Best Approximation and Perturbation Property in Hilbert Spaces[1]

Y. R. He[2] K. F. Ng[3]

Abstract. We study perturbation property of best approximation to a set defined by an abstract nonlinear constraint system. We show that at a normal point, perturbation property of best approximation is equivalent to an equality expressed in terms of normal cones. This equality is related to the strong conical hull intersection property. Our results generalize many known results in the literature on perturbation property of best approximation established for a set defined by a finite system of linear/nonlinear inequalities. The connection to minimization problem is also considered.

Key Words. Best approximation, abstract nonlinear system, strong CHIP, the Robinson constraint qualification.

1. Introduction

In recent years, there has been much interest centered on the following constraint approximation problem

$$\min \|x-y\| \text{ subject to } y \in C \text{ and } G(y) \in K, \tag{1.1}$$

where C and K are nonempty closed convex subsets of a Hilbert space X and a Banach space Y, respectively, $x \in X$, and G is a differentiable mapping from X to Y. This problem has its origin relating more concrete problems such as the shape-preserving interpolation problem that arises in curve and surface fitting (see [1]). In view of the fact that in many practical situations of (1.1) C has a very simple structure (such as a box or a nonnegative cone in a finite-dimensional space), many researchers (seeing [2–8]) were led to reformulate (1.1) as a system of nonlinear equations by using dual formulation of the "difficult" constraint $G(x) \in K$ and thus to address the issue when (1.1) has the following perturbation property

$$x_0 \in \mathrm{P}_S(x) \Longleftrightarrow x_0 = \mathrm{P}_C(x - G'(x_0)^*\lambda) \quad \text{for some } \lambda \in \mathcal{N}_K(G(x_0)), \tag{1.2}$$

where P_C denotes the projection onto C, $\mathcal{N}_K(G(x_0))$ is the normal cone of K at $G(x_0)$, $G'(x_0)$ is the derivative of G at x_0, and $G'(x_0)^*$ is the adjoint operator of $G'(x_0)$, and S denotes the feasible set of (1.1) defined by

$$S := C \cap \{x \in X : G(x) \in K\}. \tag{1.3}$$

[1]The authors thank the referees for valuable suggestions

[2]Associate Professor, Department of Mathematics, Sichuan Normal University, Chengdu, Sichuan, P. R. China. The author was partially supported by a grant from National Natural Science Foundation of China, and grants (2001)01GY051-66 and SZD0406 from Sichuan Province.

[3]Professor, Department of Mathematics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong. The author was supported by a direct grant (CUHK) and Earmarked Grant from the Research Grant Council of Hong Kong.

If $Y = \mathbb{R}^m$ and $G(x) = Ax - d$ but $K = \mathbb{R}^m_-$, Deutsch, Li and Ward [8] showed that the perturbation property (1.2) is equivalent to the strong CHIP (strong conical hull intersection property) of $\{C, \{x : Ax \leq d\}\}$. Moreover, it is shown in [9, Chapter 3] that if the graph of the multifunction $G(x) - K$ is convex then then the Robinson Constraint Qualification (RCQ) of the system (1.3) at x_0 implies that the perturbation property (1.2) holds.

In this paper, we make further theoretical development on the equivalent characterizations of perturbation property for the general nonlinear constraint system (1.3). For this purpose, we introduce a concept of normal point and prove that if $x_0 \in P_S(x)$ is normal point then the perturbation property (1.2) holds if and only if a basic constraint qualification is satisfied at x_0 (Theorem 3.2). It is also proved if S is the constraint system studied in [2–8] then all points in S are normal points and the basic constraint qualification is just the strong CHIP; see Theorem 4.4. Moreover, we also prove that if the contingent cone of S at a point $x_0 \in P_S(x)$ coincides with the closure of the radial cone of S at x_0 (which holds particularly when the graph of $G(x) - K$ is convex) and if the system (1.3) satisfies the Robinson constraint qualification (RCQ) at x_0 then x_0 is a normal point and the perturbation property (1.2) holds (Theorem 3.3). Thus our results not only recapture many of the earlier results but also demonstrate a natural nonconvex setting for the theory on perturbation property.

This paper is organized as follows. Some notations and preliminary results are presented in the next section. Several equivalent characteristics of perturbation property for the general nonlinear constraint system (1.3) are established in section 3, which are used to compare several earlier results on perturbation property for finite system of inequalities/equalities in section 4.

2. Notations, preliminaries and (RCQ)

Recall that S is defined by (1.3) with C, X, G etc as in the beginning of section 1. For any $x_0 \in S$, we use $\mathbf{T}^G(x_0)$ to denote the closed convex cone in X defined by $\mathbf{T}^G(x_0) := G'(x_0)^{-1}\mathcal{T}_K(G(x_0))$; that is

$$\mathbf{T}^G(x_0) = \{h \in X : G'(x_0)h \in \mathcal{T}_K(G(x_0))\}, \tag{2.1}$$

where $\mathcal{T}_K(G(x_0))$ denotes the tangent cone of K at $G(x_0)$. We show in Section 3 that $\mathbf{T}^G(x_0) \cap \mathcal{T}_C(x_0)$ coincides with $\mathcal{T}_S(x_0)$ if (RCQ) is satisfied. In view of a standard result in duality [10, P.64.], the cone $\mathbf{T}^G(x_0)$ defined in (2.1) equals the polar of $G'(x_0)^*\mathcal{N}_K(G(x_0))$ and consequently, by the Bipolar Theorem [10, Theorem 2.4.3],

$$\mathbf{T}^G(x_0)^{\ominus} = \overline{G'(x_0)^*\mathcal{N}_K(G(x_0))}, \tag{2.2}$$

(note that the norm-closure in (2.2) is the same as the weak*-closure by convexity and the fact that the sets are in the Hilbert space X).

Definition 2.1. $x_0 \in S$ is called a normal point of S (with respect to the constraint system of (1.3)) if

$$S \subseteq C \cap (x_0 + \mathbf{T}^G(x_0)) \subseteq C \cap (x_0 + \overline{\text{co}}\mathcal{T}_S(x_0)). \tag{2.3}$$

Given a Banach space Z, let Z^* denote its (topological) dual space and we write $\langle z^*, z\rangle$ for $z^*(z)$ where $z \in Z$ and $z^* \in Z^*$. For a linear operator $A : X \to Z$, A^* denotes the adjoint operator of A:

$$\langle A^*z^*, x\rangle = \langle z^*, Ax\rangle \text{ for each } x \in X \text{ and } z^* \in Z^*.$$

Since X is a Hilbert space, X and X^* are identified and so $\langle \cdot, \cdot \rangle$ is simply the inner product. For a set W in Z, the interior (respectively, closure, convex hull, closed convex hull, conic hull, negative polar) of W is denoted by $\operatorname{int} W$ (respectively, $\overline{W}$, $\operatorname{co} W$, $\overline{\operatorname{co}} W$, $\operatorname{cone} W$, $W^{\ominus}$). The normal cone of W at $w_0 \in W$ is denoted by $\mathcal{N}_W(w_0)$ and defined by $\mathcal{N}_W(w_0) = (W - w_0)^{\ominus}$:

$$\mathcal{N}_W(w_0) := \{z^* \in Z^* : \langle z^*, w - w_0 \rangle \leq 0 \text{ for each } w \in W\}.$$

For a set $K \subseteq Z^*$, $K^{\ominus} := \{z \in Z : \langle z^*, z \rangle \leq 0 \text{ for all } z^* \in K\}$ is the negative polar of K. The (Bouligand-) contingent cone (cf. [9–11]) of W at w_0 is denoted by $\mathcal{T}_W(w_0)$, and the radial cone of W at w_0 is denoted by $\mathcal{R}_K(x_0)$ and is defined by $\mathcal{R}_W(w_0) := \bigcup_{t>0} t^{-1}(W - w_0)$. It is known and easy to prove that

$$\mathcal{T}_W(w_0) \subseteq \overline{\mathcal{R}_W(w_0)}. \tag{2.4}$$

If W is convex then

$$\mathcal{T}_W(w_0) = \overline{\mathcal{R}_W(w_0)}, \tag{2.5}$$

and the cones $\mathcal{T}_W(w_0)$, $\mathcal{N}_W(w_0)$ are polars of each other. For $z \in Z$, $\mathrm{P}_W(z)$ denotes the projection of z to W:

$$\mathrm{P}_W(z) = \{w \in W : \|w - z\| = \operatorname{dist}(z, W)\},$$

where $\operatorname{dist}(z, W) := \inf_{y \in W} \|z - y\|$. In the case when Z is a Hilbert space, $\mathrm{P}_W(z)$ contains at most one point for each $z \in Z$. The following result, valid in Hilbert spaces, is well-known (see [12, Page 23 and Proposition 1.10]) and will be useful for our discussions.

Proposition 2.1. Let W be a closed convex set in a Hilbert space H, and let $w_0 \in W$ and $d \in H$. Then $d \in \mathcal{N}_W(w_0)$ if and only if $\mathrm{P}_W(w_0 + td) = \|td\|$ for some $t > 0$.

For the following proposition, we consider the case when W is not necessarily convex, thus $\mathcal{T}_W(w_0)$ is not necessarily convex.

Proposition 2.2. Let Z be a Banach space, and let $w_0 \in W \subseteq Z$. Then

$$\overline{\operatorname{co}} \mathcal{T}_W(w_0) = \overline{\operatorname{co}} \mathcal{R}_W(w_0) \tag{2.6}$$

if and only if

$$W \subseteq w_0 + \overline{\operatorname{co}} \mathcal{T}_W(w_0). \tag{2.7}$$

Proof. We need only to prove that (2.7) implies (2.6). In view of (2.4), to establish (2.6) it suffices to show the inclusion $\mathcal{R}_W(w_0) \subseteq \overline{\operatorname{co}} \mathcal{T}_W(w_0)$. But this inclusion follows trivially from (2.7) as $\overline{\operatorname{co}} \mathcal{T}_W(w_0)$ is a cone. □

For the study of nonconvex sets, the following definition will be useful for us.

Definition 2.2. Let Z be a Banach space and let C be a convex set in Z. Let $w_0 \in W \subseteq Z$, and let T be a convex set in Z containing the origin. W is said to be convexified at w_0 by T with respect to C if it holds that

$$W \subseteq C \cap (w_0 + T) \subseteq C \cap (w_0 + \overline{\operatorname{co}} \mathcal{T}_W(w_0)). \tag{2.8}$$

Thus in the situation of Definition 2.1, $x_0 \in S$ is a normal point of S if and only if S is convexified at x_0 by $\mathbf{T}^G(x_0)$ with respect to C. Needless to say, if W is assumed to be convex, then W is convexified at any $x_0 \in W$ by $W - x_0$.

In the following we return to the constraint system (1.3): the notations are as at the beginning of Section 1. In particular, C and K are assumed to be closed and convex; S denotes the set defined in (1.3), that is $S = C \cap G^{-1}(K)$.

Definition 2.3 ([9, 13]). Let $x_0 \in S$. The constraint system (1.3) is said to satisfy (RCQ) at x_0 if

$$G'(x_0)\mathcal{R}_C(x_0) - \mathcal{R}_K(G(x_0)) = Y. \tag{2.9}$$

Recalling the definition of $\mathbf{T}^G(x_0)$ given in (2.1), we note that

$$\mathcal{T}_S(x_0) \subseteq \mathbf{T}^G(x_0) \cap \mathcal{T}_C(x_0) \quad \text{for each } x_0 \in S. \tag{2.10}$$

Indeed, if d is a unit vector in $\mathcal{T}_S(x_0)$, then

$$\lim_n \frac{s_n - x_0}{\|s_n - x_0\|} = d$$

for some sequence $(s_n) \subseteq S$ converging to x_0. Consequently $d \in \mathcal{T}_C(x_0)$ and since $G(s_n) = G(x_0) + G'(x_0)(s_n - x_0) + o(\|s_n - x_0\|)$,

$$G'(x_0)d = \lim_n \frac{G(s_n) - G(x_0)}{\|s_n - x_0\|} \in \mathcal{T}_K(G(x_0))$$

because $G(s_n) \in K$ for each n. Therefore (2.10) is true. Part (a) of the following proposition shows that (RCQ) implies that the equality in (2.10) holds. This result is proved in [9, Corollary 2.91] for the case when $C = X$.

Proposition 2.3. Let $x_0 \in S$ and suppose that the constraint system (1.3) satisfies (RCQ) at x_0. Then the following assertions hold:

(a) $\mathcal{T}_S(x_0) = \mathbf{T}^G(x_0) \cap \mathcal{T}_C(x_0)$; in particular, $\mathcal{T}_S(x_0)$ is closed and convex.

(b) $S \subseteq x_0 + \mathcal{T}_S(x_0) \Longleftrightarrow \mathcal{T}_S(x_0) = \overline{\mathcal{R}_S(x_0)}$.

(c) $C \cap (x_0 + \mathbf{T}^G(x_0)) = C \cap (x_0 + \mathcal{T}_S(x_0))$.

(d)
$$\mathcal{N}_{C\cap(x_0+\mathbf{T}^G(x_0))}(x_0) = \mathcal{N}_C(x_0) + G'(x_0)^*\mathcal{N}_K(G(x_0)). \tag{2.11}$$

Proof. Denote Z for the product $X \times Y$, and $\mathcal{K}$ for $C \times K$. Let $\mathcal{G}$ denote the function $x \mapsto (x, G(x))$ from X into Z. Then $\mathcal{G}'(x_0)h = (h, G'(x_0)h)$ for each h and

$$\mathcal{T}_{\mathcal{K}}(x_0, G(x_0)) = \mathcal{T}_C(x_0) \times \mathcal{T}_K(G(x_0));$$

consequently, it follows from (2.1) that

$$\{h : \mathcal{G}'(x_0)h \in \mathcal{T}_{\mathcal{K}}(x_0, G(x_0))\} = \mathbf{T}^G(x_0)) \cap \mathcal{T}_C(x_0). \tag{2.12}$$

On the other hand, by (2.9), it is known (cf. [9, Page 71]) and easy to verify that

$$\mathcal{G}'(x_0)X - \mathcal{R}_{\mathcal{K}}(\mathcal{G}(x_0)) = Z.$$

Hence by [9, Corollary 2.91], the set on the left-hand side of (2.12) equals the tangent cone of $\mathcal{G}^{-1}(\mathcal{K})$ at x_0 which is exactly $\mathcal{T}_S(x_0)$ because clearly $S = \mathcal{G}^{-1}(\mathcal{K})$. Thus (a) is proved. (b) follows easily from (a) and (2.4). In view of (a), to prove (c), it suffices to show that the set on the left-hand side of (c) is contained in the set on the right-hand side. Accordingly, let $\xi \in C \cap (x_0 + \mathbf{T}^G(x_0))$. Then $\xi - x_0 \in \mathbf{T}^G(x_0)$ and $\xi \in C$. Since C is convex, it follows that $\xi - x_0 \in \mathcal{T}_C(x_0)$. Hence, by (a), $\xi - x_0 \in \mathcal{T}_S(x_0)$ and so ξ belongs to the set on the right-hand side of (c).

It remains to prove that the assertion (d). Define

$$\begin{aligned} F(x) &:= G'(x_0)(x - x_0) \quad \text{for each } x \in X, \\ D &:= \mathcal{T}_K(G(x_0)), \end{aligned}$$

and let S' denote the set of all x satisfying the new system

$$x \in C \text{ and } F(x) \in D. \tag{2.13}$$

Note that $F'(x_0) = G'(x_0)$, $F(x_0) = 0$ and $\mathcal{T}_D(F(x_0)) = \mathcal{T}_K(G(x_0))$. Hence $\mathbf{T}^F(x_0) = \mathbf{T}^G(x_0)$ and $S' = C \cap (x_0 + \mathbf{T}^G(x_0))$. Since the constraint system (1.3) satisfies (RCQ) at x_0, so does the new system (2.13). It follows from part (a) that

$$\mathcal{T}_{S'}(x_0) = \mathcal{T}_C(x_0) \cap \mathbf{T}^F(x_0) = \mathcal{T}_C(x_0) \cap \mathbf{T}^G(x_0) \tag{2.14}$$

Noting the convexity of S', making use of (2.2) and taking polars on the both sides of (2.14), one has

$$\mathcal{N}_{S'}(x_0) = \overline{\mathcal{N}_C(x_0) + \overline{G'(x_0)^* \mathcal{N}_K(G(x_0))}}$$

Thus, to prove (2.11), it remains to show that

$$\mathcal{N}_{S'}(x_0) \subseteq \mathcal{N}_C(x_0) + G'(x_0)^* \mathcal{N}_K(G(x_0)) \tag{2.15}$$

To do this, let $\xi \in \mathcal{N}_{S'}(x_0)$. Consider the following minimization problem

$$\begin{array}{rl} \min & \langle -\xi, h \rangle \\ \text{subject to} & h \in \mathcal{T}_C(x_0) \\ & G'(x_0)h \in \mathcal{T}_K(G(x_0)), \end{array} \tag{2.16}$$

and its dual problem [9, Section 2.5.6]

$$\begin{array}{rl} \max & \langle \lambda, 0 \rangle \\ \text{subject to} & \lambda \in \mathcal{N}_K(G(x_0)) \\ & \xi - G'(x_0)^* \lambda \in \mathcal{N}_C(x_0). \end{array} \tag{2.17}$$

Since $\mathcal{N}_{S'}(x_0)^{\ominus} = \mathcal{T}_{S'}(x_0)$, $h_0 := 0$ solves the problem (2.16). By the (RCQ) assumption, [9, Theorem 2.187] implies that there is no duality gap between the primal problem (2.16) and the dual problem (2.17). This shows that the feasible set of the dual problem (2.17) is nonempty. That is, there exists $\lambda \in \mathcal{N}_K(G(x_0))$ such that $\xi - G'(x_0)^* \lambda$ belongs to $\mathcal{N}_C(x_0)$. Thus

$$\xi \in \mathcal{N}_C(x_0) + G'(x_0)^* \mathcal{N}_K(G(x_0)).$$

This proves (2.15). □

Let $\{C_1, C_2, \ldots, C_m\}$ be a collection of closed convex sets and let $x_0 \in \cap_{i=1}^m C_i$. Following [4, 8], the collection is said to have the strong CHIP (strong conical hull intersection property) at x_0 if

$$\mathcal{N}_{\cap_{i=1}^m C_i}(x_0) = \sum_{i=1}^m \mathcal{N}_{C_i}(x_0).$$

3. Perturbation property of best approximation

Consider the following minimization problem:

$$(\mathcal{P}_{f,G}) \quad \begin{cases} \min & f(x) \\ \text{subject to} & x \in C, \text{ and } G(x) \in K, \end{cases} \tag{3.1}$$

where $f : X \to \mathbb{R}$ is differentiable and other notations are as in (1.3). Recall that the set $S = C \cap G^{-1}(K)$ is not necessarily convex as G is not assumed to be convex with respect to K. Let $x_0 \in S$ and let T_0 be a convex set in X containing the origin. Recall from Definition 2.2 that S is convexified at x_0 by T_0 with respect to C if

$$S \subseteq C \cap (x_0 + T_0) \subseteq C \cap (x_0 + \overline{\text{co}}\mathcal{T}_S(x_0)). \tag{3.2}$$

Proposition 3.1. Assume (3.2) and denote $C \cap (x_0 + T_0)$ by S_0. Then for any convex (and differentiable) f, the following statements are equivalent:

(i) x_0 is an optimal solution to $(\mathcal{P}_{f,G})$.

(ii) $-f'(x_0) \in \mathcal{T}_S(x_0)^{\ominus}$.

(iii) x_0 minimizes f over $C \cap (x_0 + \overline{\text{co}}(\mathcal{T}_S(x_0)))$.

(iv) x_0 minimizes f over S_0.

(v) $-f'(x_0) \in \mathcal{N}_{S_0}(x_0)$.

Proof. (i)⇒(ii). A proof can be found in [9, Lemma 3.7 (i)] even without assuming the convexity of f.

(ii)⇒(iii). Since $-f'(x_0) \in \mathcal{T}_S(x_0)^{\ominus}$, one has

$$\langle f'(x_0), y - x_0 \rangle \geq 0 \text{ for each } y \in x_0 + \overline{\text{co}}(\mathcal{T}_S(x_0)).$$

Since f is convex, the subgradient inequality gives

$$f(y) - f(x_0) \geq \langle f'(x_0), y - x_0 \rangle \geq 0 \text{ for each } y \in x_0 + \overline{\text{co}}(\mathcal{T}_S(x_0)).$$

Thus (iii) holds.

(iii)⇒ (iv)⇒ (i). This is evident as $S \subseteq S_0 \subseteq C \cap (x_0 + \overline{\text{co}}(\mathcal{T}_S(x_0)))$ and $x_0 \in S$.

(iv)⇔ (v). This is obvious due to the convexity of f and S_0. □

Below we specialize Proposition 3.1 by considering the functions of the form f_x defined by

$$f_x(w) = \frac{1}{2}\|w - x\|^2 \quad \text{for each } w \in X, \tag{3.3}$$

where $x \in X$.

Theorem 3.1. Assume (3.2) and denote $C \cap (x_0 + T_0)$ by S_0. Let N' be any nonempty subset of $T_0^{\ominus}$. Then the following statements are equivalent:

(i) $\mathcal{N}_{S_0}(x_0) = \mathcal{N}_C(x_0) + N'$.

(ii) For any $x \in X$,

$$x_0 \in \mathrm{P}_S(x) \Longleftrightarrow x_0 = \mathrm{P}_C(x - x^*) \text{ for some } x^* \in N'. \tag{3.4}$$

(iii) For any differentiable convex function $f : X \to \mathbb{R}$, x_0 minimizes f over S if and only if there exists $x^* \in N'$ such that $-f'(x_0) - x^* \in N_C(x_0)$.

Proof. (i)⇒(ii). Let $x \in X$ and consider the problem $(\mathcal{P}_{f_x,G})$. Then the first equivalence in the following is clear by Proposition 3.1:

$$\begin{aligned} x_0 \in \mathrm{P}_S(x) \iff x_0 = \mathrm{P}_{S_0}(x) \iff x - x_0 \in \mathcal{N}_{S_0}(x_0) \\ \iff x - x_0 \in \mathcal{N}_C(x_0) + N' \\ \iff x - x^* - x_0 \in \mathcal{N}_C(x_0) \text{ for some } x^* \in N' \\ \iff x_0 = \mathrm{P}_C(x - x^*) \text{ for some } x^* \in N', \end{aligned}$$

where the second and the last equivalences are due to Proposition 2.1 (and the convexity of S_0 and C).

(ii)⇒(i). Since $S_0 - x_0 \subseteq T_0$, one has $T_0^{\ominus} \subseteq \mathcal{N}_{S_0}(x_0)$. Similarly $\mathcal{N}_C(x_0) \subseteq \mathcal{N}_{S_0}(x_0)$. Thus, to establish (i), it suffices to show that $\mathcal{N}_{S_0}(x_0) \subseteq \mathcal{N}_C(x_0) + N'$. To do this, let $\xi \in \mathcal{N}_{S_0}(x_0)$. Then, as above, $x_0 \in \mathrm{P}_S(\xi + x_0)$. By (3.4) it follows that $x_0 = \mathrm{P}_C(\xi + x_0 - x^*)$ for some $x^* \in N'$. Then we have $\xi - x^* \in \mathcal{N}_C(x_0)$ and it follows that $\xi \in \mathcal{N}_C(x_0) + N'$.

(i)⇒(iii). By (i), it follows from Proposition 3.1 that

$$\begin{aligned} &- f'(x_0) - x^* \in \mathcal{N}_C(x_0) \text{ for some } x^* \in N' \iff -f'(x_0) \in \mathcal{N}_{S_0}(x_0) \\ &\iff x_0 \text{ minimizes } f \text{ over } S. \end{aligned}$$

This shows that (iii) holds.

Finally applying (iii) to the function f_x, it is clear that (iii) implies (ii). □

Recall that $\mathbf{T}^G(x_0)$ is defined in (2.1). Recall also that $x_0 \in S$ is a normal point of S with respect to the constraint system (1.3) if and only if (3.2) holds with T_0 replaced by $\mathbf{T}^G(x_0)$. By virtue of (2.2) and letting $N' = G'(x_0)^*\mathcal{N}_K(G(x_0))$, the following result follows immediately from Theorem 3.1.

Theorem 3.2. Let x_0 be a normal point of S with respect to the constraint system (1.3). Then the following statements are equivalent:

(i) $\mathcal{N}_{C\cap(x_0+\mathbf{T}^G(x_0))}(x_0) = \mathcal{N}_C(x_0) + G'(x_0)^*\mathcal{N}_K(G(x_0))$.

(ii) For any $x \in X$, $x_0 \in \mathrm{P}_S(x)$ if and only if there exists $\lambda \in \mathcal{N}_K(G(x_0))$ such that

$$x_0 = \mathrm{P}_C(x - G'(x_0)^*\lambda).$$

(iii) For any differentiable convex function $f : X \to \mathbb{R}$, x_0 minimizes f over S if and only if there exists $\lambda \in \mathcal{N}_K(G(x_0))$ such that λ is a Lagrangian multiplier of the problem $(\mathcal{P}_{f,G})$.

Similarly we have two more consequences of Theorem 3.1. For brevity sake, we omit the statement regarding Lagrange rule.

Corollary 3.1. Let $x_0 \in S$ and assume that S is convexified by $G'(x_0)^{-1}(K - G(x_0))$. Let N' be any subset of $\overline{G'(x_0)^*\mathcal{N}_K(G(x_0))}$. Then the following statements are equivalent:

(i) $\mathcal{N}_{C\cap(x_0+G'(x_0)^{-1}(K-G(x_0)))}(x_0) = \mathcal{N}_C(x_0) + N'$.

(ii) For any $x \in X$, $x_0 \in \mathrm{P}_S(x)$ if and only if there exists $x^* \in N'$ such that $x_0 = \mathrm{P}_C(x - x^*)$.

Proof. Denote $G'(x_0)^{-1}(K - G(x_0))$ by T_0. Then (3.2) is satisfied. Since $K - G(x_0) \subseteq \mathcal{T}_K(G(x_0))$ by convexity of K, we also have $T_0 \subseteq \mathbf{T}^G(x_0)$; it follows from (2.2) that $G'(x_0)^* \mathcal{N}_K(G(x_0)) \subseteq T_0^{\ominus}$. Hence N' is also a subset of $T_0^{\ominus}$. Thus one can apply Theorem 3.1 to conclude the proof. □

Corollary 3.2. Let $x_0 \in S$. Assume that S is convexified by $\overline{G'(x_0)^{-1}\mathcal{R}_K(G(x_0))}$ and that

$$C \cap \left(x_0 + \overline{G'(x_0)^{-1}\mathcal{R}_K(G(x_0))}\right) = \overline{C \cap (x_0 + G'(x_0)^{-1}\mathcal{R}_K(G(x_0)))}. \tag{3.5}$$

Let N' be any subset of $\overline{G'(x_0)^* \mathcal{N}_K(G(x_0))}$. Then the following statements are equivalent:

(i) $\mathcal{N}_{C \cap (x_0 + G'(x_0)^{-1}(K - G(x_0)))}(x_0) = \mathcal{N}_C(x_0) + N'$.

(ii) For any $x \in X$, $x_0 \in \mathrm{P}_S(x)$ if and only if there exists $x^* \in N'$ such that $x_0 = \mathrm{P}_C(x - x^*)$.

Proof. Denote $\overline{G'(x_0)^{-1}\mathcal{R}_K(G(x_0))}$ by T_0. Then (3.2) is satisfied and $T_0 \subseteq \mathbf{T}^G(x_0)$. Hence, as before, (ii) is equivalent to

$$\mathcal{N}_{C \cap (x_0 + T_0)}(x_0) = \mathcal{N}_C(x_0) + N' \tag{i*}$$

or what is the same (thanks to (3.5)):

$$\mathcal{N}_{\overline{C \cap (x_0 + G'(x_0)^{-1}\mathcal{R}_K(G(x_0)))}}(x_0) = \mathcal{N}_C(x_0) + N'. \tag{i\#}$$

Thus to complete the proof, we need only to prove the sets on the left-hand side of (i) and of (i$^\#$) are the same. It suffices to prove that

$$\mathcal{N}_{C \cap (x_0 + T_0)}(x_0) \subseteq \mathcal{N}_{C \cap (x_0 + G'(x_0)^{-1}(K - G(x_0)))}(x_0), \tag{3.6}$$

since the other inclusion relations clearly holds. To do this, let $\xi \in \mathcal{N}_{C \cap (x_0 + T_0)}(x_0)$ and $z \in C \cap \left(x_0 + G'(x_0)^{-1}\mathcal{R}_K(G(x_0))\right)$. Then there exists $t \in (0, 1)$ such that $G'(x_0)(z - x_0) \in t^{-1}(K - G(x_0))$. This implies that $t(z - x_0) \in T_0$, and hence that $x_0 + t(z - x_0) \in C \cap (x_0 + T_0)$. Consequently,

$$\langle \xi, t(z - x_0) \rangle \leq 0,$$

and so (3.6) is proved. □

For the following theorem, we do not assume (3.2) at the outset.

Theorem 3.3. Let $x_0 \in S$. Suppose that the constraint system (1.3) satisfies (RCQ) at x_0. Then each of the following statements implies the other statements and that x_0 is a normal point of S.

(i) $\mathcal{T}_S(x_0) = \overline{\mathcal{R}_S(x_0)}$.

(ii) For any $x \in X$, $x_0 \in \mathrm{P}_S(x)$ if and only if there exists $\lambda \in \mathcal{N}_K(G(x_0))$ such that

$$x_0 = \mathrm{P}_C(x - G'(x_0)^* \lambda).$$

(iii) For any differentiable convex function $f : X \to \mathbb{R}$, x_0 minimizes f over S if and only if there exists $\lambda \in \mathcal{N}_K(G(x_0))$ such that λ is a Lagrange multiplier of the problem $(\mathcal{P}_{f,G})$.

Consequently, if in addition S is assumed to be convex, then (i) and hence (ii), (iii) all hold.

Proof. Assume (i). Together with (RCQ), it follows from Proposition 2.3 that x_0 must be a normal point of S. Furthermore, by virtue of assumptions, one can apply Proposition 2.3(d) to conclude that (i) in Theorem 3.2 holds. Hence by that theorem, we have (i) $\Rightarrow$ (ii) and (iii).

(ii)$\Rightarrow$(i). In view of Proposition 2.3, to establish (i), it suffices to prove that $S \subseteq x_0 + \mathcal{T}_S(x_0)$. By [14, Lemma 3.1], the inclusion will follow provided that the following implication holds: for any $x \in X$,

$$x_0 = \mathrm{P}_{C\cap(x_0+\mathcal{T}_S(x_0))}(x) \Longrightarrow x_0 \in \mathrm{P}_S(x). \tag{3.7}$$

By Proposition 2.3, $\mathcal{T}_S(x_0)$ is convex and

$$\begin{aligned}\mathcal{N}_{C\cap(x_0+\mathcal{T}_S(x_0))}(x_0) &= \mathcal{N}_{C\cap(x_0+\mathbf{T}^G(x_0))}(x_0)\\ &= \mathcal{N}_C(x_0) + G'(x_0)^*\mathcal{N}_K(G(x_0)).\end{aligned} \tag{3.8}$$

To prove (3.7), let $x_0 = \mathrm{P}_{C\cap(x_0+\mathcal{T}_S(x_0))}(x)$. Then, by Proposition 2.1, $x - x_0 \in \mathcal{N}_{C\cap(x_0+\mathcal{T}_S(x_0))}(x_0)$ and it follows from (3.8) that $x - x_0 - G'(x_0)^*\lambda \in \mathcal{N}_C(x_0)$ for some $\lambda \in \mathcal{N}_K(G(x_0))$. By Proposition 2.1 and (ii), we then conclude that $x_0 \in \mathrm{P}_S(x)$.

The implication (iii) $\Rightarrow$ (ii) is clear by considering the function f_x defined in (3.3). $\square$

4. Applications to finite system of nonconvex inequalities

In this section, we shall apply the results in Section 3 to study the following system

$$x \in C, \text{ and } (h_i \circ F_i)(x) \leq 0 \quad \text{for each } i = 1, 2, \ldots, m, \tag{4.1}$$

where C is a closed convex set in a Hilbert space X, each $h_i : Z_i \to \mathbb{R}$ is a continuous convex function on a Banach space Z_i and each $F_i : X \to Z_i$ is a differentiable mapping.

Remark 4.1. The system has been studied by many researchers for some special cases:

(a) [15] studied the case when $X = Z_i$ for each i and F_i is the identity mapping, and also the case when each $Z_i = \mathbb{R}$ and h_i is the identity mapping.

(b) The special case when for each i, $X = Z_i$, F_i is the identity mapping and each h_i is affine has been extensively studied (see [1–7]). In particular, for such a system, Deutsch, Li and Ward established in [4] the equivalence of (BCQ) and the perturbation property.

(c) [14] studied the case when the spaces Z_i, $i = 1, 2, \ldots, m$ are the same. The perturbation property for the system (4.1) established in [14] generalizes those established in [8, 15].

Note that the system (4.1) can be converted to the system of the form

$$x \in C, \quad G(x) \in K \tag{4.2}$$

(see (1.3)). Indeed, let Y be the product space $\prod_{i=1}^m Z_i$, and define $G : X \to Y$ and $K \subseteq Y$ by

$$\begin{array}{ll} G(x) := (F_1(x), \ldots, F_m(x)), & \text{for each } x \in X \\ K := \prod_{i=1}^m K_i, & \text{where } K_i := \{y \in Z_i : h_i(y) \leq 0\}. \end{array} \tag{4.3}$$

Clearly K is closed and the feasible set S of (4.1) is the same as the set $C \cap G^{-1}(K)$. For $x_0 \in S$, let $I(x_0) = \{i : (h_i \circ F_i)(x_0) = 0\}$. Recall that $\mathbf{T}^G(x_0)$ stands for the set $G'(x_0)^{-1}\mathcal{T}_K(G(x_0))$. Then, by (4.3),

$$\mathbf{T}^G(x_0) = \bigcap_{i=1}^m F_i'(x_0)^{-1}\mathcal{T}_{K_i}(F_i(x_0)) = \bigcap_{i \in I(x_0)} F_i'(x_0)^{-1}\mathcal{T}_{K_i}(F_i(x_0)). \tag{4.4}$$

Indeed if $i \notin I(x_0)$ then $h_i(F_i(x_0)) < 0$ and so $F_i(x_0)$ is an interior point of K_i. This implies that $\mathcal{T}_{K_i}(F_i(x_0)) = Z_i$ for every $i \notin I(x_0)$ and (4.4) is clear. Consequently

$$\mathbf{T}^G(x_0)^{\ominus} = \overline{\sum_{i \in I(x_0)} \{F_i'(x_0)^{-1}\mathcal{T}_{K_i}(F_i(x_0))\}^{\ominus}} = \overline{\sum_{i \in I(x_0)} F_i'(x_0)^*\mathcal{N}_{K_i}(F_i(x_0))}. \tag{4.5}$$

We say that $x_0 \in S$ is a normal point of S with respect to (4.1) if it is a normal point with respect to (4.2). Thus we may henceforth refer this case by saying that x_0 is a normal point of S.

By standard results in convex analysis (cf. [11, Corollary 1 of Theorem 2.4.7 and Proposition 2.3.9]), one has, for each $i \in I(x_0)$,

$$\partial h_i(F_i(x_0)) \subseteq \mathcal{N}_{K_i}(F_i(x_0)) \tag{4.6}$$

and

$$\partial(h_i \circ F_i)(x_0) = F_i'(x_0)^*\partial h_i(F_i(x_0)), \tag{4.7}$$

where $\partial(h_i \circ F_i)$ is the Clarke subdifferential. Consequently it follows from (4.4) that

$$\operatorname{cone}\{\partial(h_i \circ F_i)(x_0) : i \in I(x_0)\} \subseteq \sum_{i \in I(x_0)} F_i'(x_0)^*\mathcal{N}_{K_i}(F_i(x_0)) \subseteq \mathbf{T}^G(x_0)^{\ominus}. \tag{4.8}$$

Thus, applying Theorem 3.1 to $\mathbf{T}^G(x_0)$ for T_0 with the first two sets in (4.8) respectively for N', we immediately have the following two theorems.

Theorem 4.1. Let S be the feasible set of the system (4.1), and let $x_0 \in S$ be a normal point. Then the following statements are equivalent:

(i) $\mathcal{N}_{C \cap (x_0 + \mathbf{T}^G(x_0))}(x_0) = \mathcal{N}_C(x_0) + \operatorname{cone}\{\partial(h_i \circ F_i)(x_0) : i \in I(x_0)\}$.

(ii) For any $x \in X$, $x_0 \in \mathrm{P}_S(x)$ if and only if there exist $\xi_i \in \partial(h_i \circ F_i)(x_0)$ and $\lambda_i \geq 0$, with $\lambda_i = 0$ for all $i \notin I(x_0)$ such that

$$x_0 = \mathrm{P}_C\left(x - \sum_{i \in I(x_0)} \lambda_i \xi_i\right).$$

Theorem 4.2. Let S be the feasible set of the system (4.1), and let $x_0 \in S$ be a normal point. Then the following statements are equivalent:

(i) $\mathcal{N}_{C\cap(x_0+\mathbf{T}^G(x_0))}(x_0) = \mathcal{N}_C(x_0) + \sum_{i\in I(x_0)} F_i'(x_0)^* \mathcal{N}_{K_i}(F_i(x_0))$.

(ii) For any $x \in X$, $x_0 \in \mathrm{P}_S(x)$ if and only if for every $i \in I(x_0)$, there exist $\xi_i \in \mathcal{N}_{K_i}(F_i(x_0))$ such that

$$x_0 = \mathrm{P}_C\left(x - \sum_{i\in I(x_0)} F_i'(x_0)^*\xi_i\right).$$

Following the notations in [14], let

$$\begin{aligned}\mathrm{CFD}(x_0) : &= \{d \in X : F_i(x_0) + F_i'(x_0)d \in K_i \text{ for each } i \in I(x_0)\}\\ &= \bigcap_{i\in I(x_0)} F_i'(x_0)^{-1}(K_i - F_i(x_0)).\end{aligned}$$

Note that $\mathrm{CFD}(x_0) \subseteq \bigcap_{i\in I(x_0)} F_i'(x_0)^{-1}\mathcal{T}_{K_i}(F_i(x_0))$. Taking polars of the sets on both sides and applying the Bipolar Theorem [10, Theorem 2.4.3], we have

$$\sum_{i\in I(x_0)} F_i'(x_0)^*\mathcal{N}_{K_i}(F_i(x_0)) \subseteq \mathrm{CFD}(x_0)^{\ominus}. \tag{4.9}$$

The following result was proved in [14] for the special case when Z_i, $i = 1, 2, \dots, m$ are the same.

Theorem 4.3. Let S be the feasible set of (4.1) and let $x_0 \in S$. Suppose S is convexified at x_0 by $\mathrm{CFD}(x_0)$ with respect to C. Then the following statements are equivalent:

(i) $\mathcal{N}_{C\cap(x_0+\mathrm{CFD}(x_0))}(x_0) = \mathcal{N}_C(x_0) + \mathrm{cone}\{\partial(h_i \circ F_i)(x_0) : i \in I(x_0)\}$.

(ii) For each $x \in X$, $x_0 \in \mathrm{P}_S(x)$ if and only if there exist $\xi_i \in \partial(h_i \circ F_i)(x_0)$ and $\lambda_i \geq 0$ with $\lambda_i = 0$ for all $i \notin I(x_0)$ such that

$$x_0 = \mathrm{P}_C\left(x - \sum_{i\in I(x_0)} \lambda_i\xi_i\right).$$

Proof. In view of (4.8) and (4.9), one can apply Theorem 3.1 to $\mathrm{CFD}(x_0)$ and $\mathrm{cone}\{\partial(h_i \circ F_i)(x_0) : i \in I(x_0)\}$ for T_0 and N'. □

Below we provide an example for which Theorems 4.1 and 4.2 are applicable but not Theorems 4.3.

Example 4.1. Let $X = Y = \mathbb{R}^2$, C be the ball in $\mathbb{R}^2$ with center $(1, 0)$ and radius 1. Let h, F be defined by

$$h(x_1, x_2) = x_1^2 - x_2 \text{ and } F(x_1, x_2) = (x_1, x_2^2).$$

Take $x_0 = (0, 0)$ and $K := h^{-1}(-\infty, 0]$. Then $\mathcal{T}_K(F(x_0))$ is the upper half-plane and so contains the range of the map $F'(x_0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Hence $\mathbf{T} := F'(x_0)^{-1}\mathcal{T}_K(F(x_0))$ is the whole space X. Note that

$$S := C \cap F^{-1}(K) = \{x \in \mathbb{R}^2 : (x_1 - 1)^2 + x_2^2 \leq 1,\ x_1^2 \leq x_2^2\},$$

and hence that $\overline{\mathrm{co}}\, T_S(x_0)$ is the right half-plane. Thus

$$S \subseteq C \cap (x_0 + \mathbf{T}) = C \cap (x_0 + \overline{\mathrm{co}}\mathcal{T}_S(x_0)).$$

On the other hand, $\mathrm{CFD}(x_0) = F'(x_0)^{-1}(K - F(x_0))$ is clearly the singleton $\{(0,0)\}$. Hence $S \not\subseteq C \cap (x_0 + \mathrm{CFD}(x_0))$. Thus S is not convexified at x_0 by $\mathrm{CFD}(x_0)$ with respect to C.

For the remainder of this paper we further specialize (4.1) to study a finite system of linear inequalities/equalities:

$$\begin{cases} x \in C \\ \langle a_i, x\rangle \le b_i, i = 1, \dots, \ell \\ \langle a_i, x\rangle = b_i, i = \ell + 1, \dots, m \end{cases} \tag{LS}$$

(with the obvious interpretation if $\ell = 0$ or $\ell = m$. Here x and each a_i are vectors in a Hilbert space X, each b_i is a real number and $C \subseteq X$ is a closed convex set. Let S consist of all x satisfying (LS). We assume S is nonempty and, as before, for each $x_0 \in S$ let

$$I(x_0) = \{i : 1 \le i \le \ell, \langle a_i, x\rangle - b_i = 0\} \cup \{\ell + 1, \dots, m\},$$

let $I'(x_0) := \{i : 1 \le i \le \ell, i \notin I(x_0)\}$. Denote $Y = \mathbb{R}^m$ and $G(x) = Ax - b$ where A is the linear operator from X into Y defined by: $Ax = (\langle a_i, x\rangle)_{i=1,\dots,m}$. Let $K \subseteq \mathbb{R}^m$ be defined by $K := \mathbb{R}^\ell_- \times \{0\}^{m-\ell}$. Then $\mathcal{T}_K(G(x_0)) = \mathbb{R}_-^{I(x_0)} \times \mathbb{R}^{I'(x_0)}$; more precisely $u = (u_1, \dots, u_m)$ belongs to $\mathcal{T}_K(G(x_0))$ if and only if

$$u_i \le 0 \text{ for each } i \in I(x_0).$$

Therefore $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathcal{N}_K(G(x_0))$ if and only if

$$\lambda_i \ge 0 \text{ for each } i \in I(x_0) \text{ and } \lambda_i = 0 \text{ for each } i \in I'(x_0).$$

Consequently

$$G'(x_0)^*\mathcal{N}_K(G(x_0)) = \left\{ \sum_{i \in I(x_0)} \lambda_i a_i : \lambda_i \ge 0 \text{ for each } i \right\} \tag{4.10}$$

(in particular this set is closed). Recalling the notation $\mathbf{T}^G(x_0) = G'(x_0)^{-1}\mathcal{T}_K(G(x_0))$, one has

$$x_0 + \mathbf{T}^G(x_0) = \{x \in X : \langle a_i, x\rangle \le b_i \text{ for each } i \in I(x_0)\}. \tag{4.11}$$

Thus the first inclusion in the following is clear:

$$S \subseteq C \cap (x_0 + \mathbf{T}^G(x_0)) \subseteq C \cap (x_0 + \mathcal{T}_S(x_0)). \tag{4.12}$$

To see the second inclusion, let $\xi \in \mathbf{T}^G(x_0)$. Then $\langle a_i, \xi\rangle \le 0$ for each $i \in I(x_0)$, and the equality holds for each $i = \ell + 1, \dots\dots m$. Consequently, for all $\varepsilon > 0$ small enough, one has that

$$\langle a_i, x_0 + \varepsilon\xi\rangle \le 0 \text{ for all } i = 1, \dots, \ell,$$

and the equality holds for all $i = \ell + 1, \dots\dots m$. Thus $x_0 + \varepsilon\xi \in S$ for all such ε; hence $\xi \in \mathcal{T}_S(x_0)$. The equivalences of (i)-(iii) in the following theorem are due to Deutsch, Li and Ward [4, 8]. We denote H_i the set $\{x : \langle a_i, x\rangle \le b_i\}$ for $1 \le i \le \ell$ and the set $\{x : \langle a_i, x\rangle = b_i\}$ for $\ell + 1 \le i \le m$.

Theorem 4.4. Let S be the feasible set of (LS). Then each point of S is normal. Moreover the following statements are equivalent:

(i) $\{C, H_i\ (i \in I(x_0))\}$ has the strong CHIP.

(ii) For any $x \in X$, $x_0 \in P_S(x)$ if and only if there exist $\lambda_i \geq 0$ for each $i \in I(x_0)$ such that $x_0 = P_C\left(x - \sum_{i \in I(x_0)} \lambda_i a_i\right)$.

(iii) For any differentiable convex function $f : X \to \mathbb{R}$, x_0 minimizes f over S if and only if there exist $\lambda_i \geq 0$ for each $i \in I(x_0)$ such that

$$-f'(x_0) - \sum_{i \in I(x_0)} \lambda_i a_i \in \mathcal{N}_C(x_0).$$

Proof. The first assertion follows from (4.12). Note that the set $x_0 + \mathbf{T}^G(x_0)$ is the same as the set $\cap_{i \in I(x_0)} H_i$. By (2.2), (4.10) and (4.11), the equivalences of (i)-(iii) are clear from Theorem 3.2. □

References

1. de Boor, C., *On "Best" Interpolation*, Journal of Approximation Theory, Vol. 16, pp.28–42, 1976.

2. Chui, C. K., Deutsch, F., and Ward, J. D., *Constrained Best Approximation in Hilbert Space*, Constructive Approximation, Vol. 6, pp.35–64, 1990.

3. Chui, C. K., Deutsch, F., and Ward, J. D., *Constrained Best Approximation in Hilbert Space II*, Journal of Approximation Theory, Vol. 71, pp.231–238, 1992.

4. Deutsch, F., Li, W., and Ward, J. D., *A Dual Approach to Constrained Interpolation From a Convex Subset of Hilbert Space*, Journal of Approximation Theory, Vol. 90, pp.385–414, 1997.

5. Deutsch, F., Ubhaya, V., Ward, J. D., and Xu, Y., *Constrained Best Approximation in Hilbert Space* III*: Applications to n-convex Functions*, Constructive Approximation, Vol. 12, pp.361–384, 1996.

6. Micchelli, C. A., Smith, P. W., Swetits, J., and Ward, J. D., *Constrained L_p Approximation*, Constructive Approximation, Vol. 1, pp.93–102, 1985.

7. Micchelli, C. A., and Utreras, F. I., *Smoothing and Interpolation in a Convex Subset of a Hilbert Space*, SIAM Journal on Scientific and Statistical Computing, Vol. 9, pp.728–746, 1988.

8. Deutsch, F., Li, W., and Ward, J. D., *Best Approximation From the Intersection of a Closed Convex Set and a Polyhedron in Hilbert Space, Weak Slater Conditions, and the Strong Conical Hull Intersection Property*, SIAM Journal on Optimization, Vol. 10, pp.252–268, 1999.

9. Bonnans, J. F., and Shapiro, A., *Perturbation Analysis of Optimization Problems*, Springer Series in Operations Research. Springer-Verlag, New York, 2000.

10. Aubin, J.-P., and Frankowska, H., *Set-valued Analysis*, Birkhäuser Boston Inc., Boston, MA, 1990.

11. Clarke, F. H., *Optimization and Nonsmooth Analysis*, John Wiley & Sons Inc., New York, 1983.

12. Clarke, F. H., Ledyaev, Y. S., Stern, R. J., and Wolenski, P. R., *Nonsmooth Analysis and Control Theory*, Springer-Verlag, New York, 1998.

13. Robinson, S. M., *Stability Theory for Systems of Inequalities. II. Differentiable Nonlinear Systems*, SIAM Journal on Numerical Analysis, Vol. 13, pp.497–513, 1976.

14. Li, C., and Ng, K. F., *On Best Approximation by Nonconvex Sets and Perturbation of Nonconvex Inequality Systems in Hilbert Spaces*, SIAM Journal on Optimization, Vol. 13, pp.726–744, 2002.

15. Li, C., and Jin, X.-Q., *Nonlinearly Constrained Best Approximation in Hilbert Spaces: The Strong CHIP and the Basic Constraint Qualification*, SIAM Journal on Optimization, Vol. 13, pp.228–239, 2002.