Journal of Approximation Theory **97**, 140–157 (1999)
Article ID jath.1997.3243, available online at http://www.idealibrary.com on IDEAL®

Abadie's Constraint Qualification, Hoffman's Error Bounds, and Hausdorff Strong Unicity

M. Bartelt

Department of Mathematics, Christopher Newport University, Newport News, Virginia 23606
E-mail: mbartelt@pcs.cnu.edu

and

W. Li

Department of Mathematics and Statistics, Old Dominion University, Norfolk, Virginia 23529
E-mail: wuli@math.odu.edu

Communicated by Günther Nürnberger

Received May 23, 1997; accepted in revised form January 9, 1998

The main goal of this paper is to show the connection between optimization and best approximation when studying vector-valued functions defined on a finite set. For example, Hausdorff strong unicity for best approximation is shown to be equivalent to Abadie's constraint qualification for the associated convex quadratic feasibility problem. © 1999 Academic Press

1. INTRODUCTION

Let $I=\{1, ..., r\}$ be a finite set with the discrete topology and $C(I, \mathbb{R}^m)$ be the space of vector-valued functions from the index set I to the m-dimensional Euclidean space $\mathbb{R}^m$. Since I has the discrete topology, any function from I to $\mathbb{R}^m$ is continuous. For f in $C(I, \mathbb{R}^m)$, we write $f=(f_1, ..., f_m)$, where f_j is the jth component function of f. The value of f at i in I is a vector in $\mathbb{R}^m$: $f(i)=(f_1(i), ..., f_m(i))$. Note that we can identify a function f in $C(I, \mathbb{R}^m)$ with an $r\times m$ matrix whose ith row is $f(i)$. A natural norm for functions in $C(I, \mathbb{R}^m)$ is the following mixture of the ℓ_2-norm and the ℓ_∞-norm,

$$\|f\| := \max_{i\in I} \|f(i)\|_2 = \max_{i\in I} \left(\sum_{j=1}^{m} (f_j(i))^2 \right)^{1/2}, \tag{1}$$

where $\|\cdot\|_2$ denotes the Euclidean norm on $\mathbb{R}^m$.

0021-9045/99 $30.00

Let $\{g^1, ..., g^n\}$ be n linearly independent functions in $C(I, \mathbb{R}^m)$ and let $G := \text{span}\{g^1, ..., g^n\}$ be the n-dimensional subspace of $C(I, \mathbb{R}^m)$ generated by $g^1, ..., g^n$. For a function $f \in C(I, \mathbb{R}^m)$, consider the best approximation problem of finding a function g^* in G that solves the following minimization problem:

$$\min_{g \in G} \|f - g\|. \tag{2}$$

One special case of the above best approximation problem is the best approximation of complex-valued functions on a finite set I, since $C(I, \mathbb{C})$ can be identified with $C(I, \mathbb{R}^2)$ by using the isometric mapping: $f_1(x) + \mathbf{i} f_2(x) \to (f_1(x), f_2(x))$, where $\mathbf{i} = \sqrt{-1}$. Brosowski [6, p. 215] also studied the metric projection in vector-valued function spaces with emphasis on its connection to parametric semi-infinite optimization. In [20] Pinkus had a comprehensive analysis of uniqueness of best approximation in more general vector-valued function spaces.

The set of all best approximations of f in G is denoted as

$$P_G(f) := \{g^* \in G \colon \|f - g^*\| = \text{dist}(f, G)\}, \tag{3}$$

where $\text{dist}(f, G) := \min_{g \in G} \|f - g\|$ denotes the distance from f to G. In general, $P_G(f)$ contains infinitely many elements. In fact, $P_G(f)$ is a singleton for every f in $C(I, \mathbb{R}^m)$ if and only if G satisfies the generalized Haar condition introduced by Zukhovitskii and Stechkin [27] (cf. also [2]). In such a special case, we also have the so-called strong unicity of order 2 for $P_G(f)$ [2],

$$\|f - g\|^2 \geqslant \text{dist}(f, G)^2 + \gamma \cdot \text{dist}(g, P_G(f))^2 \qquad \text{for} \quad g \in G,$$

where γ is a positive constant.

In the general case when G is a subspace of a normed linear space, we say that $P_G(f)$ is Hausdorff strongly unique of order α, if there exists a positive constant γ (depending on f, α, and G) such that

$$\|f - g\|^\alpha \geqslant \text{dist}(f, G)^\alpha + \gamma \cdot \text{dist}(g, P_G(f))^\alpha \qquad \text{for} \qquad g \in G. \tag{4}$$

If (4) holds and $P_G(f)$ is a singleton, then we say that $P_G(f)$ is strongly unique of order α.

Strong unicity was first introduced by Newman and Shapiro [17] when G is a finite dimensional Haar subspace of $C(X, Y)$, a Banach space of continuous functions from a compact Hausdorff space X to Y (which is either the real line $\mathbb{R}$ or the complex plane $\mathbb{C}$). In this setting, $P_G(f)$ is always a singleton. Moreover, Newman and Shapiro proved that (4) holds with $\alpha=1$ if $Y=\mathbb{R}$ and (4) holds with $\alpha=2$ if $Y=\mathbb{C}$ [17]. In addition, Cline [8] gave a counterexample to show that (4) does not hold in general for $\alpha=1$ if G is a finite dimensional Haar subspace of $C(X, \mathbb{C})$. In [23], Schmidt established (4) for best monotone approximations of f by algebraic polynomials with $\alpha=2$. In this case, G is not a subspace but $P_G(f)$ is always a singleton. Schmidt also used Theorem 7(viii) in this paper with $\delta=1$ and $\alpha=2$ as the definition of strong unicity of order $\frac{1}{2}$. Meanwhile, Chalmers and Taylor [7] used (4) as the definition of strong unicity of order $1/\alpha$. Theorem 7 in this paper contains many different yet equivalent formulations of Hausdorff strong unicity of order α, some of which were called local strong unicity when $P_G(f)$ is a singleton (cf. [1, 10]). The generalization of strong unicity to the case when $P_G(f)$ is not a singleton was first studied in [16] and referred to as Hausdorff strong unicity. For other results on Hausdorff strong unicity, see [3, 19].

The main goal of this paper is to reformulate (2) as a system of convex quadratic inequalities (Theorem 2) and then to see how the theory on convex quadratic inequalities is related to the theory on the best approxi mation problem (2). In particular, we will show that the optimality condition corresponds to the Kolmogorov criterion (Theorem 6), Hoffman's error bounds correspond to Hausdorff strong unicity (Theorem 9), and Abadie's constraint qualification corresponds to the strong Kolmogorov criterion (Theorem 14). As a consequence, by using recently established theory on Hoffman's error bounds for approximate solutions of convex quadratic inequalities [15, 25], we obtain new results on the Hausdorff strong unicity of $P_G(f)$ (Theorems 9 and 10).

In particular, a Kolmogorov criterion is given for best approximations (Theorem 6) and a strong Kolmogorov criterion (called strong since it refers not to unicity but to strong unicity) is given for Hausdorff strong uniqueness. A strong Kolmogorov criterion of this type was first given in [4] and in terms of linear functionals in [26] (see also [18]).

2. CONVEX QUADRATIC INEQUALITIES VERSUS THE BEST APPROXIMATION PROBLEM

Note that each function in the n-dimensional subspace G of $C(I, \mathbb{R}^m)$ can be identified with a vector in $\mathbb{R}^n$. For convenience, for $x=(x_1, ..., x_n)\in\mathbb{R}^n$, define

$$g_x := \sum_{k=1}^{n} x_k g^k. \tag{5}$$

For each fixed f in $C(I, \mathbb{R}^m)$ and each fixed index i in I, we introduce the function

$$h_i(x) = \|f(i) - g_x(i)\|_2^2 - \operatorname{dist}(f, G)^2 \qquad \text{for} \quad x \in \mathbb{R}^n. \tag{6}$$

LEMMA 1. *The function $h_i(x)$ defined in (6) is a convex quadratic function.*

Proof. Note that

$$h_i(x) = \sum_{j=1}^{m} \left(f_j(i) - \sum_{k=1}^{n} x_k g_j^k(i) \right)^2 - \operatorname{dist}(f, G)^2.$$

Clearly $h_i(x)$ is a quadratic function of x. It is easy to check that $H_i(x) = \|\sum_{k=1}^n x_k g^k(i) - f(i)\|_2$ is a convex function of x by using the triangle inequality for distance and $\varphi(t) = t^2 - \operatorname{dist}(f, G)^2$ is a convex function of t. Then it follows that $h_i(x) = (\varphi \circ H_i)(x)$ is convex, since $(F \circ H)$ is a convex function if H is a convex function on $\mathbb{R}^n$ and F is a monotone increasing and convex function on $\mathbb{R}$ (cf. [22, Theorem 5.1; 11, Proposition 2.1.8]). ∎

For convex quadratic functions $h_1(x), ..., h_r(x)$, let us consider the following system of convex quadratic inequalities:

$$h_i(x) \leqslant 0 \qquad \text{for} \quad i = 1, ..., r. \tag{7}$$

We will use $S(f)$ to denote the solution set of (7), i.e.,

$$S(f) := \{x^* \in \mathbb{R}^n : h_i(x^*) \leqslant 0 \text{ for } i = 1, ..., r\}. \tag{8}$$

Then we can prove that (7) is an equivalent reformulation of (2) and thus the best approximation problem (2) is equivalent to a convex quadratic feasibility problem.

THEOREM 2. *For any $f \in C(I, \mathbb{R}^m)$, $P_G(f) = \{g_{x^*} : x^* \in S(f)\}$, i.e., $x^* \in S(f)$ if and only if $g_{x^*} \in P_G(f)$.*

Proof. Let $x^* \in S(f)$. Then $h_i(x^*) \leqslant 0$ for $1 \leqslant i \leqslant r$, which implies $\|f - g_{x^*}\|^2 - \operatorname{dist}(f, G)^2 \leqslant 0$. That is, $\|f - g_{x^*}\| \leqslant \operatorname{dist}(f, G)$ and $g_{x^*} \in P_G(f)$. On the other hand, if $g_{x^*} \in P_G(f)$, then $\|f - g_{x^*}\| \leqslant \operatorname{dist}(f, G)$ or $\|f - g_{x^*}\|^2 - \operatorname{dist}(f, G)^2 \leqslant 0$, which implies $h_i(x^*) \leqslant 0$ for $1 \leqslant i \leqslant r$. Thus, $x^* \in S(f)$. This completes the proof of Theorem 2. ∎

Let $\langle \cdot, \cdot \rangle$ denote the dot product on $\mathbb{R}^n$ and let $\nabla h_i(x)$ denote the gradient of $h_i(x)$. The following lemma shows a useful relation between $\nabla h_i(x)$ and g_x.

LEMMA 3. *Let $f \in C(I, \mathbb{R}^m)$, $i \in I$, and $x, x^* \in \mathbb{R}^n$. Then*

$$\langle \nabla h_i(x^*), x \rangle = -2 \langle f(i) - g_{x^*}(i), g_x(i) \rangle. \tag{9}$$

Moreover, $g_x(i) = g_{x^*}(i)$ if and only if $\nabla h_i(x^*) = \nabla h_i(x)$.

Proof. By the chain rule for functions of several variables, we have

$$\frac{\partial h_i(x^*)}{\partial x_k} = -2 \langle f(i) - g_{x^*}(i), g^k(i) \rangle \qquad \text{for} \quad 1 \leqslant k \leqslant n, \tag{10}$$

which implies (9).

By (10), $\nabla h_i(x^*) = \nabla h_i(x)$ if and only if

$$-2 \langle f(i) - g_{x^*}(i), g^k(i) \rangle = -2 \langle f(i) - g_x(i), g^k(i) \rangle \qquad \text{for} \quad 1 \leqslant k \leqslant n,$$

i.e.,

$$\langle g_x(i) - g_{x^*}(i), g^k(i) \rangle = 0 \qquad \text{for} \quad 1 \leqslant k \leqslant n. \tag{11}$$

Thus, if $g_x(i) = g_{x^*}(i)$, then (11) holds, which implies $\nabla h_i(x^*) = \nabla h_i(x)$. On the other hand, if $\nabla h_i(x^*) = \nabla h_i(x)$, then (11) holds. In particular, we have

$$\begin{aligned} \|g_x(i) - g_{x^*}(i)\|_2^2 &= \langle g_x(i) - g_{x^*}(i), g_x(i) - g_{x^*}(i) \rangle \\ &= \sum_{k=1}^{n} (x_k - x_k^*) \langle g_x(i) - g_{x^*}(i), g^k(i) \rangle = 0, \end{aligned}$$

or $g_x(i) = g_{x^*}(i)$. ∎

3. OPTIMALITY CONDITION VERSUS KOLMOGOROV CRITERION

Let the set of extreme points of $(f - g)$ for any g in G be denoted by

$$E(f - g) = \{ i \in I \colon \|f(i) - g(i)\|_2 = \|f - g\| \}.$$

On the other hand, the active index set for $x^* \in S(f)$ is defined by $J(x^*) := \{ i \in I \colon h_i(x^*) = 0 \}$. For $x^* \in S(f)$ and $g_{x^*} \in P_G(f)$, we have

$$\|f(i) - g_{x^*}(i)\|_2^2 - \|f - g_{x^*}\|^2 = \|f(i) - g_{x^*}(i)\|_2^2 - \operatorname{dist}(f, G)^2 = h_i(x^*). \tag{12}$$

Thus, $J(x^*) = E(f - g_{x^*})$ for $x^* \in S(f)$. Moreover, the common extreme point set for $(f - P_G(f))$ (or the common active index set for all x^* in $S(f)$) is

$$J(f) := \bigcap_{g^* \in P_G(f)} E(f - g^*) \equiv \bigcap_{x^* \in S(f)} J(x^*). \tag{13}$$

Note that $J(f)$ is always nonempty (cf. [14, Lemma 2.2] or Lemma 13 in the paper).

THEOREM 4 (Classical Characterization). *Let $x^* \in \mathbb{R}^n$. Then the following statements are equivalent.*

(i) $x^* \in S(f)$.

(ii) *There exist $\theta_i > 0$ for i in a nonempty subset J of $J(x^*)$ such that*

$$\sum_{i \in J} \theta_i \nabla h_i(x^*) = 0. \tag{14}$$

(iii) $g_{x^*} \in P_G(f)$.

(iv) *There exist $\theta_i > 0$ for i in a nonempty subset J of $E(f - g_{x^*})$ such that*

$$\sum_{i \in J} \theta_i \langle f(i) - g_{x^*}(i), g(i) \rangle = 0 \qquad \text{for} \quad g \in G. \tag{15}$$

Proof. Define the convex piecewise quadratic function,

$$h(x) = \max_{1 \leqslant i \leqslant n} h_i(x).$$

Then we have the following relation between $h(x)$ and $\|f - g_x\|$,

$$\|f - g_x\|^2 - \text{dist}(f, G)^2 = h(x) \qquad \text{for} \quad x \in \mathbb{R}^n. \tag{16}$$

Thus, $h(x) \geqslant 0$ for $x \in \mathbb{R}^n$; so $x^* \in S(f)$ if and only if $h(x^*) = 0$. Hence, $S(f)$ is the set of (global) minimizers of the convex function $h(x)$ on $\mathbb{R}^n$. By [22, Theorem 28.3; 11, Theorem 2.2.1, p. 253], $x^* \in S(f)$ if and only if

$$0 \in \partial h(x^*), \tag{17}$$

where $\partial h(x^*)$ denotes the subgradient of h at x^*. However, it is known (cf. [11, Corollary 4.4.4]) that

$$\partial h(x^*) = \left\{ \sum_{i \in J(x^*)} \theta_i \nabla h_i(x^*) \colon \theta_i \geqslant 0, \sum_{i \in J(x^*)} \theta_i = 1 \right\}. \tag{18}$$

Therefore, (17) holds if and only if (ii) holds. This proves the equivalence of (i) and (ii).

The equivalence of (i) and (iii) follows from Theorem 2, and the equivalence of (iii) and (iv) is the standard characterization of best approximations (cf. [9]). This completes the proof of Theorem 4. ∎

Remark. By (10) and $J(x^*) = E(f - g_{x^*})$, it is easy to see that (ii) is just a reformulation of (iv), which is a version of the characterization of best approximation given by Rivlin and Shapiro (cf. [21, Theorem 2]). For the equivalence of more general forms of (ii) and (iv) in any normed linear spaces, see [24, p. 170].

The characterization (15) for g_{x^*} in $P_G(f)$ can be reformulated as the Kolmogorov criterion. But we first need the following consequence of Gordon's Theorem (cf. [5, Corollary 1, p. 47]), which is also a consequence of the so-called generalized Farkas Lemma (cf. [11, first comment on p. 132]).

LEMMA 5. *Let K be a finite subset of $\mathbb{R}^n$ and let $\mathrm{co}(K)$ be the convex hull of K. Then $0 \in \mathrm{co}(K)$ if and only if*

$$\min_{y \in K} \langle y, x \rangle \leqslant 0 \qquad \textit{for} \quad x \in \mathbb{R}^n.$$

THEOREM 6 (Kolmogorov Criterion). *Let $x^* \in \mathbb{R}^n$. Then the following statements are equivalent.*

(i) $x^* \in S(f)$.

(ii) $\min_{i \in J(x^*)} \langle \nabla h_i(x^*), x \rangle \leqslant 0$ *for* $x \in \mathbb{R}^n$.

(iii) $g_{x^*} \in P_G(f)$.

(iv) $\max_{i \in E(f - g_{x^*})} \langle f(i) - g_{x^*}(i), g(i) \rangle \geqslant 0$ *for* $g \in G$.

Proof. Let $K := \{\nabla h_i(x^*) : i \in J(x^*)\}$. By Theorem 4, (i) holds if and only if $0 \in \mathrm{co}(K)$. However, by Lemma 5 and the choice of K, we know that $0 \in \mathrm{co}(K)$ if and only if (ii) holds. Therefore, (i) is equivalent to (ii). The equivalence of (i) and (iii) is Theorem 2. Also, by (9) and $J(x^*) = E(f - g_{x^*})$, it is easy to see that (ii) holds if and only if (iv) holds. ∎

Remark. Notice that the equivalence of (iii) and (iv) is a generalization of the classical Kolmogorov criterion [13] in the sense that we use the inner product for vector-valued functions instead of the scalar product for real-valued functions.

4. VARIOUS FORMS OF HAUSDORFF STRONG UNICITY

Recall that $P_G(f)$ is called Hausdorff strongly unique of order α, if there exists a positive constant $\gamma(f, \alpha, G)$ such that

$$\|f-g\|^\alpha \geqslant \operatorname{dist}(f, G)^\alpha + \gamma(f, \alpha, G) \cdot \operatorname{dist}(g, P_G(f))^\alpha \qquad \text{for} \quad g \in G. \tag{19}$$

The following theorem gives equivalent formulations of (19).

THEOREM 7. *Let $f \in C(I, \mathbb{R}^m)\backslash G$ and $\alpha \geqslant 1$. Then the following statements are equivalent.*

(i) *$P_G(f)$ is Hausdorff strongly unique of order α.*

(ii) *For any $\eta > 0$, there exists $\gamma(\eta) > 0$ such that*

$$\|f-g\|^\alpha \geqslant \operatorname{dist}(f, G)^\alpha + \gamma(\eta) \cdot \operatorname{dist}(g, P_G(f))^\alpha \qquad \textit{for} \quad g \in G,\ \|g\| \leqslant \eta.$$

(iii) *For any $\eta > 0$ and any $\delta > 0$, there exists $\gamma(\eta, \delta) > 0$ such that*

$$\|f-g\|^\delta \geqslant \operatorname{dist}(f, G)^\delta + \gamma(\eta, \delta) \cdot \operatorname{dist}(g, P_G(f))^\alpha \qquad \textit{for} \quad g \in G,\ \|g\| \leqslant \eta.$$

(iv) *For any $\varepsilon > 0$ and any $\delta > 0$, there exists $\gamma(\varepsilon, \delta) > 0$ such that*

$$\|f-g\|^\delta \geqslant \operatorname{dist}(f, G)^\delta + \gamma(\varepsilon, \delta) \cdot \operatorname{dist}(g, P_G(f))^\alpha \qquad \textit{for} \quad g \in G,\ \operatorname{dist}(g, P_G(f)) \leqslant \varepsilon. \tag{20}$$

(v) *For some $\varepsilon > 0$ and some $\delta > 0$, there exists $\gamma(\varepsilon, \delta) > 0$ such that (20) holds.*

(vi) *For any $\beta \geqslant \alpha$, there exists $\gamma(\beta) > 0$ such that*

$$\|f-g\|^\beta \geqslant \operatorname{dist}(f, G)^\beta + \gamma(\beta) \cdot \operatorname{dist}(g, P_G(f))^\alpha \qquad \textit{for} \quad g \in G. \tag{21}$$

(vii) *For some $\beta \geqslant \alpha$, there exists $\gamma(\beta) > 0$ such that (21) holds.*

(viii) *For some fixed $\delta > 0$ and any $\eta > 0$, there exists $\gamma(\eta, \delta) > 0$ such that*

$$\|f-g\|^\delta \geqslant \operatorname{dist}(f, G)^\delta + \gamma(\eta, \delta) \cdot \operatorname{dist}(g, P_G(f))^\alpha \qquad \textit{for} \quad g \in G,\ \|g\| \leqslant \eta. \tag{22}$$

Proof. For two positive numbers s and t, define

$$\varphi_{s,t}(g) := \frac{\|f-g\|^s - \operatorname{dist}(f, G)^s}{\|f-g\|^t - \operatorname{dist}(f, G)^t}$$

and

$$\psi_{s,t}(g) := \frac{\|f-g\|^s - \operatorname{dist}(f, G)^s}{\operatorname{dist}(g, P_G(f))^t}.$$

Then by L'Hôpital's Rule applied to the quotient of real-valued functions

$$\frac{F(w)}{G(w)} = \frac{w^s - \operatorname{dist}(f, G)^s}{w^t - \operatorname{dist}(f, G)^t},$$

we get

$$\lim_{\operatorname{dist}(g, P_G(f)) \to 0} \varphi_{s,t}(g) = \frac{s \cdot \operatorname{dist}(f, G)^{s-1}}{t \cdot \operatorname{dist}(f, G)^{t-1}} > 0.$$

Thus, there exist positive constants $\varepsilon(s, t)$ and $\kappa(s, t)$ such that

$$\varphi_{s,t}(g) \geqslant \kappa(s, t) \qquad \text{for} \quad g \in G, \operatorname{dist}(g, P_G(f)) \leqslant \varepsilon(s, t). \tag{23}$$

Let $s \geqslant t > 0$, $\|g\| > \|f\|$, and $g^* \in P_G(f)$. Then

$$\begin{aligned} \psi_{s,t}(g) &\geqslant \frac{\|f-g\|^s - \|f-g^*\|^s}{\|g-g^*\|^t} \geqslant \frac{(\|g\| - \|f\|)^s - \|f-g^*\|^s}{(\|g\| + \|g^*\|)^t} \\ &\geqslant \frac{(\|g\| - \|f\|)^s - \|f-g^*\|^s}{(\|g\| + \|g^*\| + 1)^t} \geqslant \frac{(\|g\| - \|f\|)^s - \|f-g^*\|^s}{(\|g\| + \|g^*\| + 1)^s}. \end{aligned}$$

It follows that

$$\liminf_{\|g\| \to \infty} \psi_{s,t}(g) \geqslant \lim_{\|g\| \to \infty} \frac{(\|g\| - \|f\|)^s - \|f-g^*\|^s}{(\|g\| + \|g^*\| + 1)^s} = 1.$$

Thus, for any $s \geqslant t > 0$, there exists a positive constant $\eta(s, t)$ such that

$$\psi_{s,t}(g) \geqslant \tfrac{1}{2} \qquad \text{for} \quad g \in G, \|g\| \geqslant \eta(s, t). \tag{24}$$

For any $\eta > 0$ and $\varepsilon > 0$, $\varphi_{s,t}(g)$ and $\psi_{s,t}(g)$ are positive continuous functions on the compact subset $\{g \in G\colon \operatorname{dist}(g, P_G(f)) \geqslant \varepsilon, \|g\| \leqslant \eta\}$. Thus, there is a positive constant $\lambda(s, t, \varepsilon, \eta)$ such that

$$\varphi_{s,t}(g) \geqslant \lambda(s, t, \varepsilon, \eta) \qquad \text{for} \quad g \in G, \quad \operatorname{dist}(g, P_G(f)) \geqslant \varepsilon, \quad \|g\| \leqslant \eta, \tag{25}$$

$$\psi_{s,t}(g) \geqslant \lambda(s, t, \varepsilon, \eta) \qquad \text{for} \quad g \in G, \quad \operatorname{dist}(g, P_G(f)) \geqslant \varepsilon, \quad \|g\| \leqslant \eta. \tag{26}$$

Now we are ready to prove the equivalence of all the statements.

(i) $\Rightarrow$ (ii). It is obvious that (i) implies (ii).

(ii) $\Rightarrow$ (iii). For any $\eta > 0$ and any $\delta > 0$, by applying (23) and (25) with $s = \delta$, $t = \alpha$, and $\varepsilon = \varepsilon(\delta, \alpha)$, we get that for $g \in G$ with $\|g\| \leqslant \eta$,

$$\|f - g\|^\delta - \operatorname{dist}(f, G)^\delta \geqslant \min\{\kappa(\delta, \alpha), \lambda(\delta, \alpha, \varepsilon(\delta, \alpha), \eta)\}(\|f - g\|^\alpha - \operatorname{dist}(f, G)^\alpha).$$

Thus, (ii) implies (iii) with $\gamma(\delta, \eta) = \gamma(\eta) \min\{\kappa(\delta, \alpha), \lambda(\delta, \alpha, \varepsilon(\delta, \alpha), \eta)\}$.

(iii) $\Rightarrow$ (iv) $\Rightarrow$ (v). It is easy to see that (iii) $\Rightarrow$ (iv) $\Rightarrow$ (v).

(v) $\Rightarrow$ (vi). Let $\beta \geqslant \alpha$. Then (24) holds for $s = \beta$ and $t = \alpha$. That is,

$$\|f - g\|^\beta \geqslant \operatorname{dist}(f, G)^\beta + \tfrac{1}{2} \operatorname{dist}(g, P_G(f))^\alpha \qquad \text{for} \quad g \in G, \quad \|g\| \geqslant \eta(\beta, \alpha). \tag{27}$$

By applying (23) with $s = \beta$ and $t = \delta$, we get

$$\|f - g\|^\beta - \operatorname{dist}((f, G)^\beta \geqslant \kappa(\beta, \delta)(\|f - g\|^\delta - \operatorname{dist}(f, G)^\delta) \qquad \text{for} \quad g \in G, \quad \operatorname{dist}(g, P_G(f)) \leqslant \varepsilon,$$

which, along with (v), implies

$$\|f - g\|^\beta - \operatorname{dist}(f, G)^\beta \geqslant \kappa(\beta, \delta) \cdot \gamma(\delta, \varepsilon) \cdot \operatorname{dist}(g, P_G(f))^\alpha \qquad \text{for} \quad g \in G, \quad \operatorname{dist}(g, P_G(f)) \leqslant \varepsilon. \tag{28}$$

By applying (26) with $s = \beta$, $t = \alpha$, and $\eta = \eta(\beta, \alpha)$, we get that

$$\|f - g\|^\beta - \operatorname{dist}(f, G)^\beta \geqslant \lambda(\beta, \delta, \varepsilon, \eta(\beta, \alpha)) \cdot \operatorname{dist}(g, P_G(f))^\alpha, \tag{29}$$

whenever $g \in G$, $\operatorname{dist}(g, P_G(f)) \geqslant \varepsilon$, and $\|g\| \leqslant \eta(\beta, \alpha)$. Thus, (vi) follows from (27)–(29). This proves that (v) implies (vi).

(vi) $\Rightarrow$ (vii). It is obvious that (vi) implies (vii).

(vii) $\Rightarrow$ (i). Note that (vii) implies (v) with $\delta = \beta$. By (v) $\Rightarrow$ (vi) and (vi) $\Rightarrow$ (i), we know that (vii) $\Rightarrow$ (i).

(viii) $\Leftrightarrow$ (i). It is obvious that (iii) $\Rightarrow$ (viii). Since (i) $\Leftrightarrow$ (iii), we have (i) $\Rightarrow$ (viii). On the other hand, (viii) $\Rightarrow$ (v). By (v) $\Leftrightarrow$ (i), we get (viii) $\Rightarrow$ (i). ∎

Remark. The proof of the above theorem is valid if G is a finite-dimensional subspace of any normed linear space. When $\alpha = 2$, $\delta = 1$, and $m = 2$ (i.e., $C(I, \mathbb{C})$), Newman and Shapiro also stated what is easily seen to be equivalent to Theorem 7(viii) as a characterization of strong unicity of order $\frac{1}{2}$.

5. HOFFMAN'S ERROR BOUNDS VERSUS HAUSDORFF STRONG UNICITY

In this section, we show that Hausdorff strong unicity corresponds to Hoffman's error bounds for approximate solutions of convex quadratic inequalities (7). Then by using the reformulation (7) of (2) and a result on Hoffman's error bounds for convex quadratic inequalities [25], we prove that $P_G(f)$ is always Hausdorff strongly unique of order 2^{m+1}.

LEMMA 8. *There is a positive constant c such that*

$$\frac{1}{c}\cdot\|x\|_2\leqslant\|g_x\|\leqslant c\cdot\|x\|_2 \qquad \text{for} \quad x\in\mathbb{R}^n, \tag{30}$$

$$\frac{1}{c}\cdot\operatorname{dist}(x, S(f))\leqslant\operatorname{dist}(g_x, P_G(f))\leqslant c\cdot\operatorname{dist}(x, S(f)) \qquad \text{for} \quad x\in\mathbb{R}^n. \tag{31}$$

Proof. Since $\{g^1, ..., g^n\}$ is a basis of G, for any $x\in\mathbb{R}^n$, $\|x\|_2$ is also a norm for the function g_x in G. Therefore, there exists a positive constant c such that (30) holds, since any two norms on the finite-dimensional subspace G are equivalent. By Theorem 2, we obtain

$$\operatorname{dist}(g_x, P_G(f))=\min_{g^*\in P_G(f)}\|g_x-g^*\|=\min_{x^*\in S(f)}\|g_x-g_{x^*}\|. \tag{32}$$

By (30), we obtain

$$\begin{aligned}\frac{1}{c}\cdot\operatorname{dist}(x, S(f))&=\frac{1}{c}\cdot\min_{x^*\in S(f)}\|x-x^*\|_2\leqslant\min_{x^*\in S(f)}\|g_x-g_{x^*}\|\\&\leqslant c\cdot\min_{x^*\in S(f)}\|x-x^*\|_2\leqslant c\cdot\operatorname{dist}(x, S(f)).\end{aligned}$$

This, along with (32), proves (31). ∎

By Lemma 8, we can reformulate Hausdorff strong unicity of order α in terms of Hoffman's error bounds.

THEOREM 9. *Let $f\in C(I, \mathbb{R}^m)\backslash G$ and $\alpha\geqslant 1$. Then the following statements are equivalent.*

(i) *$P_G(f)$ is Hausdorff strongly unique of order α.*

(ii) *There exists $\lambda>0$ such that*

$$\operatorname{dist}(x, S(f))\leqslant\lambda\big(\max_{1\leqslant i\leqslant r} h_i(x)+[\max_{1\leqslant i\leqslant r} h_i(x)]^{1/\alpha}\big) \qquad \text{for} \quad x\in\mathbb{R}^n. \tag{33}$$

(iii) *For any* $\eta > 0$, *there exists* $\lambda(\eta) > 0$ *such that*

$$\operatorname{dist}(x, S(f)) \leqslant \lambda(\eta) (\max_{1 \leqslant i \leqslant r} h_i(x))^{1/\alpha} \qquad \text{for} \quad x \in \mathbb{R}^n,\ \|x\|_2 \leqslant \eta. \tag{34}$$

Proof. Let $h(x) = \max_{1 \leqslant i \leqslant r} h_i(x)$. Then

$$\|f - g_x\|^2 - \operatorname{dist}(f, G)^2 = h(x) \qquad \text{for} \quad x \in \mathbb{R}^n. \tag{35}$$

By Theorem 7, $P_G(f)$ is Hausdorff Strongly unique of order α if and only if for any $\hat{\eta} > 0$, there is a positive constant $\gamma(\hat{\eta})$ such that

$$\|f - g_x\|^2 - \operatorname{dist}(f, G)^2 \geqslant \gamma(\hat{\eta}) \cdot \operatorname{dist}(g_x, P_G(f))^\alpha \qquad \text{for} \quad x \in \mathbb{R}^n, \quad \|g_x\| \leqslant \hat{\eta}. \tag{36}$$

By (35) and Lemma 8, (36) holds if and only if for any $\eta > 0$, there is a positive constant $\lambda(\eta)$ such that

$$\operatorname{dist}(x, S(f))^\alpha \leqslant \lambda(\eta) \cdot h(x) \qquad \text{for} \quad x \in \mathbb{R}^n, \quad \|x\|_2 \leqslant \eta. \tag{37}$$

This proves the equivalence of (i) and (iii).

(ii) $\Rightarrow$ (iii). For any $\eta > 0$, since the continuous function $h(x)$ is bounded on the compact set $\{x \in \mathbb{R}^n : \|x\|_2 \leqslant \eta\}$, there is a positive constant $\kappa(\eta)$ such that $h(x) \leqslant \kappa(\eta)$ for $\|x\|_2 \leqslant \eta$. Thus, it follows from (ii) that

$$\begin{aligned} \operatorname{dist}(x, S(f)) &\leqslant \lambda(h(x) + h(x)^{1/\alpha}) \leqslant \lambda(h(x)^{1/\alpha} \kappa(\eta)^{(\alpha-1)/\alpha} + h(x)^{1/\alpha}) \\ &= \lambda(\kappa(\eta)^{(\alpha-1)/\alpha} + 1)\, h(x)^{1/\alpha}, \end{aligned}$$

whenever $\|x\|_2 \leqslant \eta$. So, (iii) holds.

(iii) $\Rightarrow$ (ii). By applying (24) with $s = 2$ and $t = 1$, there is a positive constant $\hat{\eta}$ such that

$$h(x) \geqslant \tfrac{1}{2} \cdot \operatorname{dist}(g_x, P_G(f)) \qquad \text{for} \quad \|g_x\| \geqslant \hat{\eta}.$$

By (30) and (31), the above inequality implies

$$h(x) \geqslant \frac{1}{2c} \cdot \operatorname{dist}(x, S(f)) \qquad \text{for} \quad \|x\|_2 \geqslant c\hat{\eta}. \tag{38}$$

By applying (iii) with $\eta = c\hat{\eta}$, we obtain

$$\operatorname{dist}(x, S(f)) \leqslant \lambda(c\hat{\eta}) \cdot h(x)^{1/\alpha} \qquad \text{for} \quad x \in \mathbb{R}^n,\ \|x\|_2 \leqslant c\hat{\eta}. \tag{39}$$

It follows from (38) and (39) that (ii) holds with $\lambda = \max\{c\hat{\eta}, 2c\}$. ∎

Remark. When $\alpha = 1$ and the functions h_i are affine functions, (33) is the original Hoffman's error bound for approximate solutions of linear inequalities [12].

In general, (33) does not hold for $\alpha = 1$ [15, 25]. However, it follows from a general result on error bounds for approximate solutions of convex quadratic inequalities by Wang and Pang [25] that (33) holds with $\alpha = 2^{m+1}$. Then by Theorem 9 we obtain the following result on Hausdorff strong unicity of order 2^{m+1}.

THEOREM 10 [25]. *For any f and G, (33) holds with $\alpha = 2^{m+1}$. That is, $P_G(f)$ is Hausdorff strongly unique of order 2^{m+1}.*

6. ABADIE'S CONSTRAINT QUALIFICATION VERSUS THE STRONG KOLMOGOROV CRITERION

DEFINITION 11. We say that the convex quadratic inequality system (7) satisfies Abadie's constraint qualification if

$$N_{S(f)}(x^*) = \left\{ \sum_{i \in J(x^*)} \theta_i \nabla h_i(x^*) \colon \theta_i \geqslant 0 \right\} \qquad \text{for} \quad x^* \in S(f), \tag{40}$$

where $N_{S(f)}(x^*)$ is the normal cone of $S(f)$ at x^* defined as

$$N_{S(f)}(x^*) := \{ y \in \mathbb{R}^n \colon \langle y, z - x^* \rangle \leqslant 0 \text{ for } z \in S(f) \}.$$

Remark. It is always true that

$$N_{S(f)}(x^*) \supset \left\{ \sum_{i \in J(x^*)} \theta_i \nabla h_i(x^*) \colon \theta_i \geqslant 0 \right\} \qquad \text{for} \quad x^* \in S(f). \tag{41}$$

Thus, (40) is actually equivalent to

$$N_{S(f)}(x^*) \subset \left\{ \sum_{i \in J(x^*)} \theta_i \nabla h_i(x^*) \colon \theta_i \geqslant 0 \right\} \qquad \text{for} \quad x^* \in S(f). \tag{42}$$

LEMMA 12 [15, Theorem 12]. *The convex quadratic inequality system (7) satisfies Abadie's constraint qualification if and only if there is a positive constant λ such that*

$$\operatorname{dist}(x, S(f)) \leqslant \lambda \cdot \max_{1 \leqslant i \leqslant r} (h_i(x))_+ \qquad \textit{for} \quad x \in \mathbb{R}^n. \tag{43}$$

Recall that $J(f)$ denotes the common extreme point set for $(f-P_G(f))$. The following result about $J(f)$ is known (cf. [14, Lemma 2.2]). In fact, any g^* in the relative interior of $P_G(f)$ satisfies (44).

LEMMA 13. *For any $f \in C(I, \mathbb{R}^m)$, there exists a function g^* in $P_G(f)$ such that*

$$E(f-g^*)=J(f)\subset\{i\in I\colon g^*(i)=g(i) \quad \textit{for all} \quad g\in P_G(f)\}. \tag{44}$$

THEOREM 14. *Let $f \in C(I, \mathbb{R}^m)\backslash G$ and $x^* \in S(f)$. Then the following statements are equivalent.*

(i) *The convex quadratic inequality system* (7) *satisfies Abadie's constraint qualification.*

(ii) *The following strong Kolmogorov criterion is satisfied,*

$$\max_{i\in J(f)} \langle f(i)-g_{x^*}(i), g(i)\rangle > 0, \tag{45}$$

whenever $g \in G$ is not identical to 0 *on $J(f)$.*

(iii) *The following strong Kolmogorov criterion is satisfied,*

$$\min_{i\in J(f)} \langle \nabla h_i(x^*), x\rangle < 0, \tag{46}$$

whenever g_x is not identical to 0 *on $J(f)$.*

(iv) *Hoffman's error bound* (43) *holds.*

(v) *$P_G(f)$ is Hausdorff strongly unique* (*of order $\alpha=1$*).

Proof. Let x^* denote the particular element in $S(f)$ given in Lemma 13 so that we have

$$h_i(x^*)<0 \qquad \text{for} \quad i\notin J(f). \tag{47}$$

By Theorem 9 and Lemma 12, we know that (i) $\Leftrightarrow$ (iv) $\Leftrightarrow$ (v). By (9) we know that (ii) is equivalent to (iii).

Next we prove (iii) $\Rightarrow$ (i). Consider the following system of linear equalities and convex quadratic inequalities:

$$\langle \nabla h_i(x^*), x-x^*\rangle = 0 \qquad \text{for} \quad i\in J(f) \quad \text{and} \quad h_i(x)\leqslant 0 \quad \text{for } i\notin J(f). \tag{48}$$

If $x \in S(f)$, then $g_x(i)-g_{x^*}(i)=0$ for $i\in J(f)$ (cf. Lemma 13). By (9), we have

$$J(f)=\{i\in J(f)\colon \langle \nabla h_i(x^*), x-x^*\rangle = 0\} \qquad \text{for} \quad x\in S(f). \tag{49}$$

Thus, x satisfies (48). On the other hand, assume that x satisfies (48). By (iii), the first set of equalities in (48) implies $g_x(i)=g_{x^*}(i)$ for $i\in J(f)$. Thus, $h_i(x)=h_i(x^*)\leqslant 0$ for $i\in J(f)$. This, along with the second set of inequalities in (48), implies $x\in S(f)$. Thus, $S(f)$ is the set of solutions to (48). By (47), (48) satisfies the weak Slater condition and hence Abadie's constraint qualification (cf. [11, p. 311]),

$$N_{S(f)}(x)=\left\{\sum_{i\in J(f)}\theta_i\,\nabla h_i(x^*)+\sum_{i\in \hat{J}(x)\setminus J(f)}\theta_i\,\nabla h_i(x)\colon \theta_i\geqslant 0\right\}\quad \text{for}\quad x\in S(f), \tag{50}$$

where

$$\hat{J}(x)=\{i\notin J(f)\colon h_i(x)=0\}\cup J(f).$$

Since $J(f)\subset J(x)$ for $x\in S(f)$, we have $\hat{J}(x)=J(x)$. Thus, we get the following representation of $N_{S(f)}(x)$:

$$N_{S(f)}(x)=\left\{\sum_{i\in J(x)}\theta_i\,\nabla h_i(x)\colon \theta_i\geqslant 0\right\}\qquad \text{for}\quad x\in S(f).$$

That is, (7) satisfies Abadie's constraint qualification.

Finally, we prove that (v) implies (ii). If (ii) does not hold, then there exists $\hat{g}\in G$ and $\bar{\imath}\in J(f)$ such that $\hat{g}(\bar{\imath})\neq 0$ but

$$\max_{i\in J(f)}\langle f(i)-g_{x^*}(i),\, \hat{g}(i)\rangle\leqslant 0, \tag{51}$$

By (47), there is a positive constant ε such that

$$\|f(i)-[g_{x^*}(i)-t\hat{g}(i)]\|_2^2<\operatorname{dist}(f,G)^2\qquad \textit{for}\quad i\notin J(f),\, 0\leqslant t\leqslant\varepsilon. \tag{52}$$

For $i\in J(f)$, we have

$$\|f(i)-[g_{x^*}(i)-t\hat{g}(i)]\|_2^2=\|f(i)-g_{x^*}(i)\|_2^2 +2t\langle f(i)-g_{x^*}(i),\,\hat{g}(i)\rangle+t^2\,\|\hat{g}(i)\|_2^2. \tag{53}$$

It follows from (51) and (53) that

$$\|f(i)-[g_{x^*}(i)-t\hat{g}(i)]\|_2^2\leqslant\operatorname{dist}(f,G)^2+t^2\,\|\hat{g}(i)\|_2^2\qquad \text{for}\quad i\in J(f). \tag{54}$$

Thus, by (52) and (54), we obtain

$$\|f-[g_{x^*}-t\hat{g}]\|^2\leqslant\operatorname{dist}(f,G)^2+t^2\,\|\hat{g}\|^2\qquad \text{for}\quad 0\leqslant t\leqslant\varepsilon. \tag{55}$$

However, by Lemma 13,

$$\operatorname{dist}([g_{x^*}-t\hat{g}], P_G(f)) \geqslant \|[g_{x^*}(\bar{\imath})-t\hat{g}(\bar{\imath})]-g_{x^*}(\bar{\imath})\|_2 = t\,\|\hat{g}(\bar{\imath})\|_2. \qquad (56)$$

But by (v) and Theorem 7, there is a positive constant γ such that

$$\|f-g\|^2 \geqslant \operatorname{dist}(f, G)^2+\gamma \cdot \operatorname{dist}(g, P_G(f)) \qquad \text{for} \quad g \in G. \qquad (57)$$

Replacing g by $[g_{x^*}-t\hat{g}]$ in (57) and using inequalities (55) and (56), we get

$$\operatorname{dist}(f, G)^2+t^2\,\|\hat{g}\|^2 \geqslant \operatorname{dist}(f, G)^2+t\gamma\,\|\hat{g}(\bar{\imath})\|_2 \qquad \text{for} \quad 0 \leqslant t \leqslant \varepsilon,$$

i.e.,

$$t\,\|\hat{g}\|^2 \geqslant \gamma\,\|\hat{g}(\bar{\imath})\|_2 \qquad \text{for} \quad 0 \leqslant t \leqslant \varepsilon,$$

which is impossible since $\gamma\,\|\hat{g}(\bar{\imath})\|_2>0$. The contradiction proves that (ii) holds. ∎

Remark. The classical strict Kolmogorov criterion is given in the following way [18],

$$\max_{i \in E(f-g_{x^*})} \langle f(i)-g_{x^*}(i), g(i)\rangle > 0 \qquad \text{for} \quad g \in G, \qquad (58)$$

which is usually used to study the uniqueness of best approximation [18, Lemma 2.1]. However, in our special setting, $E(f-g_{x^*})$ is a finite set. As a consequence, (58) is equivalent to the following strong Kolmogorov criterion,

$$\max_{i \in E(f-g_{x^*})} \langle f(i)-g_{x^*}(i), g(i)\rangle \geqslant \beta\,\|g\| \qquad \text{for} \quad g \in G, \qquad (59)$$

where β is a positive constant. By the characterization of strongly unique best approximation by Wulbert [26] for real normed linear spaces or by Bartelt and McLaughlin [4] for complex normed linear spaces, $P_G(f)$ is strongly unique (of order $\alpha=1$) if and only if (59) (or (58)) holds. Thus, the equivalence of (ii) and (v) can be considered as an extension of the Wulbert–Bartelt–McLaughlin characterization of strong unicity in a set-valued setting.

In infinite dimensional cases, the strong Kolmogorov criterion given in Theorem 14(ii) does not always guarantee strong uniqueness. For example, Bartelt and McLaughlin [4] constructed an infinite dimensional subspace G of $C(X, \mathbb{R})$ and a function f in $C(X, \mathbb{R})$ such that Theorem 14(ii) holds, $P_G(f)=\{0\}$, but $P_G(f)$ is not strongly unique.

A special case of Theorem 14 is when $P_G(f)$ is a singleton. Then the characterization of strong unicity of $P_G(f)$ can be simplified as shown in the following corollary.

COROLLARY 15. *Let $f \in C(I, \mathbb{R}^m) \setminus G$ and $x^* \in \mathbb{R}^n$. Then g_{x^*} is a strongly unique best approximation (of order $\alpha = 1$) to f in G if and only if*

$$\left\{ \sum_{i \in J(x^*)} \theta_i \nabla h_i(x^*) : \theta_i \geqslant 0 \right\} = \mathbb{R}^n. \tag{60}$$

Proof. If (60) holds, then $x^* \in S(f)$ (cf. Theorem 4). Thus, it follows from (41) and (60) that

$$N_{S(f)}(x^*) = \left\{ \sum_{i \in J(x^*)} \theta_i \nabla h_i(x^*) : \theta_i \geqslant 0 \right\} = \mathbb{R}^n.$$

This implies $S(f) = \{x^*\}$. Moreover, by Theorem 14 (cf. Definition 11), $P_G(f)$ is Hausdorff strongly unique. From Theorem 2 we know that $P_G(f) = \{g_{x^*}\}$ is a singleton. So $P_G(f)$ is strongly unique.

On the other hand, if $P_G(f) = \{g_{x^*}\}$ is strongly unique, then $S(f) = \{x^*\}$ and $N_{S(f)}(x^*) = \mathbb{R}^n$. By Theorem 14 and strong unicity of $P_G(f)$, we obtain

$$\left\{ \sum_{i \in J(x^*)} \theta_i \nabla h_i(x^*) : \theta_i \geqslant 0 \right\} = N_{S(f)}(x^*).$$

Thus, (60) holds. ∎

REFERENCES

1. J. Angelos and A. Egger, Strong uniqueness in L^P spaces, *J. Approx. Theory* **42** (1984), 14–26.
2. M. Bartelt and W. Li, Haar theory in vector-valued continuous function spaces, *in* "Approximation Theory, VIII. Vol. 1. Approximation and Interpolation" (C. K. Chui and L. L. Schumaker, Eds.), pp. 39–46, World Scientific, New York, 1995.
3. M. Bartelt and W. Li, Error estimates and Lipschitz constants for best approximation in continuous function spaces, *Comput. Math. Appl.* **30** (1995), 255–268.
4. M. Bartelt and H. McLaughlin, Characterizations of strong unicity in approximation theory, *J. Approx. Theory* **9** (1973), 255–266.
5. M. S. Bazaraa, H. D. Shevali, and C. M. Shetty, "Nonlinear Programming, Theory and Applications," 2nd ed., Wiley, New York, 1993.
6. B. Brosowski, "Parametric Semi-Infinite Optimization," Verlag Peter Lang, Frankfurt am Main, 1987.

7. B. L. Chalmers and G. D. Taylor, On the existence of strong unicity of arbitrarily small order, *in* "Approx. Theory, III" (E. Cheney, Ed.), pp. 293–298, Proceedings, Conf. Austin, Texas, 1980.
8. A. K. Cline, Lipschitz conditions on uniform approximation operators, *J. Approx. Theory* **8** (1973), 160–172.
9. F. Deutsch, Best approximation in the space of continuous vector-valued functions, *J . Approx. Theory* **53** (1988), 112–116.
10. M. S. Henry, D. Schmidt, and J. Swetits, Uniform strong unicity for rational approximation, *J. Approx. Theory* **33** (1981), 131–146.
11. J.-B. Hiriart-Urruty and C. Lemaréchal, "Convex Analysis and Minimization Algorithms, I," Springer-Verlag, New York, 1993.
12. A. J. Hoffman, Approximate solutions of systems of linear inequalities, *J. Res. Nat. Bur. Standards* **49** (1952), 263–265.
13. A. N. Kolmogorov, A remark on the polynomials of P. L. Chebyshev deviating the least from a given function, *Uspekhi Mat. Nauk* **3** (1948), 216–221. [In Russian]
14. W. Li, Various continuities of metric projections of $C_0(T, X)$, *J. Approx. Theory* **57** (1989), 150–168.
15. W. Li, Abadie's constraint qualification, metric regularity, and error bounds for differentiable convex inequalities, *SIAM J. Optim.*, in press.
16. W. Li, Strong uniqueness and Lipschitz continuity of metric projection: A generalization of the classical Haar theory, *J. Approx. Theory* **56** (1989), 164–184.
17. D. J. Newman and H. S. Shapiro, Some theorems on Chebyshev approximation, *Duke Math. J.* **30** (1963), 673–681.
18. G. Nürnberger, Unicity and strong unicity in approximation theory, *J. Approx. Theory* **26** (1979), 54–70.
19. S. H. Park, Uniform Hausdorff strong uniqueness, *J. Approx. Theory* **55** (1989), 78–89.
20. A. Pinkus, Uniqueness in vector-valued approximation, *J. Approx. Theory* **73** (1993), 17–92.
21. T. J. Rivlin and H. S. Shapiro, A unified approach to certain problems of polynomial approximation and minimization, *SIAM J. Appl. Math.* **9** (1961), 670–699.
22. R. T. Rockafellar, "Convex Analysis," Princeton Univ. Press, Princeton, NJ, 1970.
23. D. Schmidt, Strong unicity and Lipschitz conditions of order $\frac{1}{2}$ for monotone approximation, *J. Approx. Theory* **27** (1979), 346–354.
24. I. Singer, "Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces," Springer-Verlag, New York, 1970.
25. T. Wang and J.-S. Pang, Global error bounds for convex quadratic inequality systems, *Optimization* **31** (1994), 1–12.
26. D. Wulbert, Uniqueness and differential characterization of approximation from manifolds of functions, *Amer. J. Math.* **90** (1971), 350–366.
27. S. I. Zukhovitskii and S. B. Stechkin, On approximation of abstract functions with values in Banach space, *Dokl. Akad. Nauk. SSSR* **106** (1956), 773–776.