad \forall x \in C \tag{3.6}$$

iff the corresponding g_f is subdifferentiable at d.

Proof. Assume that (3.6) holds, set $y := Ax$. Then

$$(z^*, y - d) \leqslant f - f(x_0), \qquad \text{on} \quad C \cap A^{-1}y.$$

Since $f(x_0) = g_f(d)$, it follows that $(z^*, y - d) \leqslant g_f(y) - g_f(d)$ for all $y \in AC$, while this inequality always holds for $y \notin AC$ since for such y, $g_f(y) = \infty$ by definition. Consequently, $z^* \in \partial g_f(d)$.

On the other hand, if $z^* \in \partial g_f(d)$, then

$$(z^*, y - d) \leqslant g_f(y) - g_f(d).$$

For every $x \in C \cap A^{-1}y$,

$$(z^*, A(x - x_0)) \leqslant g_f(y) - g_f(d).$$

Note $f(x) \geqslant g_f(y)$ for $x \in C \cap A^{-1}y$ and $g_f(d) = f(x_0)$, so (3.6) holds. ∎

As a consequence of Proposition 3.1, Proposition 3.2, and Theorem 3.1, we have the following theorem. The result for the case $Y = \mathbb{R}^n$ appears already in [N70].

THEOREM 3.2. *Suppose $d \in \operatorname{int} A(\operatorname{dom} f \cap C)$ and either one of the following conditions holds*

(1) $Y = \mathbb{R}^n$;

(2) *A is open.*

Then $x_0 \in K := C \cap A^{-1}d$ *is a solution to* (3.1) *iff there exists* $z_0^* \in Y^*$ *such that*

$$(z_0^*, A(x - x_0)) \leq f(x) - f(x_0) \tag{3.7}$$

for all $x \in C$.

Proof. It is clear that (3.7) implies that x_0 is a solution to (3.1).

Conversely, according to Proposition 3.1 and Proposition 3.2, the convex function g_f is continuous at d. Therefore, g_f is subdifferentiable at d (cf. [H75, p. 84]). Thus the proof is completed by applying Theorem 3.1. ∎

Consequently, we obtain the Kuhn–Tucker condition to characterize the solutions to (3.1).

THEOREM 3.3. *Assume the conditions of Theorem* 3.2. *Then* $x_0 \in K := C \cap A^{-1}d$ *is a solution to* (3.1) *iff there exist* $y_0^* \in S^0(C, x_0)$ *and* $z_0^* \in Y^*$ *such that either*

$$A^*z_0^* - y_0^* \in \partial f(x_0) \tag{3.8}$$

if f *is continuous at some point in* C *or*

$$\nabla f(x_0) = A^*z_0^* - y_0^* \tag{3.9}$$

if f *is Gateaux differentiable at* x_0.

Proof. As an immediate consequence of Theorem 3.2, $x_0 \in K$ is a solution to (3.1) iff it is a solution to

$$\min\{f(x) - (A^*z_0^*, x) + I_C(x) : x \in X\}$$

for some $z_0^* \in Y^*$. Here I_C is the indicator function for C. Therefore, $x_0 \in K$ is a solution to (3.1) iff

$$0 \in \partial(f - A^*z_0^* + I_C)(x_0).$$

If f is continuous at some point in C, since $\partial I_C(x_0) = S^0(C; x_0)$, it follows (cf. [H70, p. 25]) that

$$\partial(f - A^*z_0^* + I_C)(x_0) = \partial f(x_0) - A^*z_0^* + S^0(C; x_0).$$

If f is Gateaux differentiable at x_0, by (3.7), for any $x \in C$ and any $0 < t < 1$,

$$(z_0^*, A(x - x_0)) \leq (f(x_0 + t(x - x_0)) - f(x_0))/t,$$

since $x_0+t(x-x_0)\in C$. Let t go to 0, we obtain that

$$(z_0^*, A(x-x_0)) \leqslant (\nabla f(x_0), x-x_0), \qquad \forall x\in C.$$

This implies that $A^*z_0^* - \nabla f(x_0) \in S^0(C; x_0)$. Therefore, in this case we have (3.9). ▮

Remark. When $Y=\mathbb{R}^n$, as is pointed out in the remarks for Proposition 3.2, g_f is continuous on $\operatorname{int}\operatorname{dom} g_f$ without the assumption that f is lower semicontinuous. Therefore, in this case, we still have (3.9) if f is Gateaux differentiable at x_0.

COROLLARY 3.1. *Suppose X is a real Banach space and* ran A *is a locally convex topological linear space of second category. If d is an interior point of AC relative to ran A, then $x_0\in K$ is a solution to* (1.1) *iff there exist $y_0^*\in S^0(C; x_0)$ and $z_0^*\in(\operatorname{ran} A)^*$ such that* (3.8) *holds with $f:=\|\cdot - z\|$. Furthermore, if $\|\cdot - z\|$ is Gateaux differentiable at x_0, then* (3.8) *becomes* (3.9).

So far, we have proved that for the minimization problem (3.1) with an open map A or with $Y=\mathbb{R}^n$, the Lagrange multiplier exists when the Slater condition (3.3) is satisfied. The following theorem shows that the Slater condition (3.3) is satisfied automatically if C is a convex cone with vertex 0, $d\neq 0$, and X is the smallest complete space for the minimization problem (3.1). In the rest of this section, the convex set C in (3.1) is assumed to be a closed convex cone with vertex 0, X_0 is the smallest closed subspace containing $K:=C\cap A^{-1}d$, and $C_0:=C\cap X_0$. We continue to denote $A_|: X_0\to AX_0$ and $f_|$, respectively, the restrictions of A and f to X_0.

THEOREM 3.4. *Suppose in* (3.1) *that C is a closed convex cone with vertex* 0 *and $d\neq 0$. Then $x_0\in K:=C\cap A^{-1}d$ is a solution to* (3.1) *iff there exists $z_0^*\in(A_| X_0)^*$ such that*

$$(A_|^* z_0^*, x-x_0) \leqslant f_|(x_0) - f_|(x_0), \qquad \forall x\in C_0. \tag{3.10}$$

Proof. By its definition, X_0 is the closure of $\operatorname{span}(K)$. Since $A\operatorname{span}(K)=\operatorname{span}(d)$, it follows that $A_| X_0=\operatorname{span}(d)$ and $A_| C_0=\{td : t\geqslant 0\}$. It is clear that $d\in\operatorname{int} A_| C_0$ because $d\neq 0$. We complete the proof by Theorem 3.2. ▮

As immediate results of Theorem 3.3 and Theorem 3.4, we have

THEOREM 3.5. *Assume the conditions of Theorem* 3.4. *Then $x_0\in K:=C\cap A^{-1}d$ is a solution to* (3.1) *iff there exist $y_0^*\in C_0^0\cap x_0^\perp$ and $z_0^*\in(A_| X_0)^*$ such that either*

$$A_|^* z_0^* - y_0^* \in \partial f_|(x_0) \tag{3.11}$$

if $f_|$ is continuous at some point in the cone C_0 or

$$\nabla f_|(x_0) = A_|^* z_0^* - y_0^* \tag{3.12}$$

if $f_|$ is Gateaux differentiable at x_0.

COROLLARY 3.2. *Suppose X is a real Banach space and Y is a locally convex topological linear space. Then, $x_0 \in K$ is a solution to* (3.1) *iff there exist $y_0^* \in C_0^0 \cap x_0^\perp$ and $z_0^* \in (A_| X_0)^*$ such that* (3.11) *holds with $f := \|\cdot - z\|$. If $f_|$, the restriction of $\|\cdot - z\|$, is Gateaux differentiable at x_0, then* (3.11) *becomes* (3.12).

When (3.3) is not satisfied, as is proved in the following, for any solution x_0, there exists $z^* \in X^*$ such that

$$z^* - (z^*, x_0) \leqslant f - f(x_0) \tag{3.13}$$

and

$$(z^*, A(x - x_0)) \leqslant (z^*, x - x_0), \qquad \forall x \in C_0. \tag{3.14}$$

THEOREM 3.6. *Suppose X_0 is a subspace of X and $C \subset X$ is a convex cone with vertex 0. Let $C_0 := C \cap X_0$. If $z_0^* \in (AX_0)^*$ is such that $A^* z_0^* \leqslant f$ on C_0 and $f(x_0) = (z_0^*, Ax_0)$ for some $x_0 \in C_0 \cap \operatorname{int} \operatorname{dom} f$, then there exists $z^* \in X^*$ such that* (3.13) *and* (3.14) *hold.*

Proof. Let

$$B := \{[x, (z_0^*, Ax)] \in X \times \mathbb{R} : x \in C_0\}.$$

So B is convex and $B \cap \operatorname{int} \operatorname{epi}(f) = \varnothing$. Since $\operatorname{int} \operatorname{dom} f \neq \varnothing$ and f is lower semicontinuous, we have that $\operatorname{int} \operatorname{epi}(f) \neq \varnothing$. Therefore there exist $y^* \in X^*$ and $\lambda \in \mathbb{R}$, not all zero, such that

$$(y^*, x) + \lambda r \geqslant (y^*, u) + \lambda (z_0^*, Au) \tag{3.15}$$

for all $[x, r] \in \operatorname{epi}(f)$ and $u \in C_0$.

If, in (3.15), we take $x = u = x_0 \in C_0$, then we obtain $\lambda(r - (z_0^*, Ax_0)) \geqslant 0$ for all $r > f(x_0)$. So $\lambda \geqslant 0$. If $\lambda = 0$, then

$$(y^*, x - x_0) \geqslant 0, \qquad \forall x \in \operatorname{dom} f,$$

since $x_0 \in \operatorname{int} \operatorname{dom} f$, this implies $y^* = 0$ and we get a contradiction.

Therefore, $\lambda > 0$. Without loss of generality, assume $\lambda = 1$. Thus we have

$$(y^*, x) + f(x) \geqslant (y^*, u) + (z_0^*, Au), \qquad \forall x \in X, u \in C_0, \tag{3.16}$$

and $(-y^*, x - x_0) \leqslant f(x) - f(x_0)$ for all $x \in X$.

If we take $x = x_0$ in (3.16), we obtain

$$(z_0^*, A(u - x_0)) \leqslant (-y^*, u - x_0), \qquad \forall u \in C_0.$$

With the choice $z^* := -y^*$, we obtain (3.13) and (3.14). ∎

THEOREM 3.7. *Let $X_0 := A^{-1}\,\mathrm{span}(d)$. Then $x_0 \in C \cap A^{-1}d \cap \mathrm{int\,dom}\, f$ is a solution to* (3.1) *iff there exists $z^* \in X^*$ such that*

$$(z^*, x - x_0) \leqslant f(x) - f(x_0), \qquad \forall x \in X, \tag{3.17}$$

and $(z^*, x - x_0) \geqslant 0$ on $C \cap X_0$.

Proof. As proved in Theorem 3.4, there exists $z_0^* \in (AX_0)^*$ such that $A^* z_0^* - (z_0^*, d) \leqslant f - f(x_0)$ on $C \cap X_0$. The proof is completed by applying Theorem 3.6 to z_0^* and $f - f(x_0) + (z_0^*, d)$. ∎

4. Projection Problem in Hilbert Spaces

Let X be a Hilbert space and $K \subset X$ a closed convex set. For $z \in X$, $P_K(z) \in K$ is called the projection of z to K if

$$\|z - P_K(z)\| \leqslant \|z - x\|, \qquad \forall x \in K.$$

It is well known that

$$x_0 = P_K(z) \Leftrightarrow (x - x_0, z - x_0) \leqslant 0, \qquad \forall x \in K. \tag{4.1}$$

Equivalently, x_0 is the projection of z to K iff there exists $y_0 \in S^0(K; x_0)$ such that $z = x_0 + y_0$.

In this section, we consider the special case where the convex set K is the intersection of a closed convex set with a linear manifold. Suppose Y is a locally convex linear topological space and $A: X \to Y$ is a continuous linear map. For $d \in AC$ and a closed convex set $C \subset X$, denote by K the closed convex set $C \cap A^{-1}d$. The goal of this section is to derive the following more informative formula

$$P_K(z) = P_C(z_0 + z) \tag{4.2}$$

for any $z_0 \in \mathrm{ran}\, A^*$ satisfying

$$P_C(z_0 + z) \in A^{-1}d. \tag{4.3}$$

Actually, one only needs to prove that there exists $z_0 \in \mathrm{ran}\, A^*$ satisfying (4.2) since (4.3) implies $P_C(z_0 + z) = P_K(z)$ as is verified in the following.

Let $x_1 := P_C(z_0 + z) \in A^{-1}d$ with $z_0 \in \operatorname{ran} A^*$. Then for every $x \in C$,

$$(z_0 + z - x_1, x - x_1) \leqslant 0.$$

Especially, for $x \in K$, since $x - x_1 \in \ker A$ and $(z_0, x - x_1) = 0$, we have

$$(z - x_1, x - x_1) \leqslant 0.$$

This means that $x_1 = P_K(z)$ by (4.1).

THEOREM 4.1. *Suppose X is a real Hilbert space, Y is a locally convex linear topological space of second category, $A: X \to Y$ is a continuous linear map and $C \subset X$ is a closed convex set. If d is an interior point of AC, then $x_0 = P_K(z)$ iff there exists $z_0 \in \operatorname{ran} A^*$ satisfying $P_C(z + z_0) \in A^{-1}d$ such that*

$$x_0 = P_C(z_0 + z). \tag{4.4}$$

Proof. If $z \in K$, then let $z_0 = 0$ in (4.4). Otherwise, let $f := \|\cdot - z\|$, then $\nabla f(x_0) = (x_0 - z)/f(x_0)$. From Corollary 3.1 we know that $x_0 = P_K(z)$ iff there exist $z_0 \in \operatorname{ran} A^*$ and $y_0 \in S^0(C; x_0)$ such that $x_0 - z = z_0 - y_0$ because $f(x_0) > 0$. With this we obtain (4.4) from (4.1). ∎

As mentioned in Section 1, when Y is a Hilbert space, the conclusion corresponding to Theorem 4.1 has been reached in [C89].

THEOREM 4.2. *Suppose X_0 is the smallest closed subspace of X containing $K := C \cap A^{-1}d$ and $C \subset X$ is a closed convex cone with vertex 0. Denote by $A_|: X_0 \to AX_0$ the restriction of A to X_0 and $C_0 = C \cap X_0$. Then*

$$P_K(z) = P_{C_0}(z + z_0) \tag{4.5}$$

for any $z_0 \in \operatorname{ran} A_|^*$ such that $P_{C_0}(z + z_0) \in A^{-1}d$.

Proof. If $z \in K$, let $z_0 = 0$ in (4.5).

For $z \notin K$, if $d = 0$, then $K = C \cap \ker A$ is also a closed convex cone with vertex 0. Suppose $z = P_{X_0}(z) + z_1$ and $P_{X_0} = P_{C_0} P_{X_0}(z) + z_2$ for some $z_1 \in X_0^\perp$ and $z_2 \in C_0^0 \cap (P_{C_0} P_{X_0}(z))^\perp$. Then we have

$$z = P_{C_0} P_{X_0}(z) + (z_1 + z_2). \tag{4.6}$$

Since $z_1 + z_2 \in C_0^0 \cap (P_{C_0} P_{X_0}(z))^\perp$, the equality $P_K(z) = P_{C_0} P_{X_0}(z)$ follows from (4.1). So we only need to choose $z_0 = 0$ since $P_{C_0} P_{X_0} = P_{C_0}$.

For $d \neq 0$, take $f := \|\cdot - z\|$ in Corollary 3.2. Suppose $z = z_1 + z_2$ for some $z_1 \in X_0$ and $z_2 \in X_0^\perp$. Then the restriction of f to X_0 is $f_| = (\|\cdot - z_1\|^2 + \|z_2\|^2)^{1/2}$. So the Gateaux derivative for $f_|$ at x_0 is

$(x_0 - z_1)/f_|(x_0)$. From Corollary 3.2 it follows that $x_0 = P_K(z)$ iff there exist $y_0 \in C_0^0 \cap x_0^\perp$ and $z_0 \in \operatorname{ran} A_|^*$ such that

$$x_0 - z_1 = z_0 - y_0$$

because $f_|(x_0) > 0$. By (4.1) we obtain

$$P_K(z) = P_{C_0}(P_{X_0}(z) + z_0) = P_{C_0}(z + z_0),$$

since P_{X_0} is linear and $P_{C_0} P_{X_0} = P_{C_0}$.

Now suppose $Y = \mathbb{R}^n$ and $Ax := \sum_{i=1}^n (\Phi_i, x)\, e_i$ for $\Phi_i \in X$, where e_i is the ith unit vector for $\mathbb{R}^n$. Then $\operatorname{ran} A^* = \operatorname{span}(\Phi_i)$. Theorem 4.2 implies

THEOREM 4.3. *Suppose X is a Hilbert space and $A: X \to \mathbb{R}^n$ is defined by $Ax := \sum_{i=1}^n (\Phi_i, x)\, e_i$. Assume C is a closed convex cone with vertex 0 and $d \in \mathbb{R}^n$. Let $K := C \cap A^{-1}d$. Then for $z \in X$,*

$$P_K(z) = P_{C_0}\left(z + \sum_{i=1}^n c(i)\, \Phi_i\right), \tag{4.7}$$

for any $c \in \mathbb{R}^n$ satisfying $P_{C_0}(z + \sum c(i)\, \Phi_i) \in A^{-1}d$. Here $X_0 = (K^\perp)^\perp$ and $C_0 = C \cap X_0$.

Proof. Since P_{X_0} is a linear map, $A_| = \sum e_i P_{X_0} \Phi_i$, by Theorem 4.2 we have (4.7). ∎

EXAMPLE. We apply Theorem 4.3 to the problem

$$\min\{ \|x - z\| : x \in C \cap A^{-1}d \}, \tag{4.8}$$

where

$C := \{x \in L_2[0, 3] : x \geqslant 0\}$,
$Ax := \sum_{i=1}^2 (\Phi_i, x)\, e_i$, $\Phi_i(t) := (1 - |t - i|)_+$, $i = 1, 2$,
$d = [1, 0] \in \mathbb{R}^2$, and $z(t) := t^2$.

Since $\Phi_2(t) > 0$ for $1 < t < 3$, it follows that $X_0 = \chi_{[0,1]} L_2[0, 3]$. From Theorem 4.3 we know that the solution x_0 has the representation

$$x_0(t) = (t^2 - ct)_+\, \chi_{[0,1]}(t)$$

because $\Phi_1(t) = t$ on $[0, 1]$. Since

$$\int_0^1 t(t^2 - ct)_+ = \begin{cases} \dfrac{1}{4} - \dfrac{c}{3}, & \text{if } c \leqslant 0; \\ 0, & \text{if } c \geqslant 1; \\ \dfrac{1}{3}(1 - c) - \dfrac{1}{12}(1 - c^4), & \text{if } 0 < c < 1, \end{cases} \tag{4.9}$$

by the interpolation condition, we obtain $c = -9/4$. So the solution is

$$x_0(t) = (t^2 + \tfrac{9}{4}t)\, \chi_{[0,1]}(t). \tag{4.10}$$

5. Nonnegative Best Interpolation in L_p Norm

In this section, we apply the results obtained in the preceding sections to the nonnegative best interpolation problem in $L_p(\Omega)$ for $1 < p < \infty$.

Suppose Y is a locally convex linear topological space, $A: L_p(\Omega) \to Y$ is a continuous linear map, and C is the cone consisting of all nonnegative functions in $L_p(\Omega)$. For a nonzero $d \in AC$, we consider the nonnegative best interpolation problem

$$\min\{\|x\| : x \in K := C \cap A^{-1}d\}. \tag{5.1}$$

Lemma 5.1 [H70]. *$f := (1/p)\, \|\cdot\|^p$ is Gateaux differentiable everywhere with the Gateaux derivative $\nabla f(x) = |x|^{p-1}\, \mathrm{sgn}(x)$.*

Now we are ready to give the representation of the solution of the problem (5.1). As an immediate consequence of Corollary 3.1 we have the following theorem.

Theorem 5.1. *Suppose $d \in \operatorname{int} AC$ and Y is of second category. Then x_0 is the solution to (5.1) iff there exists $z_0^* \in \operatorname{ran} A^*$ such that*

$$x_0 = (z_0^*)_+^{1/(p-1)} \in A^{-1}d. \tag{5.2}$$

Proof. From Corollary 3.1 it follows that $x_0 \in K$ is the solution to (5.1) iff there exists $z^* \in \operatorname{ran} A^*$ and $y^* \in C^0 \cap x_0^\perp$ such that

$$\nabla f(x_0) = z^* - y^*,$$

where $f := \|\cdot\|$. By applying Lemma 5.1 and the Chain Rule, we obtain

$$x_0^{p-1} = z_0^* - y_0^*,$$

where $z_0^* = \|x_0\|^{1-p}\, z^*$, $y_0^* = \|x_0\|^{1-p}\, y^*$.

Since $y_0^* \in C^0 \cap x_0^\perp$, it follows that $y_0^* \leqslant 0$ and $y_0^*(t) = 0$ if $x_0(t) > 0$. Therefore

$$x_0 = (z_0^*)_+^{1/(p-1)}. \quad \blacksquare \tag{5.3}$$

Let $q := p/(p-1)$. For the case $Y = \mathbb{R}^n$ and $A := \sum_{i=1}^n e_i \varphi_i$ with $\varphi_i \in L_q(\Omega)$, [M85] obtains the characterization (5.2) under the following condition: There exists $g \in K$ such that $\varphi_1, \ldots, \varphi_n$ are linearly independent on

$$G := \{t \in \Omega : g(t) > 0\}.$$

Actually, the above condition is equivalent to $d \in \operatorname{int} AC$ as proved in the following proposition.

PROPOSITION 5.1. *When $Y = \mathbb{R}^n$ and $A = \sum_{i=1}^n e_i \varphi_i$, $d \in \operatorname{int} AC$ iff there exists $g \in K$ such that $\varphi_1, \ldots, \varphi_n$ are linearly independent on*

$$G := \{t \in \Omega : g(t) > 0\}.$$

Proof. Suppose $\varphi_1, \ldots, \varphi_n$ are linearly independent on G for some $g \in K$. If $d \notin \operatorname{int} AC$, since AC is convex and $Y = \mathbb{R}^n$, there exists nonzero $b \in \mathbb{R}^n$ such that

$$(b, d) \leqslant (b, Ax), \qquad \forall x \in C.$$

Since $0 \in C$ and $d \in AC$, it follows that

$$(b, d) = 0 \leqslant (b, Ax), \qquad \forall x \in C.$$

Let $s := A^* b = \sum_{i=1}^n b_i \varphi_i = s_+ - s_-$, then $s_-^{q/p} \in C$. Therefore, we have

$$\int_\Omega s s_-^{q/p} = \int_\Omega -s_-^{(p+q)/p} \geqslant 0. \tag{5.4}$$

This shows that $s \geqslant 0$. Hence, $(s, g) = (b, d) = 0$ implies that

$$s = 0 \qquad \text{on } G.$$

This contradicts that $\varphi_1, \ldots, \varphi_n$ are linearly independent on G.

Conversely, if $d \in \operatorname{int} Ac$, then there exist $g_i \in C$, $i = 0, \ldots, n$, such that

$$\sigma := \operatorname{conv}(Ag_0, \ldots, Ag_n)$$

is a simplex such that $d \in \operatorname{int} \sigma$. Therefore, there exist $0 < t_i < 1$, $i = 0, \ldots, n$, such that

$$d = \sum_{i=0}^n t_i A g_i$$

and $\sum_{i=0}^n t_i = 1$. Choose

$$g := \sum_{i=0} t_i g_i. \tag{5.5}$$

Since $g \in C$ and $Ag = \sum t_i A g_i = d$, it follows that $g \in K$. Note $g(t) = 0$ implies that $g_i(t) = 0$ for any i. Hence, if there exists $b \in \mathbb{R}^n$ such that $s := \sum_{i=1}^n b_i \varphi_i = 0$ on G, then

$$(b, Ag_i) = (A^* b, g_i) = (s, g_i) = 0, \qquad i = 0, ..., n.$$

Since $d \in \operatorname{int} \sigma$, $Ag_1 - Ag_0, ..., Ag_n - Ag_0$ are linearly independent. Therefore, $b = 0$. Thus, we finish the proof. ∎

In general, we have the following result:

THEOREM 5.2. *Let X_0 be the smallest closed subspace containing $K := C \cap A^{-1}d$. Denote by $A_|: X_0 \to AX_0$ the restriction of A to X_0. Then x_0 is the solution to* (5.1) *iff there exists $z_0^* \in \operatorname{ran} A_|^*$ such that*

$$x_0 = (z_0^*)_+^{1/(p-1)} \in A_|^{-1} d. \tag{5.6}$$

Proof. As shown in Theorem 3.4, $d \neq 0$ implies that $d \in \operatorname{int} A_| C_0$. Since $A_| X_0 = \operatorname{span}(d)$, (5.6) is asserted by Theorem 5.1. ∎

When $Y = \mathbb{R}^n$, $A := \sum_{i=1}^n e_i \Phi_i$ for $\Phi_i \in L_q(\Omega)$, we have the following corollary as an immediate result of Theorem 5.2.

COROLLARY 5.1. *x_0 is the solution of* (5.1) *iff there exists $c \in \mathbb{R}^n$ such that*

$$x_0 = \left(\sum_{i=1}^{m} (e_i, \varphi) \varphi_i \right)_+^{1/(p-1)} \in A_|^{-1} d, \tag{5.7}$$

where X_0 is the smallest closed subspace of X containing $C \cap A^{-1}d$ and $A_| := \sum_{i=1}^n e_i \varphi_i$ is the restriction of A to X_0.

ACKNOWLEDGMENTS

I take this opportunity to thank Professor Carl de Boor for the helpful discussions. Also my thanks go to Professor Stephen Robinson for the references.

REFERENCES

[B78] V. BARBU AND TH. PRECUPANU, "Convexity and Optimization in Banach Spaces," Stijthoff & Noordhoff, Bucuresti, 1978.

[C89] C. CHUI, F. DEUTSCH, AND J. WARD, "Constrained Best Approximation in Hilbert Space, II," CAT Rep. No. 193, 1989.

[C90] C. CHUI, F. DEUTSCH, AND J. WARD, Constrained best approximation in Hilbert space, *Constr. Approx.* **6** (1990), 35–64.

[H70] R. Holmes, "A Course on Optimization and Best Approximation," Springer-Verlag, New York, 1970.

[H75] R. Holmes, "Geometric Functional Analysis and Its Applications," Springer-Verlag, New York, 1975.

[M85] C. Micchelli, P. Smith, J. Swetits, and J. Ward, Constrained L_p approximation, *Constr. Approx.* **1** (1985), 93–102.

[M88] C. Micchelli and F. Utreras, Smoothing and interpolation in a convex subset of a Hilbert space, *SIAM J Sci. Statist. Comput.* **9** (1988), 728–746.

[N70] L. Neustadt, Sufficiency conditions and dual theory for mathematical programming problems in arbitrary linear spaces, *in* "Nonlinear Programming" (J. Rosen, O. Mangasarian, and K. Ritter, Eds.), pp. 323–348, Academic Press, New York, 1970.

[R74] R. Rockafellar, Conjugate duality and optimization, *in* "SIAM Regional Conference Series in Applied Mathematics, 1974."