JOURNAL OF APPROXIMATION THEORY 75, 107–111 (1993)

Note

A. J. Hoffman's Theorem and Metric Projections in Polyhedral Spaces

WU LI

Department of Mathematics and Statistics, Old Dominion University, Norfolk, Virginia 23529, U.S.A.

Communicated by Aldric L. Brown

Received August 28, 1991; accepted in revised form June 2, 1992

A polyhedral space is a Banach space whose unit ball is the convex hull of a finite set [24]. A polyhedral space X can be considered as $\mathbb{R}^n$ endowed with a polyhedral norm $\|\cdot\|$; i.e, there exist $E_1, ..., E_m \in \mathbb{R}^n$ such that [24]

$$\|z\| = \max_{1 \leqslant i \leqslant m} E_i^T z \qquad \text{for} \quad z \in \mathbb{R}^n,$$

where the superscript "T" denotes the transpose of column vectors. For example, the l_1-norm $\|z\|_1 := \sum_{i=1}^n |z_i|$ and the l_∞-norm $\|z\|_\infty := \max_{1 \leqslant i \leqslant n} |z_i|$ are polyhedral norms. For two vectors x and y, $x \leqslant y$ means that $x_i \leqslant y_i$ for all indices i. We use x_+ to denote the vector whose ith component is $\max\{x_i, 0\}$.

Let B be an $r \times n$ matrix, $c \in \mathbb{R}^r$, and $K := \{x \in \mathbb{R}^n : Bx \leqslant c\}$ a convex polyhedral subset of $\mathbb{R}^n$. There is a best approximation problem associated with the set K. Namely, for any point z, we can ask for a best approximation to z in K. This leads us to define the distance from z to K by writing

$$d(z, K) := \min\{\|z - x\| : x \in K\} \qquad \text{for} \quad z \in \mathbb{R}^n \tag{2}$$

and to define the metric projection P_K from X to K by the equation

$$P_K(z) := \{x \in K : \|z - x\| = d(z, K)\}.$$

Our objective is to derive the Lipschitz continuity and the Hausdorff strong uniqueness of P_K. For this purpose, we use the following theorem of A. J. Hoffman [11] on approximate solutions of linear systems.

0021-9045/93 $5.00

HOFFMAN'S THEOREM. *Let A be $k \times n$ matrix. Then there exists a constant $\gamma > 0$, depending only on A, such that, if $b \in \mathbb{R}^k$ and $\{y \in \mathbb{R}^n : Ay \leqslant b\} \neq \emptyset$, then*

$$\min\{\|x-y\|_\infty : Ay \leqslant b\} \leqslant \gamma \cdot \|(Ax-b)_+\|_\infty \qquad \textit{for} \quad x \in \mathbb{R}^n. \tag{3}$$

Remark. There are many papers on estimation of γ in (3) for various norms, for example, [30, 20, 7, 23, 3, 17, 18]. Hoffman's theorem is essential in the convergence analysis of certain descent algorithms for linearly constrained optimization problems [12, 13, 19].

THEOREM 1. *The metric projection P_K is uniformly Hausdorff strongly unique and consequently Lipschitz continuous; i.e., there exists a positive constant β, depending only on B and X, such that*

$$\|z-x\| \geqslant d(z, K) + \beta \cdot d(x, P_K(z)) \qquad \textit{for} \quad x \in K, z \in X, \tag{4}$$

and consequently

$$H(P_K(w), P_K(z)) := \max\{\max_{y \in P_K(w)} d(y, P_K(z)), \max_{x \in P_K(z)} d(x, P_K(w))\}$$

$$\leqslant \frac{2}{\beta} \cdot \|w-z\| \qquad \textit{for} \quad w, z \in X. \tag{5}$$

Proof. By (1), we have

$$P_K(z) = \{y \in \mathbb{R}^n : By \leqslant c,\ E_i^T(z-y) \leqslant d(z, K), \text{ for } 1 \leqslant i \leqslant m\}.$$

Let $x \in K$. Then $Bx \leqslant c$ and $E_i^T(z-x) \leqslant \|z-x\|$. By applying Hoffman's Theorem to $P_K(z)$ which is a polyhedron, we obtain

$$\min\{\|x-y\|_\infty : y \in P_K(z)\} \leqslant \gamma \cdot \max_{1 \leqslant i \leqslant m} |(E_i^T(z-x) - d(z, K))_+|$$

$$\leqslant \gamma \cdot (\|z-x\| - d(z, K)),$$

where $\gamma > 0$ depends only on B and $E_1, ..., E_m$. Since any two norms on $\mathbb{R}^n$ are equivalent, there exist $\beta > 0$, depending only on B and X, such that

$$\|z-x\| \geqslant d(z, K) + \beta \cdot d(x, P_K(z)) \qquad \text{for} \quad x \in K, z \in \mathbb{R}^n.$$

Now we can use Cheney's argument (cf. page 82 of [5]) to prove (5). For $x \in P_K(w) \subset K$,

$$d(x, P_K(z)) \leqslant \frac{1}{\beta} \cdot (\|z-x\| - d(z, K)) \leqslant \frac{1}{\beta} \cdot (\|z-x\| - d(w, K) + \|z-w\|)$$

$$\leqslant \frac{1}{\beta} \cdot (2\|z-w\| + \|w-x\| - d(w, K)) = \frac{2}{\beta} \cdot \|z-w\|,$$

where the first inequality is (4) and the next two follow from the triangle inequality for distance. Similarly, for any $y \in P_K(z)$, $d(y, P_K(w)) \leqslant (2/\beta) \|z - w\|$. Thus, (5) holds. ∎

Remarks. Theorem 1 is a generalization of Theorem 2.1 in [9], where K was assumed to be a subspace. A subspace K of X is called a Chebyshev subspace if $P_K(x)$ is a singleton for every $x \in X$. When X is $\mathbb{R}^n$ with the l_∞-norm and K is a Chebyshev subspace of X, Theorem 1 reduces to a result of Cline [6] (a correct proof is given in [2]). In the case that X is $\mathbb{R}^n$ with the l_1-norm and K is a subspace of X, Theorem 1 extends a result by Angelos and Schmidt [1]. When $P_K(z)$ is a singleton, (4) is due to Jittorntrum and Osborn [14].

In order to recount the history and terminology related to (4), let us consider the general minimization problem

$$f_{\min} := \min\{f(x) : x \in K\}, \tag{6}$$

where f is a real-valued function on a subset K of a topological space X. Let

$$S := \{x \in K : f(x) = f_{\min}\} \neq \varnothing.$$

It sometimes happens that there exists a positive constant β such that

$$f(x) - f_{\min} \geqslant \beta \cdot d(x, S) \qquad \text{for} \quad x \in K. \tag{7}$$

For the moment, let us assume that X is the Banach space $C(T)$ of continuous real-valued functions on a compact Hausdorff space T, that K is a finite-dimensional subspace of X, and that $f(x) := \max_{t \in T} |g(t) - x(t)|$ for some $g \in C(T)$. When K is a Haar subspace, (7) was proved by Newman and Shapiro [25]. See also [4]. In this case, S is a singleton and is said to be strongly unique by Newman and Shapiro. Since then, there have been many papers devoted to the study of strong uniqueness of best approximations. The main topics include estimates of the minimum β in (7) (which is called the strong unicity constant), representations of the strong unicity constant, the dependency of the strong unicity constant on g, T, K, characterizations of when (7) holds, relationships between the strong unicity constant and the Lipschitz constant, etc. A complete list of references is too long to include here. When K is not a Haar subspace, S has in general more than one element. In this case, (7) is referred to as Hausdorff strong uniqueness of $P_K(g)$ [15]. (Note that $S = P_K(g)$ here.) This was used in studying characterizations of lower semicontinuity of $P_K(g)$ with respect to g. As seen from the proof of Theorem 1, the Hausdorff strong uniqueness provides a natural technique for establishing the Lipschitz continuity of metric projections in a normed linear space (cf. also [28, 9]).

Now let us assume that $X = \mathbb{R}^n$, that K is a polyhedral subset of $\mathbb{R}^n$, and that f is a linear functional. Then (6) leads to a linear programming problem. For this case, Mangasarian and Meyer [22] established (7). (As a matter of fact, (4) can also be derived from Mangasarian and Meyer's result, since (2) can be reformulated as a linear program [26, 16].) But they did not attach any terminology to (7). Meanwhile, in deriving (4) for a singleton $P_K(z)$, Jittorntrum and Osborne [14, 26] proved (7) for a singleton S. They defined S to be strongly unique if S is a singleton and (7) holds [14, 26].

The terminology "sharp minimum" was coined by Polyak [29]. The general minimization problem (6) is said to have a sharp minimum if S is a singleton and (7) holds [29]. In a recent paper on the finite termination of the proximal point algorithm, Ferris [10] used the same terminology even if S is not a singleton. However, Mangasarian [21] prefers to say that (6) has a weak sharp minimum if (7) holds. The concept of strong uniqueness or sharp minimum is also essential in convergence analysis of some algorithms for solving optimization problems [8, 22, 14, 26, 29, 27, 10].

Even though there is no standard terminology to describe (7), the importance of the strong uniqueness or the sharp minimum in applications is obvious.

ACKNOWLEDGMENT

The author thanks Professor E. W. Cheney for many helpful comments.

REFERENCES

1. J. R. ANGELOS AND D. SCHMIDT, Strong uniqueness in $L^1(X, \Sigma, \mu)$, *in* "Aproximation Theory IV" (C. Chui, L. Schumaker and J. Ward, Eds.), pp. 297–302, Academic Press, New York, 1983.
2. M. W. BARTELT, On Lipschitz conditions, strong unicity and a theorem of A. K. Cline, *J. Approx. Theory* **14** (1975), 245–250.
3. C. BERGTHALLER AND I. SINGER, The distance to a polyhedron, *Linear Algebra Appl.* **169** (1992), 111–129.
4. E. W. CHENEY, "Some Relationships between the Tchebycheff Approximations on an Interval and on a Discrete Subset of That Interval," Mathematical Notes No. 262, Boeing Scientific Research Laboratories, Seattle, Washington, 1962.
5. E. W. CHENEY, "Introduction to Approximation Theory," 2nd ed., Chelsea, New York, 1982.
6. A. K. CLINE, Lipschitz conditions on uniform approximation operators, *J. Approx. Theory* **8** (1973), 160–172.
7. W. COOK, A. M. H. GERARDS, A. SCHRIJVER, AND É. TARDOS, Sensitivity theorems in integer linear programming, *Math. Programming* **34** (1986), 251–264.
8. L. CROMME, Strong uniqueness: A far reaching criterion for convergence analysis of iterative processes, *Numer. Math.* **29** (1978), 179–193.

9. F. DEUTSCH AND W. LI, Strong uniqueness, Lipschitz continuity, an continuous selections of metric projections in $L_1(T, \mu)$, *J. Approx. Theory* **66** (1991), 198–224.
10. M. C. FERRIS, Finite termination of the proximal point algorithm, *Math. Programming* **50** (1991), 359–366.
11. A. J. HOFFMAN, Approximate solutions of systems of linear inequalities, *J. Res. Nat. Bur. Standards* **49** (1952), 263–265.
12. J. L. GOFFIN, The relaxation method for solving systems of linear inequalities, *Math. Oper. Res.* **5** (1980), 388–414.
13. N. IUSEM AND A. R. DE PIERRO, On the convergence properties of Hildreth's quadratic programming algorithm, *Math. Programming* **47** (1990), 37–51.
14. K. JITTORNTRUM AND M. R. OSBORNE, Strong uniqueness and second order convergence in nonlinear discrete approximation, *Numer. Math.* **34** (1980), 439–455.
15. W. LI, Strong uniqueness and Lipschitz continuity of metric projections: A generalization of the classical Haar theory, *J. Approx. Theory* **56** (1989), 164–184.
16. W. LI, Best approximations in polyhedral spaces and linear programs, *in* "Approximation Theory: Proceedings of the Sixth Southeastern International Conference" (G. Anastassiou, Ed.), pp. 393–400, Dekker, New York, 1993.
17. W. LI, Sharp Lipschitz constants for basic optimal solutions and basic feasible solutions of linear programs, SIAM *J. Control Optim.* **32** (1994), to appear.
18. W. LI, The sharp Lipschitz constants for feasible and optimal solutions of a perturbed linear program, *Linear Algebra Appl.* **187** (1993), 15–40.
19. ZHI-QUAN LUO AND P. TSENG, On the linear convergence of descent methods for convex essentially smooth optimization, *SIAM J. Control Optim.* **30** (1992), 408–425.
20. O. L. MANGASARIAN, A condition number of linear inequalities and equalities, Methods of Operations Research 43, *in* "Proceedings of 6. Symposium über Operations Research, Universität Augsburg, September 7–9, 1981" (G. Bamberg and O. Opitz, Eds.), pp. 3–15, Verlagsgruppe Athennäum/Hain/Scriptor/Hanstein, Konigstein, 1981.
21. O. L. MANGASARIAN, Error bounds for nondegenerate monotone linear complementarity problems, *Math. Programming* **48** (1990), 437–445.
22. O. L. MANGASARIAN AND R. R. MEYER, Nonlinear perturbation of linear programs, *SIAM J. Control Optim* **17** (1979), 745–752.
23. O. L. MANGASARIAN AND T.-H. SHIAU, Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems, *SIAM J. Control Optim.* **25** (1987), 583–595.
24. P. H. MASERICK, Dual convex polytopes in Banach spaces, *J. Math. Anal. Appl.* **19** (1967), 263–273.
25. D. J. NEWMAN AND H. S. SHAPIRO, Some theorems on Chebyshev approximation, *Duke Math. J.* **30** (1963), 673–684.
26. M. R. OSBORNE, "Finite Algorithms in Optimization and Data Analysis," Wiley Series in Probability and Mathematical Statistics, Wiley, New York, 1985.
27. M. R. OSBORNE AND R. S. WOMERSLEY, Strong uniqueness in sequential linear programming, *J. Austral. Math. Soc. Ser. B* **31** (1990), 379–384.
28. SUN-HO PARK, Uniform Hausdorff strong uniqueness, *J. Approx. Theory* **58** (1989), 78–89.
29. B. T. POLYAK, "Introduction to Optimization," Optimization Software, Inc., New York, 1987.
30. S. M. ROBINSON, Bounds for error in the solution set of a perturbed linear program, *Linear Algebra Appl.* **6** (1973), 69–81.