ON CONSTRAINT QUALIFICATION FOR AN INFINITE SYSTEM
OF CONVEX INEQUALITIES IN A BANACH SPACE

CHONG LI† AND K. F. NG‡

Abstract. For a general infinite system of convex inequalities in a Banach space, we study the
basic constraint qualification and its relationship with other fundamental concepts, including various
versions of conditions of Slater type, the Mangasarian–Fromovitz constraint qualification, as well
as the Pshenichnyi–Levin–Valadier property introduced by Li, Nahak, and Singer. Applications are
given in the restricted range approximation problem, constrained optimization problems, as well as
in the approximation problem with constraints by conditionally positive semidefinite functions.

Key words. system of convex inequalities, basic constraint qualification, Slater condition,
MFCQ, PLV property, optimality condition, best approximation

AMS subject classifications. Primary, 90C34; Secondary, 90C46, 41A29

DOI. 10.1137/S1052623403434693

1. Introduction. Let \( X \) be a Banach space over the real field \( \mathbb{R} \), \( C \) be a closed
convex subset of \( X \), and let \( \{ g_i : i \in I \} \) be a family of continuous convex functions
on \( X \), where \( I \) is an arbitrary (but nonempty) index-set. We consider the following
system of convex inequalities:

\[
\text{(1.1)} \quad g_i(x) \leq 0, \quad i \in I,
\]

and the associated system for the sup-function \( G \),

\[
\text{(1.2)} \quad G(x) \leq 0,
\]

where \( G \) is defined by

\[
\text{(1.3)} \quad G(x) := \sup_{i \in I} g_i(x) \quad \text{for all } x \in X.
\]

In this paper, we always assume that the sup-function \( G(\cdot) \) is continuous on \( X \).
It is clear that if \( X \) is of finite dimension and \( G(x) \) is finite for each \( x \in X \), or if
\( \{ g_i : i \in I \} \) is locally uniformly bounded, then the condition that \( G(\cdot) \) is continuous
on \( X \) is automatically satisfied. Throughout we use \( S \) to denote the solution set of
(1.1) (equivalently, of (1.2)):

\[
\text{(1.4)} \quad S := \{ x \in X : g_i(x) \leq 0 \quad \text{for all } i \in I \} = \{ x \in X : G(x) \leq 0 \}.
\]

Let

\[
\text{(1.5)} \quad K := C \cap S;
\]
that is, \( K \) is the solution set of (1.1) with the constraint

\[(1.6) \quad x \in C.\]

Many important problems in mathematical programming are based on relationships between the normal cone \( N_K(x) \) in relation to \( N_C(x) \) and the subdifferentials of \( G, g_i \) with \( i \in I(x) \). Here \( x \in X \), and \( I(x) \) is the set of “active indices” at \( x \) defined by

\[(1.7) \quad I(x) := \{ i \in I : g_i(x) = G(x) \}.

In the classical case (that is, when \( I \) is finite, \( X \) is a Euclidean space, and each \( g_i \) is differentiable), many fundamental notions of so-called constraint qualifications for (1.1) have been introduced, such as the basic constraint qualification (BCQ), the Slater condition (SC), and the Mangasarian–Fromovitz constraint qualification (MFCQ); their interrelationships are well known and play a very important part not only in optimization problems but also in many other areas, such as best approximation theory. A big step forward from the classical theory was recently achieved by Li, Nahak, and Singer in [16], where detailed studies on the BCQ and the SC have been made for the case when the index-set \( I \) is not restricted to be finite. They introduced and studied the Pshenichnyi–Levin–Valadier (PLV) property and the weak PLV property; in particular, they made use of these properties in extending the classical theory to the case when \( I \) is infinite. Another direction of extension was done in [12, 13], where \( I \) was assumed to be finite but \( X \) was allowed to be an infinite dimensional Hilbert space; moreover, the constraint (1.6) was considered.

In this paper, we propose to study the general case: \( X \) is a Banach space, \( C \) is a closed convex subset, and \( I \) is an arbitrary set. In this general case we introduce and study various notions of constraint qualifications. In particular, in the stated general setting, the notions of the BCQ and the PLV are introduced in section 2. We show in Theorem 2.4 that, in the presence of the PLV, (1.1) satisfies the BCQ if and only if (1.2) satisfies the BCQ. Our main results are presented in section 3, where we establish some sufficient conditions to ensure the BCQ. Along with the extended notions of the SC and the MFCQ, the weak versions (weak SC and weak MFCQ) are also introduced. Some of their relationships are established. We show in particular that, under the weak SC, the PLV implies the BCQ for (1.1). We remark that many of the results obtained in this paper, for example, Theorems 3.8 and 3.10, are new even for the case when \( C = X \) is a finite dimensional space. In particular, Theorem 3.10 provides a set of conditions ensuring that the following implication holds:

Each finite subsystem of (1.1) satisfies the BCQ \( \implies \) the system (1.1) does.

Meanwhile, Examples 3.2, 3.3, and 3.4 show that the above implication is no longer true if any of the conditions in the theorem are dropped. The last three sections of the paper are devoted to applications of the main results. In section 4, we study the problem of minimizing a real-valued function \( f(x) \) subjected to \( x \in C \cap S \) and we show that, for an optimal point \( x_0 \), the BCQ of (1.1) relative to \( C \) at \( x_0 \) is closely related to the KKT property. In section 6, we address an approximation problem studied by H. Strauss. We give a counterexample to show that [22, Theorem 6.2] is not true, and we provide a new version which is established under the added assumption of the BCQ.

One of the reasons why we study the proposed general case instead of the more restrictive case is that there are many approximation and optimization problems which
are in the general Banach space setting with infinitely many convex constraints. For example, it was shown in [14] that the restricted Chebyshev approximation problem (see [11, 21] and references therein) in the Banach space of complex-valued continuous functions on a compact Hausdorff space can be reformulated as a constrained approximation problem with a system of convex inequalities. Section 5 is devoted to another application dealing with the best restricted range approximation problem in $L^p$ space studied by Levasseur and Levis: a main result from [10] is extended and improved under much weaker conditions.

2. Preliminaries, the BCQ, and the PLV property. For a set $Z$ in $X$ (or in $\mathbb{R}^n$), $\text{ext } Z$ denotes the set of all extreme points of $Z$ and the interior (resp., relative interior, closure, convex hull, convex cone hull, linear hull, negative polar, boundary) of $Z$ is defined by $\text{int } Z$ (resp., $\text{ri } Z$, $\overline{Z}$, $\text{conv } Z$, $\text{cone } Z$, $\text{span } Z$, $Z^\ominus$, $\text{bd } Z$). We use $|Z| (\leq \infty)$ to denote the number of elements of $Z$ and use $\dim Z$ to denote the dimension of $\text{span } (Z - z)$, where $z$ is an element of $Z$. For $x \in X$, in the case when $Z$ is closed and convex, the normal cone of $Z$ at $x$ is denoted by $N_Z(x)$ and defined by $N_Z(x) = (Z - x)^\ominus$. $\mathbb{R}_-$ denotes the subset of $\mathbb{R}$ consisting of all nonpositive real numbers.

For a proper convex function on $X$ and $x \in X$, the subdifferential of $f$ at $x$ is defined by

$$\partial f(x) := \{z^\ast \in X^* : f(x) + \langle z^\ast, y - x \rangle \leq f(y) \quad \text{for all } y \in X\}.$$  

In particular, $N_Z(z) = \partial I_Z(z)$ for each $z \in Z$. Here and throughout, $I_Z$ denotes the indicator function of $Z$: $I_Z(x) = 0$ if $x \in Z$ and $I_Z(x) = +\infty$ if $x \notin Z$.

Recall that the directional derivative $f'(x, d)$ of the function $f$ at $x$ in the direction $d$ is defined by

$$f'(x, d) = \lim_{t \to 0^+} \frac{f(x + td) - f(x)}{t}. \quad (2.1)$$

Then (cf. [6, Proposition 2.2.7]) if $f$ is convex and if $x$ is a continuity point of $f$,

$$\partial f(x) = \{z^\ast \in X^* : \langle z^\ast, d \rangle \leq f'(x, d) \quad \text{for all } d \in X\} \quad (2.2)$$

and

$$f'(x, d) = \max\{\langle z^\ast, d \rangle : z^\ast \in \partial f(x)\}. \quad (2.3)$$

**Remark 2.1.** For a continuous convex function $f$ on $X$, it is easy to show that cone $\partial f(x) \subseteq N_{f^{-1}(\mathbb{R}_-)}(x)$ if $f(x) = 0$ and that the equality holds in both of the following cases: (a) $f$ is affine, (b) $x$ is not a minimizer of $f$. See [6, Corollary 1, p. 56].

**Definition 2.1.** Let $x \in X$, and let $I(x)$ be defined by (1.7). Assuming $I(x) \neq \emptyset$, define $D'(x)$ and $N'(x)$ by

$$D'(x) := \text{conv} \left( \bigcup_{i \in I(x)} \partial g_i(x) \right), \quad (2.4)$$

$$N'(x) := \text{cone} \left( \bigcup_{i \in I(x)} \partial g_i(x) \right). \quad (2.5)$$
also we adopt the convention that
(2.6) \[ N'(x) = \{0\} \text{ if } I(x) = \emptyset. \]

Part (ii) of the following proposition is well known while the other parts can be verified easily by definitions.

**Proposition 2.2.** The following assertions hold:
(i) If \( i \in I(x) \), then \( \partial g_i(x) \subseteq \partial G(x) \). Hence
(2.7) \[ N'(x) \subseteq \text{cone} \partial G(x) \text{ for all } x \in X \]
and
(2.8) \[ D'(x) \subseteq \partial G(x) \text{ for all } x \in X. \]

(ii) If \( I \) is finite, then
(2.9) \[ D'(x) = \partial G(x) \text{ for all } x \in X. \]

(iii) If \( G(x) = 0 \) (e.g., if \( x \in \text{bd } S \)), then
(2.10) \[ \partial G(x) \subseteq N_K(x) \]
and
(2.11) \[ N_C(x) + N'(x) \subseteq N_C(x) + \text{cone} \partial G(x) \subseteq N_K(x). \]

Below we define the conversed relations of the inclusions given in the preceding proposition.

**Definition 2.3.** Let \( x \in X \). The system (1.1) is said to satisfy
(a) the BCQ at \( x \) relative to \( C \) if
(2.12) \[ N_C(x) + N'(x) \supseteq N_K(x); \]
(b) the PLV at \( x \) relative to \( C \) if
(2.13) \[ N_C(x) + D'(x) \supseteq N_C(x) + \partial G(x). \]

The system (1.1) is said to satisfy the BCQ relative to \( C \) if it satisfies the BCQ at each point of \( K \). In addition, the wording “relative to \( C \)” need not be mentioned if \( C = X \).

The following theorem (the proof of which is routine and hence omitted) extends a result of Li, Nahak, and Singer for the case when \( C = X \) is finite dimensional.

**Theorem 2.4.** Let \( x \in C \cap \text{bd } S \). Suppose that the system (1.1) satisfies the PLV at \( x \) relative to \( C \). Then the system (1.1) satisfies the BCQ at \( x \) relative to \( C \) if and only if (1.2) satisfies the BCQ at \( x \) relative to \( C \).

Part (ii) of the following theorem is well known; see, for example, [9, Theorem 4.4.2] and [16, Theorem 3.1]. The proof given in [16] is valid for any compact space \( I \) satisfying the first axiom of countability (in the sense that there exists a countable local base at each point of \( I \)). For wider scope of applications (such as that in section 5), we do not assume that \( I \) is metrizable.

**Theorem 2.5.** Suppose that \( I \) is a compact space satisfying the first axiom of countability and that the function: \( i \mapsto g_i(x) \) is upper semicontinuous for each \( x \in X \). Then \( I(x) \neq \emptyset \) for each \( x \in X \) and the following assertions hold:
(i) $\partial G(x) = \overline{D'(x)}$.
(ii) $\partial G(x) = D'(x)$, provided that $I(x)$ is finite or $X$ is finite dimensional.
(iii) If $x \in \text{span } C$ and if $\text{span } C$ is finite dimensional, then the system (1.1) satisfies the PLV relative to $C$ at $x$.

Furthermore, if in addition $\{g_i : i \in I\}$ is equicontinuous on $X$, then the assumption on $I$ about the first axiom of countability can be dropped.

Proof. Let $x, d \in X$. By assumption, it is easy to verify that $I(x) \neq \emptyset$. For each $t > 0$, define

$$I_t := \left\{ i \in I : \frac{g_i(x + td) - G(x)}{t} \geq G'(x, d) \right\}.$$  

By convexity, $I(x + td) \subseteq I_t$ and it follows that $I_t$ is nonempty. Moreover, $\{I_t : t > 0\}$ has the “finite intersection property” (and hence $\cap_{t>0}I_t \neq \emptyset$) because the function

$$t \mapsto \frac{g_i(x + td) - G(x)}{t} = \frac{g_i(x + td) - g_i(x)}{t} + \frac{g_i(x) - G(x)}{t}$$

is nondecreasing on $(0, +\infty)$ (the functions represented in the two terms of the right-hand side of (2.15) are nondecreasing by the convexity of $g_i$ and the fact that $g_i(x) - G(x) \leq 0$). Pick $i_0 \in \cap_{t>0}I_t$. Then passing to the limits as $t \to 0$ in

$$g_{i_0}(x + td) - G(x) \geq tG'(x, d),$$

we see that $g_{i_0}(x) = G(x)$ (that is, $i_0 \in I(x)$) and so (2.16) entails $g'_{i_0}(x, d) \geq G'(x, d)$. Consequently, it follows from (2.3) that

$$G'(x, d) = \max_{i \in I(x)} \{g_i'(x, d)\} = \max_{i \in I(x)} \left\{ \max_{x^* \in \partial g_i(x)} \langle x^*, d \rangle \right\} = \max_{x^* \in D'(x)} \langle x^*, d \rangle.$$  

Thus (i) must hold by [6, Proposition 2.1.4]. Note that we have not used the assumption on the first axiom of countability. For (iii), write $X$ for $\text{span } C$; let $\tilde{G}$ and $\tilde{g}_i$, respectively, denote the restrictions of $G$ and $g_i$ to $\tilde{X}$. Clearly, $\tilde{G}$ is the sup-function of $\{\tilde{g}_i : i \in I\}$ and it follows from (ii) that, for $x \in \tilde{X}$,

$$\partial \tilde{G}(x) = \text{conv } \{\partial \tilde{g}_i(x) : i \in I(x)\}.$$  

To prove (iii), we need only show that

$$\partial G(x) \subseteq N_C(x) + D'(x).$$

To do this, let $y^* \in \partial G(x)$, and let $\tilde{y}^*$ denote the restriction of $y^*$ to $\tilde{X}$. Then $\tilde{y}^* \in \partial \tilde{G}(x)$, and so it follows from (2.18) that $\tilde{y}^* = \sum_{j=1}^{m} \lambda_j \tilde{y}_j^*$ for some $\{\lambda_j\} \subseteq [0, 1]$ and $\{\tilde{y}_j^*\} \subseteq \tilde{X}^*$ with $\sum_{j=1}^{m} \lambda_j = 1$, $\tilde{y}_j^* \in \partial \tilde{g}_i(x)$, and $i_1, i_2, \ldots, i_m \in I(x)$. Noting

$$\langle \tilde{y}_j^*, z \rangle \leq \tilde{g}'_{i_j}(x, z) = g'_{i_j}(x, z) \quad \text{for all } z \in \tilde{X},$$

one can apply the Hahn–Banach theorem to obtain a continuous linear extension $y_j^*$ of $\tilde{y}_j^*$ such that

$$\langle y_j^*, \cdot \rangle \leq g'_{i_j}(x, \cdot) \quad \text{on } X.$$
Thus \( y_j^* \in \partial g_i(x) \) and

\[
\langle y^*, \cdot \rangle = \left( \sum_{j=1}^{m} \lambda_j y_j^* \right) \quad \text{on } \hat{X}.
\]

Since \( C - x \subseteq \hat{X} \), this implies that \( y^* - \sum_{j=1}^{m} \lambda_j y_j^* \in N_C(x) \). Therefore \( y^* \in N_C(x) + N'(x) \) and (2.19) is proved.

Now we drop the assumption that \( I \) satisfies the first axiom of countability but assume that \( \{g_i : i \in I\} \) is equicontinuous on \( X \). By the above proof, (iii) will be true if (ii) holds. Hence, to complete the proof it suffices to show that \( \partial G(x) = D'(x) \) when \( X \) is finite dimensional. For this purpose, we need to introduce a new topology on the set \( I \). For each \( i_0 \in I, x \in X \), and each \( r \in \mathbb{R}, r > 0 \), set

\[
O_x(i_0, r) = \{ i \in I : g_i(x) - g_{i_0}(x) < r \}.
\]

Since \( X \) is finite dimensional, there exists a countable subset \( A \) of \( X \) such that the closure \( \overline{A} \) is equal to \( X \). Let \( \tau_N \) denote the topology generated by \( \{ O_x(i_0, r) : x \in A, i_0 \in I, r \in \mathbb{R}, r > 0 \} \). Then, under the topology \( \tau_N \), the following assertions hold:

1. \( I \) is compact;
2. \( I \) satisfies the first axiom of countability;
3. for each \( x \in X \), the function \( i \mapsto g_i(x) \) is upper semicontinuous.

In fact, (2) is trivial, and (1) holds because \( I \) is compact under the original topology, which is stronger than the new topology \( \tau_N \), as each function \( i \mapsto g_i(x) \) is upper semicontinuous. (3) follows from the density of \( A \) in \( X \) and the equicontinuity of \( \{g_i : i \in I\} \) on \( X \). In fact, let \( x \in X \) and \( \epsilon > 0 \). By the equicontinuity, there exists \( \delta > 0 \) such that

\[
|g_i(x) - g_i(y)| < \frac{\epsilon}{3} \quad \text{for each } i \in I \text{ and each } y \text{ with } \|x - y\| < \delta.
\]

Pick \( x_\epsilon \in A \) such that \( \|x_\epsilon - x\| < \delta \). Let \( i_0 \in I \). Noting that, by (2.21),

\[
g_i(x) - g_{i_0}(x) < \frac{2}{3}\epsilon + g_i(x_\epsilon) - g_{i_0}(x_\epsilon),
\]

we get

\[
g_i(x) - g_{i_0}(x) < \epsilon \quad \text{for each } i \in O_x(i_0, \epsilon/3).
\]

This shows that the function \( i \mapsto g_i(x) \) is upper semicontinuous at \( i_0 \), and hence (3) holds. Thus, by the first part of the proof, it is easy to see that \( \partial G(x) = D'(x) \) and the proof is complete.

**Remark 2.2.** In general, \( D'(x) \) need not be \( w^*\)-closed, and hence the \( w^*\)-closure in (i) cannot be dropped. For example, take \( X = l^2 \), the Hilbert space consisting of all infinite real sequences \( (x_i) \) satisfying \( \sum_{j=1}^{\infty} |x_j|^2 < \infty \). Write

\[
e_{j} = (0, \ldots, 0, 1, 0, \ldots) \quad \text{for all } j = 1, 2, \ldots,
\]

that is, the \( j \)th component of \( e_j \) is 1 while all other components are 0. Let \( I = \{0, 1, 1/2, \ldots, 1/j, \ldots\} \) and define \( \{g_i\}_{i \in I} \) as follows:

\[
g_i(x) = \begin{cases} 0, & i = 0, \\ \langle ie_1/i, x \rangle, & i \in I \setminus \{0\}. \end{cases}
\]
Let \( x = 0 \). Then \( I(x) = I \),

\[
\partial g_i(x) = \begin{cases} 
\{0\}, & i = 0, \\
\{ie_1/i\}, & i \neq 0,
\end{cases}
\]

and hence,

\[
D'(x) = \left\{ \sum_{j=1}^{m} \lambda_j \frac{1}{j} e_j : m > 0, \lambda_j \geq 0, \sum_{j=1}^{m} \lambda_j \leq 1 \right\}.
\]

Clearly, \( D'(x) \) is not closed.

**3. The SC and the MFCQ.** We continue to consider the system (1.1) relative to a closed convex set \( C \) in \( X \). Notation is as in the introduction. A main aim of this section is to provide sufficient conditions ensuring that (1.1) satisfies the BCQ relative to \( C \). These conditions are of two types (the SC and the MFCQ): one is expressed in terms of the functions \( g_i \), while the other is expressed in terms of the directional derivatives of \( g_i \). Let us first extend several known concepts to our present general case from the semi-infinite case (cf. [2, 12, 15] for the former type and [12, 17, 23] for the latter type). Recall that \( S \) and \( K \) are defined by (1.4) and (1.5).

**Definition 3.1.** Let \( x \in C \cap \text{bd} S \). An element \( d \in X \) is called

(a) a linearized feasible direction of (1.1) at \( x \) if

\[
\langle z^*, d \rangle \leq 0 \quad \text{for all } z^* \in \text{ext } \partial g_i(x) \text{ and for all } i \in I(x);
\]

(b) a sequentially feasible direction of \( K \) at \( x \) if there exist a sequence \( \{d_k\} \) with \( d_k \rightarrow d \) and a sequence \( \{\delta_k\} \) of positive real numbers with \( \delta_k \rightarrow 0 \) such that \( \{x + \delta_k d_k\} \subseteq K \).

Let \( \text{LFD}(x) \) (resp., \( \text{SFD}(x) \)) denote the set of all \( d \) satisfying (a) (resp., (b)). Note that \( \text{LFD}(x) \) and \( \text{SFD}(x) \) are both closed convex cones.

**Definition 3.2.** Let \( K_S(x) \), \( K_L(x) \) be, respectively, defined by

\[
K_S(x) = (x + \text{SFD}(x)) \cap C
\]

and

\[
K_L(x) = (x + \text{LFD}(x)) \cap C.
\]

Note that these two sets are closed convex sets.

**Proposition 3.3.** Let \( x \in C \cap \text{bd} S \). Then \( \text{SFD}(x) \subseteq \text{LFD}(x) \), and hence

\[
K \subseteq K_S(x) \subseteq K_L(x).
\]

**Proof.** Since \( K \subseteq K_S(x) \), (3.4) follows from the first assertion and the fact that \( \text{LFD}(x) \) is closed convex. To prove the first assertion, let \( d \in \text{SFD}(x) \), and let \( \{d_k\}, \{\delta_k\} \) be as in Definition 3.1(b). In particular, for each \( i \in I(x) \), one has \( g_i(x + \delta_k d_k) \leq 0 \) and hence, for each \( z^* \in \partial g_i(x) \), we have

\[
\langle z^*, \delta_k d_k \rangle \leq g_i(x + \delta_k d_k) - g_i(x) \leq 0
\]

thanks to the fact that \( G(x) = 0 \) as \( x \in \text{bd} S \). This implies that \( \langle z^*, d_k \rangle \leq 0 \). Letting \( k \rightarrow +\infty \) we obtain \( \langle z^*, d \rangle \leq 0 \). Since \( i \in I(x) \) and \( z^* \in \partial g_i(x) \) are arbitrary, \( d \in \text{LFD}(x) \), and the proof is complete. \( \Box \)

**Definition 3.4.** Let \( x \in C \cap \text{bd} S \). We say that the system (1.1) satisfies
(a) the MFCQ relative to \( C \) at \( x \) if the following conditions are satisfied:
0. \( (x + \text{LFD}(x)) \cap \text{ri} \, C \neq \emptyset \);
1. \( \{d \in X : g_i^*(x, d) < 0 \text{ for all } i \in I(x)\} \cap \text{span} (C - x) \neq \emptyset \);

(b) the weak MFCQ relative to \( C \) at \( x \) if there exists a finite subset \( I_0 \) of \( I(x) \) such that
0. \( (x + \text{LFD}(x)) \cap \text{ri} \, C \neq \emptyset \);
1. \( \{d \in X : g_i^*(x, d) < 0 \text{ for all } i \in I(x) \setminus I_0\} \cap \text{span} (C - x) \neq \emptyset \);
2. \( g_i \) is affine for each \( i \in I_0 \).

We shall also say that \( x \) satisfies the MFCQ (resp., the weak MFCQ) for (1.1) relative to \( C \) when (a) (resp., (b)) occurs.

For the following definitions, we need additional notation. Let

\[
\tilde{I}(x) = \{i \in I(x) : g_i(x) = 0\}.
\]

Clearly, \( \tilde{I}(x) = I(x) \) in the case when \( x \in \text{bd} \, S \).

**Definition 3.5.** The system (1.1) is said to satisfy

(a) the SC relative to \( C \) if there exists \( \bar{x} \in \text{ri} \, C \) such that \( G(\bar{x}) < 0 \);

(b) the weak SC relative to \( C \) if there exists \( \bar{x} \in \text{ri} \, C \cap S \) such that

\[
\begin{align*}
(1) & \quad \tilde{I}(\bar{x}) \text{ is finite and } g_i \text{ is affine for each } i \in \tilde{I}(\bar{x}) ; \\
(2) & \quad G_0 \text{ is continuous and } G_0(\bar{x}) < 0, \text{ where } G_0 \text{ denotes the sup-function of } \{g_i : i \in I \setminus \tilde{I}(\bar{x})\} ; \\
(3) & \quad G_0(z) = \sup_{i \in I \setminus \tilde{I}(\bar{x})} g_i(z), \quad z \in X.
\end{align*}
\]

A point \( \bar{x} \) with the property in (a) (resp., (b)) is called a Slater (resp., weak Slater) point for the system (1.1) relative to \( C \).

As the nomenclature suggests, it is clear that (a) \( \Rightarrow \) (b) in Definitions 3.4 and 3.5.

**Theorem 3.6.** Suppose there exists a weak Slater point \( \bar{x} \) for the system (1.1) relative to \( C \), and let \( x \in C \cap \text{bd} \, S \). Then (1.1) satisfies the BCQ relative to \( C \) at \( x \) if (1.1) satisfies the PLV relative to \( C \) at \( x \).

**Proof.** Let \( G_0 \) be defined as in Definition 3.5(b). Set

\[
S_0 = \{y \in X : g_i(y) \leq 0, \ i \in I \setminus \tilde{I}(\bar{x})\}
\]

and

\[
S_1 = \{y \in X : g_i(y) \leq 0, \ i \in \tilde{I}(\bar{x})\}.
\]

Then \( G_0(\bar{x}) < 0, \tilde{I}(\bar{x}) \) is finite, and \( g_i \) is affine for each \( i \in \tilde{I}(\bar{x}) \). Hence

\[
N_{S_1}(x) = \begin{cases} 
\sum_{i \in I(x) \setminus \tilde{I}(\bar{x})} \text{cone} \, \partial g_i(x) & \text{if } I(x) \cap \tilde{I}(\bar{x}) \neq \emptyset, \\
0 & \text{if } I(x) \cap \tilde{I}(\bar{x}) = \emptyset
\end{cases}
\]

and

\[
N_{S_0}(x) = \begin{cases} 
\text{cone} \, \partial G_0(x) & \text{if } G_0(x) = 0, \\
0 & \text{if } G_0(x) < 0.
\end{cases}
\]

Note in particular that \( N_{S_1}(x) \) and \( N_{S_0}(x) \) are contained in \( \text{cone} \, \partial G(x) \). Moreover, by [8, Proposition 2.3 and Theorem 5.1] (the proofs given there are valid for Banach
spaces, though the results were stated in a Hilbert space setting), \( \{C, S_1\} \) and \( \{C \cap S_1, S_0\} \) have the strong conical hull intersection property (the strong CHIP). Hence

\[
N_K(x) = N_{C \cap S_1}(x) + N_{S_0}(x) = N_C(x) + N_{S_1}(x) + N_{S_0}(x)
\]

\[
\subseteq N_C(x) + \text{cone} \partial G(x) \subseteq N_C(x) + N'(x),
\]

where the last inclusion holds because of the PLV property. \( \square \)

For convenience of reference, we note the following immediate consequence of Theorem 2.5(iii) and Theorem 3.6.

**Corollary 3.7.** Suppose that \( I \) is a compact space satisfying the first axiom of countability, and that the function: \( i \mapsto g_i(x) \) is upper semicontinuous for each \( x \in X \). Let \( x \in K \). Suppose that \( \dim C < \infty \) or \( I(x) \) is finite and that there exists a weak Slater point for the system (1.1) relative to \( C \). Then (1.1) satisfies the BCQ relative to \( C \) at \( x \).

Furthermore, if in addition \( \{g_i : i \in I\} \) is equicontinuous on \( X \), then the assumption on \( I \) about the first axiom of countability can be dropped.

**Theorem 3.8.** For the system (1.1) and relative to \( C \), the following assertions hold:

(i) If the weak SC is satisfied, then at each point in \( C \cap \text{bd} S \) the weak MFCQ is satisfied. If \( I \) is finite and if \( x \in C \cap \text{bd} S \) satisfies the weak MFCQ, then there exist weak Slater points arbitrarily near \( x \).

(ii) If the SC is satisfied, then, at each point in \( C \cap \text{bd} S \), the MFCQ is satisfied. Conversely, if \( x \in C \cap \text{bd} S \) satisfies the MFCQ and if in addition \( x \) satisfies the PLV, then there exist Slater points arbitrarily near \( x \) (in particular, (1.1) satisfies the BCQ relative to \( C \)).

**Proof.** (i) Let \( \bar{x} \in \text{ri} C \cap S \) be a weak Slater point for the system (1.1) relative to \( C \), and let \( x \in C \cap \text{bd} S \) (hence \( G(x) = 0 \)). Set \( I_0 = I(x) \cap \bar{I}(\bar{x}) \). Then \( I_0 \) is finite (as \( \bar{I}(\bar{x}) \)), and \( x \) with this \( I_0 \) satisfies (1)–(3) of Definition 3.4(b). Indeed, (3) is evident. Suppose \( I(x) \subseteq \bar{I}(\bar{x}) \). Then (2) of Definition 3.4(b) holds trivially. (1) also holds as \( \bar{x} \in (x + \text{LFD}(x)) \cap \text{ri} C \) because, for each \( i \in I_0 = I(x) \subseteq \bar{I}(\bar{x}) \), one has

\[
g_i(x) = 0 \quad \text{and} \quad g_i(\bar{x}) = G(\bar{x}).
\]

It follows from the convexity that \( g_i'(x, \bar{x} - x) \leq 0 \); that is, \( \bar{x} - x \in \text{LFD}(x) \). We can therefore assume henceforth that \( I(x) \not\subseteq \bar{I}(\bar{x}) \). Consequently \( G_0(x) = G(\bar{x}) = 0 \), where \( G_0 \) is the function on \( X \) defined by (3.7). Noting \( G_0(\bar{x}) < 0 \), we have for each \( i \in I(x) \setminus I_0 \) that

\[
g_i'(x, \bar{x} - x) \leq g_i(\bar{x}) - g_i(x) \leq G_0(\bar{x}) - G_0(x) < 0.
\]

This implies that the intersection in (2) of Definition 3.4(b) is an empty set (containing \( \bar{x} - x \)) and that \( \bar{x} - x \in \text{LFD}(x) \) as \( g_i \) is affine for each \( i \in I_0 \) (because \( I_0 \subseteq \bar{I}(\bar{x}) \) and \( \bar{x} \) is a weak Slater point). Therefore \( x \) is a weak MFCQ point relative to \( C \). Conversely, we suppose that \( x \) satisfies the weak MFCQ relative to \( C \) and, in addition, that \( I \) is finite. Let \( I_0 \subseteq I(x) \) satisfy (2) and (3) of Definition 3.4(b). Then by definition, \( x \) satisfies the MFCQ relative to \( C \) for the subsystem defined by

\[
g_i(i) \leq 0 \quad \text{for all} \quad i \in I \setminus I_0.
\]

Note that \( x \) satisfies the PLV for the subsystem (3.8) as \( I \) is finite. Hence, assuming the results in (ii), we can apply (ii) to conclude that there exists \( \bar{x} \in \text{ri} \{c \} \) arbitrarily near \( x \) such that \( \sup \{g_i(\bar{x}) : i \in I \setminus I_0\} < 0 \). Clearly then \( \bar{x} \) is a weak Slater point for (1.1) relative to \( C \).
(ii) Let $\bar{x}$, $x$ be as in the first part of (i) but assume the stronger assumption that $\bar{x}$ is a Slater point relative to $C$. Then $\bar{I}(\bar{x})$ must be empty and so must $I_0$. In this case, (2) of Definition 3.4(b) coincides with (2) of Definition 3.4(a). Therefore $x$ is an MFCQ point relative to $C$ by (i). Conversely, suppose that $x$ satisfies the MFCQ for the system (1.1) relative to $C$ and, in addition, that $x$ also satisfies the PLV. Then, by the former, there exist $d_1$, $d_2$ such that

$$x + d_1 \in (x + \text{LFD}(x)) \cap \text{ri } C, \ d_2 \in \text{span}(C - x),$$

and

$$g_i'(x, d_2) < 0 \quad \text{for all } i \in I(x).$$

Also, by the PLV property,

$$\partial G(x) \subseteq N_C(x) + \text{conv } \{\partial g_i(x) : i \in I(x)\}.$$  

Take $t_0 > 0$ small enough such that $x + (1-t_0)d_1 + t_0d_2 \in \text{ri } C$. Set $d := (1-t_0)d_1 + t_0d_2$. Then, for any $i \in I(x),

$$g_i'(x, d) \leq (1-t_0)g_i'(x, d_1) + t_0g_i'(x, d_2) \leq t_0g_i'(x, d_2) < 0.$$  

It follows that

$$G'(x, d) < 0.$$  

Indeed, by (2.3) there exists $z^* \in \partial G(x)$ such that $G'(x, d) = \langle z^*, x \rangle$. By (3.10), there exist $t_1, \ldots, t_m > 0$ and $z_{j_1}^*, \ldots, z_{j_m}^* \in X^*$ with $\sum_{j=1}^m t_j = 1$, and $z_j^* \in \partial g_i(x)$ where $i_1, \ldots, i_m \in I(x)$ such that $z^* - \sum_{j=1}^m t_j z_j^* \in N_C(x)$; in particular, $\langle z^* - \sum_{j=1}^m t_j z_j^* \rangle = 0$. It follows from (3.11) and (2.2) that

$$\langle z^*, d \rangle \leq \sum_{j=1}^m t_j g_j'(x, d) < 0;$$

this proves (3.12). Noting $G(x) = 0$ (as $x \in \text{bd } S$), (3.12) implies that $G(x + \hat{t}d) < 0$ for all sufficiently small $\hat{t} \in (0, 1)$. Clearly $x + \hat{t}d \in \text{ri } C$ (provided $\hat{t} < t_0$) and so $x + \hat{t}d$ is a Slater point for the system (1.1) relative to $C$. □

**Remark 3.1.** Example 3.1(a) below will show that the finiteness assumption cannot be dropped in the second assertion of Theorem 3.8(i), and the condition of the PLV in the second assertion of Theorem 3.8(ii) cannot be dropped. Example 3.1(b) provides an example of (1.1) which satisfies the PLV and the weak MFCQ relative to $C$ at some point $x_0 \in C \cap \text{bd } S$ but fails for the weak SC and the BCQ relative to $C$ at $x_0$. (In particular, the finiteness assumption in the second assertion of Theorem 3.8(i) cannot be replaced by the PLV.)

**Example 3.1.** (a) Let $X = C = \mathbb{R}$, $I = [-1, 0) \cup (0, +1]$, and let $\{g_i : i \in I\}$ be defined by

$$g_i(x) = \begin{cases} ix, & i \in [-1, 0) \\ 2ix - i^2, & i \in (0, +1] \end{cases} \quad \text{for all } x \in \mathbb{R}.$$  

Then the corresponding sup-function is given by

$$G(x) = \begin{cases} 2x - 1, & x \geq 1 \\ x^2, & 0 \leq x \leq 1 \\ -x, & x \leq 0 \end{cases} \quad \text{for all } x \in \mathbb{R}.$$  

(3.13)
Take \( x_0 = 0 \). Then, \( S = \{0\} \), \( I(x_0) = [-1, 0) \), and
\[
\partial g_i(x_0) = \{i\} \quad \text{for all } i \in [-1, 0).
\]

Moreover,
\[
g'_i(x_0, 1) = i < 0 \quad \text{for all } i \in [-1, 0)
\]
and
\[
LFD(x_0) = [0, +\infty).
\]

For the system (1.1) relative to \( C \) at the point \( x_0 \), the following assertions hold:

1. The MFCQ is satisfied.
2. The PLV is not satisfied.
3. The weak SC is not satisfied.
4. The BCQ at \( x_0 \) is not satisfied.

Indeed, (1) is evident by (3.15) and (3.16). Next, by (3.14),
\[
D'(x_0) = [-1, 0), \quad N'(x_0) = (-\infty, 0],
\]
so that
\[
\partial G(x_0) = [-1, 0] = D'(x_0) \subseteq (-\infty, 0] = N'(x_0), \quad \partial G(x_0) \neq D'(x_0).
\]

Hence, (2) holds. (3) is also evident because \( G_0(\bar{x}) = 0 \) whenever \( \bar{x} \in S \) and \( G_0 \) is the sup-function of \( \{g_i : i \in I \setminus I_0\} \) for some finite subset \( I_0 \) of \( I \).

Finally, (4) holds because
\[
N_K(x_0) = N_S(x_0) = \mathbb{R}, \quad N'(x_0) = (-\infty, 0].
\]

(b) Let \( g_0 \) be defined by \( g_0(x) = 0 \) for all \( x \in X \). We add this function \( g_0 \) to the system in (a) above. The new system satisfies the PLV relative to \( C \) at \( x_0 \) and the weak MFCQ relative to \( C \) at \( x_0 \). However, the weak SC on \( C \) does not hold; hence the BCQ relative to \( C \) at \( x_0 \) is not satisfied.

It is well known (see, e.g., [16]) that the following implication is not true in general:

\((\ast)\) Each finite subsystem of (1.1) satisfies the BCQ \( \implies \) the system (1.1) does.

Therefore the remainder of this section is devoted to studying the following natural question: When does this implication hold? The following theorem (which will be useful for our discussion in section 5) provides a set of conditions ensuring the above implication \((\ast)\). To facilitate the proof of this theorem, we recall the following proposition which was established independently by Deutsch [7] and Rubenstein [20] (see also [4]).

For a closed convex subset \( Z \) of \( X \), let \( P_Z \) denote the projection operator defined by
\[
P_Z(x) = \{y \in Z : \|x - y\| = d_Z(x)\},
\]
where \( d_Z(x) \) denotes the distance from \( x \) to \( Z \). Recall that the duality map \( J \) from \( X \) to \( 2^X^* \) is defined by
\[
J(x) := \{x^* \in X^* : \langle x^*, x \rangle = \|x\|^2, \|x^*\| = \|x\|\}.
\]
In fact, $J(x) = \partial \phi(x)$, where $\phi(x) := \frac{1}{2}\|x\|^2$. Thus a Banach space $X$ is smooth if and only if for each $x \in X$ the duality map is single-valued.

**Proposition 3.9.** Let $Z$ be a closed convex set in $X$. Then for any $x \in X$, $z_0 \in P_Z(x)$ if and only if for each $z \in Z$ and there exists $x^* \in J(x-z_0)$ such that $\langle x^*, z - z_0 \rangle \leq 0$ for any $z \in Z$, that is, $J(x-z_0) \cap N_Z(z_0) \neq \emptyset$. In particular, when $X$ is smooth, $z_0 \in P_Z(x)$ if and only if $z_0 \in Z$ and $J(x-z_0) \cap N_Z(z_0) \neq \emptyset$.

**Theorem 3.10.** Let $I$ be a compact metric space and let $C$ be a closed convex subset of $X$ such that $\dim C := l < +\infty$. Suppose that

(i) for each $x \in X$, the function $i \mapsto g_i(x)$ is continuous;

(ii) for any finite subset $J$ of $I$, the subsystem

$$g_j(x) \leq 0, \quad j \in J,$$

satisfies the BCQ relative to $C$;

(iii) for any finite subset $J$ of $I$ with $|J| \leq l$, the subsystem (3.18) satisfies the SC relative to $C$.

Then (1.1) satisfies the BCQ relative to $C$.

**Proof.** Since $I$ is compact and metrizable, there exists a sequence $\{I_k\}$ of subsets of $I$ such that

(a) each $I_k$ is finite;

(b) $I_k \subseteq I_{k+1}$ for $k = 1, 2, \ldots$;

(c) $\bigcup_{k=1}^{\infty} I_k = I$.

Set $J_k = \cap_{i \in I_k} \{x \in X : g_i(x) \leq 0\}$ for each $k$. Then

$$K = C \cap S \subseteq C \cap S_k.$$  

Let $x_0 \in K$ and $x^* \in N_K(x_0)$. We have to show that $x^* \in N_C(x_0) + N'(x_0)$. We will first show that there exist $\{x_k\} \subseteq C$ with $x_k \to x_0$ and $x_k^* \in N_C \cap S_k(x_k)$ such that $\{x_k\}$ is bounded and

$$\lim_{k \to \infty} \langle x_k^*, y \rangle = \langle x^*, y \rangle \quad \text{for each } y \in \text{span} (C-x_0).$$

In fact, since $Z := \text{span} (C-x_0)$ is of finite dimension, we may assume, without loss of generality, that the norm restricted on $Z$ is both strictly convex and smooth. Clearly, we may assume that $x^*|_Z \neq 0$. Take $z_0 \in Z$ such that $\langle x^*, z_0 \rangle = \|x^*\| \cdot \|z_0\| = \|z_0\|^2$. Write $x = x_0 + z_0$. Then, $x^*|_Z = J(x-x_0)|_Z$ and, by Proposition 3.9, $x_0 = P_K(x)$. Let $x_k = P_{C \cap S_k}(x)$. Then

$$\|x_k\| \leq \|x_k - x\| + \|x\| \leq \|x-x_0\| + \|x\|.$$ 

hence $\{x_k\} \subseteq C$ is bounded. Without loss of generality, assume that $x_k \to \bar{x}$. Let $i \in I$ and $\epsilon > 0$. By (b) and (c), there exists $\{i_k\} \subseteq I$ with $i_k \in I_k$ for each $k$ such that $i_k \to i$. By assumption (i), there exists an integer $k_0$ such that

$$g_i(\bar{x}) - g_{i_{k_0}}(\bar{x}) < \epsilon.$$

By (b), for each $k > k_0$, $i_{k_0} \in I_{k_0} \subseteq I_k$; hence, by the fact that $x_k \in S_k \subseteq S_{k_0}$,

$$g_{i_{k_0}}(\bar{x}) = g_{i_{k_0}}(\bar{x}) - g_{i_{k_0}}(x_k) + g_{i_{k_0}}(x_k) \leq g_{i_{k_0}}(\bar{x}) - g_{i_{k_0}}(x_k).$$

Consequently, by (3.21) and (3.22), we have that, for each $k > k_0$,

$$g_i(\bar{x}) < \epsilon + g_{i_{k_0}}(\bar{x}) - g_{i_{k_0}}(x_k).$$
Taking the limit as \( k \to \infty \) we get \( g_i(\bar{x}) \leq \epsilon \); hence \( \bar{x} \in K \) as \( \epsilon > 0 \) and \( i \in I \) are arbitrary. Because

\[
|\bar{x}-\bar{y}| = \lim_{k \to \infty} |x-k| \leq \|x-y\| \quad \text{for each } y \in K,
\]

\( \bar{x} = P_k(x) = x_0 \). Let \( x_k^* \in X^* \) with \( x_k^*|_Z = J(x-x_k)|_Z \) and \( \|x_k^*\| = \|x-x_k\| \). Then, by Proposition 3.9, \( x_k^* \in N_{C\cap S_k}(x_k) \). By the smoothness, the mapping \( x \mapsto J(x)|_Z \) is continuous and so (3.20) holds.

Now let \( G_k \) denote the sup-function corresponding to the subsystem (3.18) with \( J = I_k \); that is,

\[
G_k(x) := \sup_{i \in I_k} g_i(x) \quad \text{for each } x \in X.
\]

If there exists a subsequence \( \{k_j\} \) of \( \{k\} \) such that \( G_{k_j}(x_0) < 0 \) for each \( j \), then \( \sum_{i \in I_k \cap S_k} (x_{k_j}) = N_{C}(x_{k_j}) \), and hence \( x_{k_j}^* \in N_{C}(x_{k_j}) \). This implies that \( x^* \in N_{C}(x_{k_j}) \) by (3.20) (and so \( x^* \in N_{C}(x_0) + N'(x_0) \)). Therefore we may assume that, for each \( k \), \( x_k^* \notin N_{C}(x_k) \), and that \( G_k(x_k) = 0 \). Then \( I_k(x_k) := \{i \in I_k : g_i(x_k) = G_k(x_k)\} \neq \emptyset \).

By assumption (ii), \( x_k^* \in N_{C\cap S_k}(x_k) = N_{C}(x_k) + \text{cone} \{\partial g_i(x_k) : i \in I_k(x_k)\} \), and it follows from [19, Corollary 17.2.2] that there exist \( z_k^* \in N_{C}(x_k) \), \( i_{j_k}^* \in I_k(x_k) \), \( y_{i_{j_k}}^* \in \partial g_{i_{j_k}}(x_k) \), and \( \lambda_{j_k}^* \geq 0 \) (\( j = 1, 2, \ldots, l \)) such that

\[
x_k^* = z_k^* + \sum_{j=1}^l \lambda_{j_k}^* y_{i_{j_k}}^* \quad \text{on } Z, \quad k = 1, 2, \ldots.
\]

Let \( \lambda^k = \sum_{j=1}^l \lambda_{j_k}^* \). Then \( \{\lambda^k\} \) is bounded. Indeed, if not, by considering a subsequence if necessary, we have that \( \lim_{k \to \infty} \lambda^k = +\infty \). Thus \( \frac{z_k^*}{\lambda^k} \to 0 \) as \( k \to \infty \).

Furthermore, without loss of generality, we may assume that, as \( k \to \infty \),

\[
i_{j_k}^* \to i_j \quad \text{and} \quad \frac{\lambda_{j_k}^*}{\lambda^k} \to \lambda_j, \quad j = 1, 2, \ldots, l.
\]

Note in particular that \( \sum_{j=1}^l \lambda_j = 1 \). Since \( y_{i_{j_k}}^* \in \partial g_{i_{j_k}}(x_k) \subseteq \partial G(x_k) \) and since \( x_k \to x_0 \), \( \{y_{i_{j_k}}^*\} \) is bounded. Consequently, by (3.24), \( \{\tilde{z}_0^*, z \in Z\} \) is bounded too. Thus, we may also assume that there exist \( \tilde{z}_0^*, \tilde{y}_{i_{j_k}}^* \in Z_* \) such that

\[
\frac{z_k^*}{\lambda^k} \to \tilde{z}_0^*; \quad y_{i_{j_k}}^* \to \tilde{y}_{i_{j_k}}^* \quad \text{on } Z
\]

as \( k \to \infty \). Then

\[
\langle \tilde{z}_0^*, z - x_0 \rangle \leq 0 \quad \text{for each } z \in C
\]

and, by assumption (i), (3.25), and (3.26), for each \( j \),

\[
\langle \tilde{y}_{i_{j_k}}^*, z - x_0 \rangle \leq g_{i_{j_k}}(z) \leq g_{i_{j_k}}(z) - g_{i_{j_k}}(x_0) \quad \text{for each } z \in Z + x_0.
\]

Hence by the Hahn–Banach theorem, \( \tilde{z}_0^* \) and \( \tilde{y}_{i_{j_k}}^* \) can be extended to \( z_0^* \in N_C(x_0) \) and \( y_{i_{j_k}}^* \in \partial g_{i_{j_k}}(x_0) \). By (3.24), (3.25), and (3.26),

\[
0 = z_0^* + \sum_{j=1}^l \lambda_j y_{i_{j_k}}^* \quad \text{on } Z.
\]
By assumption (iii), take $\bar{y} \in C$ such that $g_{i}(\bar{y}) < 0$ for each $j = 1, 2, \ldots, l$. Consequently, $\langle y_{i}^{*}, \bar{y} - x_{0} \rangle < 0$ for each $j = 1, 2, \ldots, l$, and hence,

$$\left\langle z_{0}^{*} + \sum_{j=1}^{l} \lambda_{j} y_{i,j}^{*}, \bar{y} - x_{0} \right\rangle < 0,$$

which contradicts (3.29). Hence $\{\lambda^{k}\}$ is bounded. Thus, taking the limits on the two sides of (3.24) and using the similar arguments as above (if necessary, use subsequences), we get that

(3.30) $$x^{*} = z_{0}^{*} + \sum_{j=1}^{l} \lambda_{j} y_{i,j}^{*} \text{ on } Z$$

for some $\lambda_{j} \geq 0$, $z_{0}^{*} \in N_{C}(x_{0})$, and $y_{i,j}^{*} \in \partial g_{i_{j}}(x_{0})$ ($j = 1, 2, \ldots, l$). Note that, by (3.28), $y_{i,j}^{*} \in N_{C \cap g_{i_{j}}^{-1}(R_{-})}(x_{0})$ for each $j = 1, 2, \ldots, l$. Since $\langle y_{i,j}^{*}, \bar{y} - x_{0} \rangle < 0$, it follows that $y_{i,j}^{*} \in I$ for each $j = 1, 2, \ldots, l$. Let $y^{*} = x^{*} - z_{0}^{*} - \sum_{j=1}^{l} \lambda_{j} y_{i,j}^{*}$. Then $y^{*} \in N_{C}(x_{0})$ by (3.30), and so $x^{*} \in N_{C}(x_{0}) + N'(x_{0})$. The proof is complete.

Remark 3.2. Example 3.2 below will show that the condition dim $C < \infty$ cannot be dropped, while Examples 3.3 and 3.4 will show that conditions (i) and (iii) also cannot be dropped. In each of these examples, $I$ is a compact subset of $R$: $I = \{0, \frac{1}{2}, \ldots, \frac{1}{4}, \ldots\}$.

Example 3.2. Let $C = X = \{x = (x_{1}, x_{2}, \ldots, x_{k}, \ldots) : \lim_{k} x_{k} \text{ exists}\}$ with the norm defined by

$$\|x\| = \sup_{k} |x_{k}|, \quad x = (x_{k}) \in X.$$ 

Define

$$g_{i}(x) = \begin{cases} \lim_{k} x_{k}, & i = 0, \\ x_{i}, & i \in I \setminus \{0\}, \end{cases} \quad x = (x_{k}) \in X.$$ 

Since $|g_{i}(x) - g_{i}(y)| \leq \|x - y\|$ for each $i \in I$, the corresponding sup-function $G$ is continuous on $X$. Note that $G(\bar{x}) = -1 < 0$ for $\bar{x} = (-1) \in X$. Hence conditions (ii) and (iii) in Theorem 3.10 hold. Clearly, condition (i) in Theorem 3.10 is satisfied. Let $x_{0} = 0$. Then $x_{0} \in C \cap S = S$ and $I(x_{0}) = I$. It is easy to see that $N'(x_{0})$ is not closed; hence this system does not satisfy the BCQ at $x_{0}$.

Example 3.3. Let $C = X = R^{2}$. Define

$$g_{i}(x) = \begin{cases} x_{1} + x_{2}, & i = 0, \\ x_{1} + ix_{2}, & i \in I \setminus \{0\}, \end{cases} \quad x = (x_{1}, x_{2}) \in X.$$ 

Note that $|g_{i}(x) - g_{i}(y)| \leq 2 \|x - y\|$ for each $i \in I$; thus the corresponding sup-function $G$ is continuous on $X$. Let $x_{0} = 0$. Then $K := C \cap S = S = \{(x_{1}, x_{2}) : x_{1} \leq 0, x_{1} + x_{2} \leq 0\}, x_{0} \in bdS$, and $I(x_{0}) = I$. Since $G((-1, \frac{1}{2})) = -\frac{1}{2} < 0$, conditions (ii) and (iii) in Theorem 3.10 hold. However, $N_{K}(x_{0}) = N_{S}(x_{0}) = \{(t_{1}, t_{2}) : 0 \leq t_{2} \leq t_{1}\}$ and $N'(x_{0}) = \{(t_{1}, t_{2}) : 0 < t_{2} \leq t_{1}\} \cup \{(0, 0)\}$. Therefore this system does not satisfy the BCQ at $x_{0}$. Note that condition (i) is not satisfied.
Example 3.4. Let $C = X = \mathbb{R}^2$ and define

\[ g_i(x) = \begin{cases} x_2, & i = 1, \\ ix_1 - x_2 - i^2, & i \in I \setminus \{1\}, \end{cases} \quad x = (x_1, x_2) \in X. \]

It is easy to verify that $|g_i(x) - g_i(y)| \leq 2|x - y|$ for each $i \in I$; thus the corresponding sup-function $G$ is continuous on $X$. Let $x_0 = 0$. Then $I(x_0) = \{0, 1\}$ and $K = S = \{(x_1, 0) \in \mathbb{R}^2 : x_1 \leq 0\}$. Hence

\[ N_K(x_0) = \{(t_1, t_2) \in \mathbb{R}^2 : t_1 \geq 0\} \]

and

\[ N'(x_0) = \text{cone} \{(0, -1), (0, 1)\} = \{(t_1, t_2) \in \mathbb{R}^2 : t_1 = 0\}. \]

Consequently, this system does not satisfy the BCQ at $x_0$. Note that conditions (i) and (ii) in Theorem 3.10 are satisfied but condition (iii) is not.

4. Optimality conditions for minimization problems. The main purpose of this section is to apply our results of the earlier sections to provide optimality conditions for optimization problems of the following type. For a real-valued function $f$ on $X$, a typical problem is to

\begin{equation}
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad x \in C \cap S,
\end{align*}
\end{equation}

where $C$ and $S$ are as at the beginning of section 1. We will consider several kinds of functions on $X$. Let $\mathcal{A}(X)$ (resp., $\mathcal{C}(X)$) denote the class of all continuous real-valued affine (resp., convex) functions on $X$. Moreover, for $x_0 \in X$, let $\mathcal{G}_{x_0}(X)$ denote the class of real-valued functions $f$ on $X$ such that $f$ is Gateaux differentiable (with the Gateaux derivative $Df(x_0)$) at $x_0$ and the restriction of $f$ to $C$ is convex. Recall that $K := C \cap S$.

Theorem 4.1. Let $x_0 \in K$. Then the following statements are equivalent:

(i) The system (1.1) satisfies the BCQ relative to $C$ at $x_0$.

(ii) For any $f \in \mathcal{G}_{x_0}(X)$, $x_0$ is an optimal solution of (4.1) if and only if

\begin{equation}
Df(x_0) + x^* + \sum_{i \in I_0} \lambda_i x_i^* = 0 \quad \text{on} \quad X
\end{equation}

for some finite subset $I_0$ of $I(x_0)$, $x^* \in N_C(x_0)$, $\lambda_i \geq 0$, and $x_i^* \in \partial g_i(x_0)$ (for all $i \in I_0$).

(ii*) Same as (ii) but with (4.2) replaced by

\begin{equation}
Df(x_0) + x^* + \sum_{i \in I_0} \lambda_i x_i^* = 0 \quad \text{on} \quad \text{span } C.
\end{equation}

(iii) For any $f \in \mathcal{C}(X)$, $x_0$ is an optimal solution of (4.1) if and only if

\begin{equation}
\gamma^* + x^* + \sum_{i \in I_0} \lambda_i x_i^* = 0 \quad \text{on} \quad X
\end{equation}

for some finite subset $I_0$ of $I(x_0)$, $\gamma^* \in \partial f(x_0)$, $x^* \in N_C(x_0)$, $\lambda_i \geq 0$, and $x_i^* \in \partial g_i(x_0)$ (for all $i \in I_0$).
(iii*) Same as (ii*) but with (4.4) is replaced by

\[ y^* + x^* + \sum_{i \in I_0} \lambda_i x_i^* = 0 \text{ on span } C. \]

(iv) Same as (iii) but with \( f \in A(X) \).

(iv*) Same as (iii*) but with \( f \in A(X) \).

Proof. Since each of the statements in the above list is true (with \( I_0 = \emptyset \)) when \( x_0 \in C \cap \text{int } S \), we assume henceforth that \( x_0 \in C \cap \text{bd } S \).

(i) \implies (ii) \implies (ii*). Let \( f \in g_{x_0}(X) \) define

\[ F(x) = \langle Df(x_0), x - x_0 \rangle + I_K(x), \quad x \in X. \]

Then, by [6, Corollary 1, p. 105]

\[ \partial F(x_0) = Df(x_0) + N_K(x_0). \]

It follows from (i) that

\[ \partial F(x_0) = Df(x_0) + N_C(x_0) + N'(x_0). \]

If \( x_0 \) is an optimal solution of (4.1), \( x_0 \) is a minimal point of \( F \) on \( X \). In fact, it is clear that \( F(x) \geq F(x_0) \) for any \( x \in X \setminus K \). Moreover, if \( x \in K \), then

\[ F(x) = \langle Df(x_0), x - x_0 \rangle = \lim_{t \to 0^+} \frac{f(x_0 + t(x - x_0)) - f(x_0)}{t} \geq 0 = F(x_0) \]

because \( x_0 + t(x - x_0) \in K \) for each \( 0 < t < 1 \) (and \( x_0 \) is an optimal solution of (4.1)). Hence, by [6, Proposition 2.4.11], \( 0 \in \partial F(x_0) \). Consequently, (4.7) entails (4.2) (and (4.3)) for appropriate \( x^*, \lambda_i, x_i^* \) as stated in (ii) and (ii*). Conversely, suppose that (4.3) holds for some \( I_0, x^*, \lambda_i, x_i^* \) as stated in (ii*). Writing \( z^* \) for \( x^* + \sum_{i \in I_0} \lambda_i x_i^* \), one has from (2.11) that \( z^* \in N_K(x_0) \) and from (4.3) that

\[ \langle Df(x_0), \cdot \rangle = \langle -z^*, \cdot \rangle \text{ on span } C. \]

Therefore, by convexity, we have for each \( x \in K \) that

\[ f(x) - f(x_0) \geq \langle Df(x_0), x - x_0 \rangle = \langle -z^*, x - x_0 \rangle \geq 0. \]

Thus the implications (i) \implies (ii) and (ii*) are proved.

The implication (i) \implies (iii) is a direct consequences of [6, Corollary, p. 52]. The necessity part of (iii*) follows from (iii) while the sufficiency part of (iii*) is proved similarly to the proof of the corresponding part of (ii*).

Further, since the implications (ii) \implies (iv), (ii*) \implies (iv*), (iii) \implies (iv), and (iii*) \implies (iv*) are trivial, it remains to show that (iv) \implies (i) and (iv*) \implies (i). To show these, let \( y^* \in N_K(x_0) \). Then \( -y^* \in A(X) \) and \( x_0 \) is an optimal solution of (4.1) with \( f := -y^* \). Thus \( \partial f(x_0) = -y^* \) and, assuming (iv) or (iv*), there exist a finite subset \( I_0 \) of \( I(x_0) \), \( x^* \in N_C(x_0) \), \( \lambda_i > 0 \), and \( x_i^* \in \partial g_1(x_0) \) (\( i \in I_0 \)) such that

\[ -y^* + x^* + \sum_{i \in I_0} \lambda_i x_i^* = 0 \text{ on span } C, \]

that is,

\[ \langle y^*, \cdot \rangle = \langle z^*, \cdot \rangle \text{ on span } C, \]
where \( z^* := x^* + \sum_{i \in I_0} \lambda_i x_i^* \). Since \( x_0 \in C \), it follows that \( y^* - z^* \in N_C(x_0) \). Consequently,

\[
y^* \in N_C(x_0) + x^* + \sum_{i \in I_0} \lambda_i x_i^* \subseteq N_C(x_0) + N'(x_0),
\]

and (i) holds. This completes the proof. \( \square \)

5. Best restricted range approximation problem in \( L^p \) spaces. Let \( \Omega := T \cup A \), where \( A \) is a compact Hausdorff space and \( T \) is a measure space with a measure \( \mu \). As usual, let \( C_R(A) \) denote the Banach space of all real-valued continuous functions on \( A \) equipped with the uniform norm denoted by \( \| \cdot \|_A \).

Let \( X \) be the space consisting of all real functions \( x \) on \( \Omega \) such that the restrictions \( x|_T \in L_p(T, \mu) \) and \( x|_A \in C_R(A) \), where \( p \in [1, +\infty) \). Then \( X \) is a Banach space under the norm \( \| \cdot \| \) defined by

\[
\| x \| = \| x|_T \|_p + \| x|_A \|_A \quad \text{for each } x \in X.
\]

For the so-called restricted range approximation problem in \( L^p \) spaces (studied by Levasseur and Levis in [10]), one is given the following data: Let \( l \) and \( u \) be a pair of real functions which are, respectively, upper and lower semicontinuous on \( A \) such that \( l(t) \leq u(t) \) for each \( t \in A \). Let \( L \) be a finite dimensional vector subspace of \( X \) generated by \( h_1, h_2, \ldots, h_n \). Let \( z \in X \) be given. Then the problem to be considered is

\[
(5.1) \quad \text{minimize } \int_T |z(t) - x(t)|^p \, d\mu
\]

subject to

\[
(5.2) \quad x \in L, \quad l(t) \leq x(t) \leq u(t) \quad \text{for all } t \in A
\]

(one can of course replace \( \mu \) by \( \mu' \) of the form \( d\mu' = w \, d\mu \) for some positive \( w \in L^\infty(T, \mu) \)).

The aim of this section is to apply results of the preceding section to establish the following theorem, which was proved for the case when \( p = 2 \) (or \( p \) is an even integer) in [10] using a very different method and under assumptions \( H_1 - H_6 \) given in the following:

- \( H_1 \): \( A \) is a compact subset of \( \mathbb{R} \) and \( T \) is a finite union of compact intervals of \( \mathbb{R} \).
- \( H_2 \): \( z \) is a real continuous function on \( \Omega \).
- \( H_3 \): \( l \) and \( u \) are continuous on \( A \) with \( l(t) < u(t) \) for all \( t \in A \).
- \( H_4 \): \( \{h_1, h_2, \ldots, h_n\} \) is linearly independent on \( T \).
- \( H_5 \): There exists \( h \in X \) such that (5.2) is satisfied with \( x = h \).
- \( H_6 \): For each \( k = 1, 2, \ldots, n \), the set \( \{h_1, h_2, \ldots, h_k\} \) is a Chebyshev system of order \( k \) on \( A \); that is, if \( t_1, t_2, \ldots, t_k \) are distinct points in \( A \), then the determinant \( \det(h_j(t_i)) \neq 0 \).

Instead of assumptions \( H_1 - H_6 \), we consider the following two conditions:

- \( D_1 \): There exists an element \( h_0 \in L \) satisfying \( l(t) < h_0(t) < u(t) \) for all \( t \in A \).
- \( D_2 \): \( A \) is metrizable (in addition to being compact Hausdorff), \( l \) and \( u \) are continuous on \( A \), and, for any \( k \leq n \) and any distinct points \( t_1, t_2, \ldots, t_k \) in \( A \), there exists an element \( h \in L \) satisfying \( l(t_i) < h(t_i) < u(t_i) \) for all \( i = 1, 2, \ldots, k \).

We remark that \( H_1 + H_6 \Rightarrow D_2 \).
We use $Z(x)$ to denote the set of all points in $T$ such that $x(t) = 0$. Note that if $x = \tilde{x}$ a.e. on $T$, then the “symmetric difference” $(Z(x) \setminus Z(\tilde{x})) \cup (Z(\tilde{x}) \setminus Z(x))$ is of $\mu$-measure zero; hence $\int_{Z(x)} h(t) d\mu = \int_{Z(\tilde{x})} h(t) d\mu$ for each integrable function $h$ on $T$.

We shall need the following well-known results; see, for example, [3, Example 2.19] for a proof of Lemma 5.1. The other two lemmas follow immediately from Lemma 5.1 and the Riesz representation theorem.

**Lemma 5.1.** Let $W$ be a normed vector space. Then, for any $w \in W \setminus \{0\}$, the subdifferential $\partial \| \cdot \|_1(w)$ of the norm function at $w$ equals the set

$$\{ w^* \in W^* : \langle w^*, w \rangle = \|w\| \}.$$

**Lemma 5.2.** Let $W = L^1(T, \mu)$, where here and throughout, $T$ denotes a measure space with measure $\mu$. Let $w \in W \setminus \{0\}$. Then, as a subset of $L^\infty(T)$, $\partial \| \cdot \|_1(w)$ consists of all $\beta \in L^\infty(T)$ satisfying the properties that

$$\left\{ \begin{array}{l}
|\beta(t)| \leq 1 \quad \text{a.e. on } T, \\
\beta(t) = \text{sgn}[w(t)] \quad \text{a.e. on } T \setminus Z(w).
\end{array} \right.$$ 

**Lemma 5.3.** Let $W = L^p(T, \mu)$ with $p \in (1, +\infty)$. Let $w \in W \setminus \{0\}$. Then, the norm function is Fréchet differentiable at $w$ with the derivative $\| \cdot \|_p^*(w)$, which is represented by $\xi \in L^q(T, \mu)$ ($\frac{1}{p} + \frac{1}{q} = 1$) with

$$\xi(t) = \|w\|^{1-p} \cdot |w(t)|^{p-1} \cdot \text{sgn}[w(t)] \quad \text{a.e. on } T.$$ 

In particular, $\partial \| \cdot \|_p(w) = \{ \xi \}$.

Our main result in this section is now stated as follows.

**Theorem 5.4.** Suppose that $D_1$ or $D_2$ holds. Let $x_0 \in X$ satisfy (5.2) and let $z \in X$. Then $x_0$ is an optimal solution of the minimization problem (5.1) subject to (5.2) if and only if there exist a finite subset $\{t_1, t_2, \ldots, t_k\}$ of $A$ and a finite subset $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$ of $\mathbb{R}$ ($0 \leq k \leq n$) with the following properties:

(a) $x_0(t_i) = u(t_i)$ or $x_0(t_i) = l(t_i)$ for each $i = 1, 2, \ldots, k$;

(b) for each $i = 1, 2, \ldots, k$,

$$\text{sgn} \lambda_{t_i} = \left\{ \begin{array}{cl} 1 & \text{if } x_0(t_i) = u(t_i), \\
-1 & \text{if } x_0(t_i) = l(t_i);
\end{array} \right.$$ 

(c) for each $j = 1, 2, \ldots, n$,

$$\int_T |z(t) - x_0(t)|^{p-1} \text{sgn}[z(t) - x_0(t)] h_j(t) d\mu = \sum_{i=1}^k \lambda_{t_i} h_j(t_i), \quad p > 1; \quad \tag{5.3}$$

$$\int_T \text{sgn}[z(t) - x_0(t)] h_j(t) d\mu + \int_{Z(z - x_0)} \beta(t) h_j(t) d\mu = \sum_{i=1}^k \lambda_{t_i} h_j(t_i), \quad p = 1, \quad \tag{5.4}$$

for some $\beta \in L^\infty$ with $|\beta(t)| \leq 1$ a.e. on $Z(z - x_0)$. (Note: we adopt the convention that, if $k = 0$, the sums of the right-hand sides of (5.3) and (5.4) are read as zero.)

In order to prove this theorem, we first do some preparation. We make two homeomorphic copies of $A$: one is still denoted by $A$ while the other is denoted by $\tilde{A} = \{ \tilde{a} : a \in A \}$, where $a \mapsto \tilde{a}$ is a bijection. Let $I = A \cup \tilde{A}$. Thus $I$ can be
topologized such that $A$, $\hat{A}$ are two disjoint compact subsets of $I$. For each $t \in A$, define $g_t : X \to \mathbb{R}$ by

\begin{equation}
(5.5) \quad g_t(x) = e_t(x) - u(t), \quad x \in X,
\end{equation}

where $e_t : X \to \mathbb{R}$ is the “unit point functional” at $t$ defined by

\begin{equation}
(5.6) \quad e_t(x) = x(t), \quad x \in X.
\end{equation}

Similarly, for each $\tilde{t} \in \hat{A}$, define $g_{\tilde{t}} : X \to \mathbb{R}$ by

\begin{equation}
(5.7) \quad g_{\tilde{t}}(x) = l(\tilde{t}) - e_{\tilde{t}}(x), \quad x \in X.
\end{equation}

Thus $\{g_i : i \in I\}$ is a family of affine functions on $X$ such that, for each $i \in I$,

\begin{equation}
(5.8) \quad |g_i(x) - g_i(y)| \leq \|x - y\|, \quad x, y \in X.
\end{equation}

In particular, $\{g_i : i \in I\}$ is equicontinuous on $X$, and the corresponding sup-function $x \mapsto G(x) := \sup\{g_i(x) : i \in I\}$ is continuous. Note furthermore that the function $i \mapsto g_i(x)$ is upper semicontinuous for each $x \in X$. Clearly, $x \in X$ satisfies (5.2) if and only if $x \in L$ and $g_i(x) \leq 0$ for each $i \in I$. Assuming $D_1$, it is easy to verify that $h_0 \in L$ given in $D_1$ satisfies the strict inequality $G(h_0) < 0$ (thanks to the compactness of $A$); thus $h_0$ is a Slater point for the system $g_i(\cdot) \leq 0$, $i \in I$. Thus by Corollary 3.7, this system satisfies the BCQ relative to $L$ at each point $x$ in $L \cap S$, where $S := \{x \in X : g_i(x) \leq 0, i \in I\}$ (the assertion being trivial if $x \in L \cap \text{int} S$). Similarly, if $D_2$ holds, then it follows from Theorem 3.10 that this system satisfies the BCQ relative to $L$ at each point $x$ in $L \cap S$.

Note that if $x \in \text{bd} S$, then $G(x) = 0$, and $I(x)$ defined in (1.7) is equal to the set

\begin{equation}
(5.9) \quad I(x) = \{t \in A : x(t) = u(t)\} \cup \{\tilde{t} \in \hat{A} : x(t) = l(t)\}
\end{equation}

whenever $x \in \text{bd} S$. Note that, since $l(\cdot) < u(\cdot)$ on $A$, $t \in I(x)$ implies $\tilde{t} \notin I(x)$; similarly, $\tilde{t} \in I(x)$ implies $t \notin I(x)$. For easy reference, we record the following trivial fact:

\begin{equation}
(5.10) \quad g_i(t) = \begin{cases} 
  e_t, & i = t \in A, \\
  -e_{\tilde{t}}, & i = \tilde{t} \in \hat{A}.
\end{cases}
\end{equation}

Now define $f : X \to \mathbb{R}$ by

\begin{equation}
(5.11) \quad f(x) = \left( \int_T |z(t) - x(t)|^p \, d\mu \right)^{\frac{1}{p}} \quad \text{for each } x \in X.
\end{equation}

Then $f$ is a continuous convex function on $X$, and $x_0$ is a solution of the minimization problem (5.1) subject to (5.2) if and only if it is a solution of the problem of minimizing $f$ subject to

\begin{equation}
(5.12) \quad x \in L, \quad g_i(x) \leq 0 \quad \text{for each } i \in I,
\end{equation}

where $\{g_i : i \in I\}$ is defined by (5.5) and (5.6).

Since $f(x)$ defined in (5.10) depends only on the restriction of $x$ to $T$, the following result follows from Lemmas 5.2 and 5.3 together with the chain rule.
Proposition 5.5. Suppose that \((z - x_0)|_T\) is not the zero-element of \(L^p(T, \mu)\). Then the following assertions hold:

(a) The case when \(p = 1\). A functional \(y^* \in \partial f(x_0)\) if and only if there exists \(\beta \in L^\infty(T, \mu)\) such that

\[
|\beta(t)| \leq 1 \quad \text{a.e. on } T,
\]

\[
\beta(t) = \text{sgn}[z(t) - x_0(t)] \quad \text{a.e. on } T \setminus Z(z - x),
\]

and

\[
\langle y^*, x \rangle = -\int_T x(t)\beta(t)d\mu \quad \text{for each } x \in X.
\]

(b) The case when \(p \in (1, +\infty)\). A functional \(y^* \in \partial f(x_0)\) if and only if

\[
\langle y^*, x \rangle = -c\int_T x(t)|z(t) - x_0(t)|^{p-1}\text{sgn}[z(t) - x_0(t)]d\mu \quad \text{for each } x \in X,
\]

where \(c := \|(z - x_0)|_T\|_p^{1-p}\).

Proof of Theorem 5.4. Let \(f, g_i, I, \) and \(S\) be the same as those introduced after the statement of Theorem 5.4. Since the proof for the case when \(x_0 \in L \cap \text{int } S\) is easier (in this case one can take \(k = 0\), and the right-hand sides of (5.3) and (5.4) are replaced by zero), it suffices to consider the case when \(x_0 \in L \cap \text{bd } S\). As already noted, by assumption \(D_1\) or \(D_2\), the system of inequalities \(g_i(\cdot) \leq 0, i \in I,\) satisfies the BCQ at \(x_0\) relative to \(L\). Noting that \(x^* = 0\) on \(L\) whenever \(x^* \in N_L(x_0)\), it follows from Theorem 4.1 that the following statements are equivalent:

(i) \(x_0\) is a minimizer of \(f\) subject to (5.11).

(ii) There exist a finite subset \(I_0\) of \(I(x_0)\), \(y^* \in \partial f(x_0), \mu_i > 0, \) and \(x^*_i \in \partial g_i(x_0)\) (for all \(i \in I_0\)) such that

\[
y^* + \sum_{i \in I_0} \mu_i x^*_i = 0 \quad \text{on } L.
\]

(ii*) Same as (ii) but with (5.16) replaced by

\[
cy^* + \sum_{i \in I_0} \mu_i x^*_i = 0 \quad \text{on } L,
\]

where \(c\) is a positive constant.

In view of (5.8) and (5.9), we note that, in (ii) and (ii*), each \(x^*_i\) can be expressed as either \(e_t\) or \(-e_t\):

\[
x^*_i = \begin{cases} 
e_t & \text{if } i = t \in I(x_0) \cap A \text{ (i.e., if } x_0(t) = u(t)), \\
-e_t & \text{if } i = t \in I(x_0) \cap A^c \text{ (i.e., if } x_0(t) = l(t)).
\end{cases}
\]

Thus \(\mu_i x^*_i\) can be expressed in the form \(\lambda_i e_t\) for some \(\lambda_i \in \mathbb{R}\) and some \(t_i \in A\) satisfying \(x_0(t_i) = u(t_i)\) or \(x_0(t_i) = l(t_i)\). Explicitly, \(\lambda_i\) is defined by

\[
\lambda_i = \begin{cases} 
\mu_i & \text{if } x_0(t_i) = u(t_i) \text{ and } t_i \in I_0, \\
-\mu_i & \text{if } x_0(t_i) = l(t_i) \text{ and } t_i \in I_0.
\end{cases}
\]

Moreover, when (ii) (or (ii*)) holds, we can suppose \(I_0\) satisfies an additional property that \(|I_0|\), the number of elements of \(I_0\), is not greater than \(\dim L\), the dimension of
L. To see this, we suppose, without loss of generality, that \(h_1, h_2, \ldots, h_n\) are linearly independent, so \(\dim L = n\). Let \(V\) denote the subset of \(\mathbb{R}^n\) defined by

\[
\{(x^*_i, h_1), (x^*_i, h_2), \ldots, (x^*_i, h_n) : i \in I_0\}
\]

and \(U := V \cup \{(y^*, h_1), (y^*, h_2), \ldots, (y^*, h_n)\}\). Since \(L\) is spanned by \(h_1, h_2, \ldots, h_n\), (ii) implies that \(0 \notin \text{co} \, U\). Note that

\[
(x^*_i, h_0) < 0 \quad \text{for each} \quad i \in I_0.
\]

Indeed, if \(i = \tilde{t} \in I_0\) with some \(t \in A\), then \(x_0(t) = l(t)\) and

\[
(x^*_i, h_0) = (-e_t, h_0) = -h_0(t) < -l(t) = (-e_t, x_0) = (x^*_i, x_0);
\]

thus (5.18) is true, as the case when \(i = t \in I_0\) with some \(t \in A\) can be considered similarly. Since \(h_0 - x_0 \in L\) and \(L\) is spanned by \(h_1, h_2, \ldots, h_n\), it follows from (5.18) that \(0 \notin \text{co} \, V\); thus \(0 \notin \text{co} \, U \setminus \text{co} \, V\).

By the Carathéodory theorem (cf. [5]), there exist \(\bar{\mu}_0 > 0, \bar{\mu}_{ij} \geq 0 \, (j = 1, 2, \ldots, n)\) with \(\bar{\mu}_0 + \sum_{j=1}^n \bar{\mu}_{ij} = 1\) and \(i_1, i_2, \ldots, i_n \in I_0\) such that

\[
\bar{\mu}_0 y^* + \sum_{j=1}^n \mu_{ij} x^*_j = 0 \quad \text{at} \quad h_1, h_2, \ldots, h_n.
\]

Therefore, \(I_0\) in (ii) can be replaced by its subset with at most \(n\) elements. Making use of Proposition 5.5 and noting that \(L\) is spanned by \(h_1, h_2, \ldots, h_n\), we now see that (ii) (or (ii*)) holds if and only if (a), (b), and (c) of Theorem 5.4 hold for some finite subsets \(\{i_1, i_2, \ldots, i_k\} \subseteq A\) and \(\{\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_k}\} \subseteq \mathbb{R}\), where \(k \leq n\). Finally, as we already noted, (i) holds if and only if \(x_0\) is an optimal solution of the minimization problem (5.1) subject to (5.2). This completes the proof of Theorem 5.4. \(\square\)

6. Approximation problem with constraints by conditionally positive semidefinite functions. This problem is studied in a recent paper [22] by Strauss. Following his approach, let \(Q\) be a nonempty subset of a Euclidean space \(\mathbb{R}^m\) and let \(F(Q)\) denote the class of all real-valued functions on \(Q\). Let \(H\) be a vector subspace of \(F(Q)\) and suppose that \(H\) is equipped with a seminorm induced by a semi-inner product \(\langle \cdot, \cdot \rangle\) on \(H \times H\). Let \(U = \text{span} \{u_1, u_2, \ldots, u_m\}\) be an \(m\)-dimensional subspace of \(H\), and let \(K\) be a reproducing kernel with respect to \(U\) (for undefined terms, see [1, 18]). Let \(A = \{z_1, z_2, \ldots, z_n\} \subseteq Q\) and let \(C_A\) denote the \((m \times n)\)-matrix defined by

\[
C_A = (u_j(z_i))_{j=1}^m_{i=1}.
\]

Let \(W_n, W_m\), respectively, denote an \((n \times n)\)-matrix and an \((m \times m)\)-matrix; we assume that \(W_n\) and \(W_m\) are positive semidefinite. Define a seminorm \(\| \cdot \|\) on \(\mathbb{R}^n \times \mathbb{R}^m\) by

\[
\|(a, b)\|_W^2 = a^T W_n a + b^T W_m b \quad \text{for all} \quad a \in \mathbb{R}^n, b \in \mathbb{R}^m.
\]

Suppose that we are given the following data: \(h \in H, \{\xi_i : i \in I\} \subseteq \mathbb{R}, \{p_i : i \in I\} \subseteq \mathbb{R}^n, \text{and} \{q_i : i \in I\} \subseteq \mathbb{R}^m, \text{where} \ I \text{is an index set. Let}

\[
Z_I = \{(a, b) : a \in \mathbb{R}^n, b \in \mathbb{R}^m, C_A a = 0, p_i^T a + q_i^T b \leq \xi_i, i \in I\}
\]
and let $\hat{d}: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be defined by

$$
\hat{d}(a, b) = \frac{1}{2} \left( \left\| h - \sum_{i=1}^{n} a_i K(\cdot, z_i) - \sum_{j=1}^{m} b_j u_j \right\|^2 + \left\| (a, b) \right\|_{W}^2 \right),
$$

where $a = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ and $b = (b_1, b_2, \ldots, b_m) \in \mathbb{R}^m$. Under the obvious modifications, $Z_I$ and $\hat{d}$ can be defined when $m = 0$, and our discussions below are valid for this case as well as for the case when $m \neq 0$.

The problem to be considered here is

$$
\text{minimize } \hat{d}(a, b) \text{ subject to } (a, b) \in Z_I.
$$

Define

$$
U_A^\perp = \{ (a, b) \in \mathbb{R}^n \times \mathbb{R}^m : C_A a = 0 \},
$$

$$
K_A = (K(z_i, z_j))_{i,j=1}^{n},
$$

$$
h_A = (h(z_1), \ldots, h(z_n))^T,
$$

and

$$
\hat{e}(a, b) = \frac{1}{2} (a^T (K_A + W_n) a + b^T W_m b) - h_A^T a + \frac{1}{2} \langle h, h \rangle.
$$

By Proposition 2.3 and the proof of Proposition 4.1 in [22], we have

$$
\hat{d}(a, b) = \hat{e}(a, b) \text{ for all } (a, b) \in U_A^\perp
$$

and

$$
K_A \text{ is positive semidefinite with respect to } U_A^T
$$

in the sense that $a^T K_A a \geq 0$ whenever $(a, b) \in U_A^\perp$. This implies of course that $K_A + W_n$ is also positive semidefinite, and so the quadratic function $\hat{e}$ is convex on $U_A^\perp$.

We always assume that (6.1) has a solution: let $(a_0, b_0)$ denote a solution of (6.1). Let $r(C_A)$ denote the rank of the matrix $C_A$. Finally, $|J|$ denotes the number of elements of a finite set $J$. The following result was asserted as a theorem ([22, Theorem 6.2]):

Suppose that $(a_0, b_0)$ is a solution of (6.1), $r(C_A) = m$, and that there exists $(\bar{a}, \bar{b}) \in \mathbb{R}^n \times \mathbb{R}^m$ such that

$$
c_a \bar{a} = 0, \quad p_i^T \bar{a} + q_i^T \bar{b} < \xi_i \quad \text{for all } i \in I.
$$

Then there exists a finite subset $I_0$ of $I$ with $|I_0| \leq n$ such that $(a_0, b_0)$ also minimizes $\hat{d}(a, b)$ subject to $(a, b) \in Z_{I_0}$.

While we will give a counterexample below to show this result is false, we put forward the following corrected version.
Theorem 6.1. Let \((a_0, b_0)\) be a solution of (6.1). Assume that the system of linear inequalities

\[
\sum_{i \in I} p_i^T a + q_i^T b \leq \xi_i, \quad i \in I,
\]

satisfies the BCQ relative to \(U_A^\perp\) at \((a_0, b_0)\). Then there exists a finite subset \(I_0\) of \(I\) with \(|I_0| \leq m + n - r(C_A)\) such that \((a_0, b_0)\) also minimizes \(d(a, b)\) subject to \((a, b) \in U_A^\perp\).

Proof. By assumption and (6.6), \((a_0, b_0)\) is a solution of the following minimization problem:

Minimize \(\bar{d}(a, b)\) subject to \((a, b) \in U_A^\perp\) and (6.9).

We shall use \(D\bar{e}(a_0, b_0)|_{U_A^\perp}\) to denote the restriction to \(U_A^\perp\) of the Gateaux (equivalently, Fréchet) derivative of the function \(\bar{e}\) at \((a_0, b_0)\). Since, by assumption, the system (6.9) in \(X := \mathbb{R}^n \times \mathbb{R}^m\) satisfies the BCQ relative to \(U_A^\perp\) at \((a_0, b_0)\), one can apply the implication \((i) \Rightarrow (ii)\) of Theorem 4.1 to conclude that there exist a finite subset \(I_0\) of \(I(a_0, b_0)\) and a collection \(\{\lambda_i' : i \in I_0\}\) of nonnegative real numbers such that

\[
D\bar{e}(a_0, b_0)|_{U_A^\perp} + \sum_{i \in I_0} \lambda_i'(p_i^T, q_i^T)|_{U_A^\perp} = 0,
\]

because \(x^*|_{U_A^\perp} = 0\) whenever \(x^* \in N_{U_A^\perp}(a_0, b_0)\). We write \(\mu\) for the sum \(\sum_{i \in I_0} \lambda_i'\) and assume that \(\mu \neq 0\) (otherwise, \((a_0, b_0)\) must be a minimal point of \(\bar{e}(= \bar{d})\) on \(U_A^\perp\) by convexity of \(\bar{e}\) on \(U_A^\perp\)). Thus

\[
-D\bar{e}(a_0, b_0)|_{U_A^\perp} \in \text{cone}\{(p_i^T, q_i^T)|_{U_A^\perp} : i \in I_0\}.
\]

Since \(\dim U_A^\perp = n + m - r(C_A)\), it follows from [19, Corollary 17.1.2] that \(I_0\) in (6.11) can be replaced by a subset \(I_0\) with \(|I_0| \leq n + m - r(C_A)\). Note that (6.10) continues to hold when \(I_0\) is replaced by \(I_0\) (replace the family \(\{\lambda_i'\}\) by a suitable new one if necessary). Consequently, by \((ii^*)\) of Theorem 4.1, \((a_0, b_0)\) minimizes \(\bar{e}(a, b)\) on \(U_A^\perp\) subject to the system of linear inequalities

\[
p_i^T a + q_i^T b \leq \xi_i, \quad i \in I_0.
\]

This completes the proof as before. 

Example 6.1. Let \(A = Q\) consist of two distinct elements \(z_1, z_2\). We identify \(F(Q)\) with \(\mathbb{R}^2\) in the obvious manner. Let \(H = F(Q)\) with the usual inner product. Take \(h = (1, 0)^T \in H, m = 0, W_m = 0, U = \{0\}, C_A = 0,\) and \(U_A^\perp = \mathbb{R}^2\). Define \(W_n\) by \(W_n = (\frac{2}{3}, \frac{4}{5})\) and fix the reproducing kernel \(K : Q \to \mathbb{R}\) defined by

\[
K(x, y) = \begin{cases} 
1 & \text{if } x = y, \\
0 & \text{if } x \neq y.
\end{cases}
\]

Take \(I = \{1, 2, \ldots\}, \xi_i = 0,\) and \(p_i = (1, \frac{1}{2})\) for \(i \in I\). Then, (6.8) is satisfied for appropriate \(\tilde{a}\). Note

\[
\tilde{d}(a) := \tilde{d}(a, b) = \frac{1}{2} \left[ \|h - a\|^2 + \frac{1}{4} |a|^2 \right]
\]
and
\[ Z_I = \left\{ a \in \mathbb{R}^2 : a_1 + \frac{1}{i} a_2 \leq 0 \text{ for all } i \in I \right\}. \]

Clearly, \( Z_I \) is the convex set of \( \mathbb{R}^2 \) containing the 4th quadrant and bounded by the two half-lines, respectively, contained in \( a_1 + a_2 = 0 \) and \( a_1 = 0 \). It follows easily that the origin \( 0 \in \mathbb{R}^2 \) is the nearest point to \( h \) from \( Z_I \). On the other hand, if \( I_0 \) is a finite subset of \( I \), then \( Z_{I_0} \) is also bounded by the two half-lines, respectively, contained in \( a_1 + \frac{1}{i} a_2 = 0 \) and \( a_1 + \frac{1}{i} a_2 = 0 \), where \( i_- := \min I_0, \ i_+ := \max I_0 \).

Let \( a_{I_0} \) denote the nearest point to \( h \) from \( Z_{I_0} \). Then \( a_{I_0} = \frac{1}{(1 + i_+^2)} (1, -i_+)^T \) and 
\[ \|h\|^2 = \|a_{I_0}\|^2 + \|h - a_{I_0}\|^2. \]
This implies that
\[ \frac{1}{2} \|h\|^2 > \frac{1}{2} \left[ \frac{1}{4} \|a_{I_0}\|^2 + \|h - a_{I_0}\|^2 \right] = \hat{d}(a_{I_0}), \]
while
\[ \hat{d}(0) = \frac{1}{2} \|h - 0\|^2 \leq \frac{1}{2} \|h - a\|^2 \leq \hat{d}(a) \]
for all \( a \in Z_I \).

Therefore [22, Theorem 6.2] mentioned earlier is not true.

**Remark 6.1.** Let \( l_i \) be the function on \( \mathbb{R}^n \times \mathbb{R}^m \) defined by
\[ l_i(a, b) = p_i^T a_i + q_i^T b_i - \xi_i, \quad (a, b) \in \mathbb{R}^n \times \mathbb{R}^m. \]

Suppose that \( I \) can be topologized in such a way that it is a compact space satisfying the first axiom of countability and that \( i \mapsto l_i(a, b) \) is upper semicontinuous on \( I \) for each \( (a, b) \in U_{I_0}^A \). Then, by Corollary 3.7, the linear inequality system (6.9) satisfies the BCQ relative to \( U_{I_0}^A \) at \( (a_0, b_0) \) if (6.8) is satisfied. In this particular case, the conclusion of [22, Theorem 6.2] mentioned earlier does hold. Similarly, if \( I \) can be topologized in such a way that it is a compact metric space and that \( i \mapsto l_i(a, b) \) is continuous on \( i \in I \) for each \( (a, b) \in U_{I_0}^A \), then, by Theorem 3.10, the aforesaid result is also valid even if (6.8) is replaced by a weaker condition given in the following: for each subset \( J \) of \( I \) with \( |J| \leq m + n - r(C_A) \), there exists \( (\bar{a}, \bar{b}) \in \mathbb{R}^n \times \mathbb{R}^m \) such that
\[ C_A \bar{a} = 0, \quad p_i^T \bar{a} + q_i^T \bar{b} < \xi_i \quad \text{for all } i \in J. \]

**Acknowledgments.** We wish to thank two referees for their valuable comments and suggestions.

**REFERENCES**


