JOURNAL OF APPROXIMATION THEORY **12**, 234–239 (1974)

The Continuity of Metric Projections as Functions of the Data

JAMES W. DANIEL*

Departments of Mathematics and Computer Sciences, University of Texas, Austin, Texas 78712

Communicated by E. W. Cheney

Let X be a Hilbert space, and consider the point x_0 minimizing, for a given f in X, the distance $\| x - f \|$ as x ranges over a polyhedral set C defined by a finite number of real-valued equalities and inequalities. We wish to see how x_0 varies when f and C vary. It is easy to see that x_0 is Hölder continuous with exponent $\frac{1}{2}$ in its dependence on f and C; this estimate is in general sharp. We show, however, that in certain cases x_0 is actually Lipschitz continuous in its dependence on the parameters which are used to define the set C.

1. INTRODUCTION

A problem of some practical and theoretical interest is that of constrained approximation: given a point f and a subset C_0 of some normed space X, find an x_0 in C_0 minimizing $\| x - f \|$ as x ranges over C_0. When C_0 is a linear subspace or a flat, this problem is fairly well understood; we wish to consider the more general situation in which C_0 is just convex. For example, if x is a continuous function, we wish to allow constraints such as $x \geqq 0$ or $\alpha_i \leqq x(t_i) \leqq \beta_i$. In practical situations the parameters appearing in the constraints—such as α_i, β_i, and t_i above—may well contain experimental or measurement errors; it is meaningful to ask in what way these errors influence the solution x_0. Thus, in this paper we wish to investigate how x_0, the metric projection of f onto the set C_0, depends on the set C_0 itself. Apparently this question of stability or conditioning has not been considered in much detail although some related results are known [1, 2, 4, 6–8, 11, 12]; generally one more often studies the case in which C_0 is fixed but f varies. In our studies we can of course also allow f to vary since, by the simple device of replacing C_0 by $C_0 - f$, we can always assume f to be zero and we can replace perturbations in f with perturbations in C_0. With this in mind, *we hereafter assume for convenience that f is zero.*

* This research was supported in part by the Office of Naval Research under Contract N00014-67-A-0126-0015, NRO44-425; reproduction in whole or in part is permitted for any purpose of the United States government.

First we consider what sort of results one might hope to obtain. The best one might expect in any generality is that a "nice" perturbation of order ϵ in a "nice" set C_0 would lead to a perturbation of order ϵ in the solution x_0. The following extremely simple example in $\mathbf{R}^2$ seems to indicate however that no such result is possible.

Let $C_0 = \{x = (x_1, x_2) \in \mathbf{R}^2 \mid x_1 \geqq 0, x_2 \geqq 1\}$; clearly the closest point to (0, 0) in C_0 is (0, 1), where we use the l^2-norm $\| x \| = (x_1^2 + x_2^2)^{1/2}$. For any $\epsilon \geqq 0$, let $C_\epsilon = \{x \mid x_1 \geqq 0, x_2 \geqq 1, x_2 + x_1\sqrt{\epsilon} \geqq 1 + \epsilon\}$. Since $C_\epsilon \subset C_0$, clearly any x in C_ϵ is at a zero distance from C_0; conversely, any $x = (x_1, x_2)$ in C_0 is at most ϵ distant from $(x_1, \max\{x_2, 1 + \epsilon - x_1\sqrt{\epsilon}\})$ in C_ϵ. Thus in almost any sense C_ϵ is a perturbation of C_0 of order ϵ; since the closest point in C_ϵ to (0, 0) is $(\sqrt{\epsilon}, 1)$, we see that the closest point has moved by $\sqrt{\epsilon}$ although the set C_0 moved only by ϵ. Since this was a finite dimensional problem with an inner product norm and since C_0 and C_ϵ were simple polygons, it seems clear that we cannot hope to prove that the closest point is Lipschitz continuous in its dependence on the distance between C_ϵ and C_0; at most we might hope for Hölder continuity with exponent $\frac{1}{2}$. We return to this in Section 2.

It is important to note in this example however that, although C_ϵ is only ϵ away from C_0, in some sense this is a fluke since the inequalities defining C_ϵ are actually perturbations of order $\sqrt{\epsilon}$ from those inequalities defining C_0. Thus one might now hope to show that perturbations of order ϵ in the inequalities and equalities defining C_0 would lead to perturbations of order ϵ in the closest point to the origin. In [3] we extended in certain directions the work of [5, 10, 11] concerning perturbations of linear inequalities and equalities in finite dimensions, while in [4] we applied these results to study the perturbation in the solution of finite dimensional definite quadratic programs when the data are perturbed; here we apply and extend these latter results to the problem addressed in this paper, assuming X to be a real Hilbert space.

2. Hölder Continuity in the Dependence on Distance

First we pause to dispose of the simpler question of the dependence of the solution on the distance between the constraint sets. Throughout the remainder of this paper *we assume that X is a real Hilbert space with inner product* $\langle \cdot, \cdot \rangle$. We also assume throughout that C_ϵ for $\epsilon \geqq 0$ is a closed convex subset of X; actually C_ϵ need not depend explicitly on ϵ, a parameter used only to measure the "distance" between C_0 and perturbations thereof. We define, for $\epsilon \geqq 0$,

$$x_\epsilon \text{ minimizes } \| x \|^2 = \langle x, x \rangle \text{ as } x \text{ ranges over } C_\epsilon. \tag{2.1}$$

THEOREM 2.2. *In addition to the general conditions above, suppose that for each* $r > 0$ *there exists a constant* c_r *such that to each* x *in* C_0 *with* $\| x \| \leqq r$ *there corresponds an* x_ϵ' *in* C_ϵ *with* $\| x_\epsilon' - x \| \leqq c_r \epsilon$ *and to each* y_ϵ *in* C_ϵ *with* $\| y_\epsilon \| \leqq r$ *there corresponds a* y' *in* C_0 *with* $\| y' - y_\epsilon \| \leqq c_r \epsilon$. *Then there exists a constant* c *such that the* x_ϵ *defined by Eq.* (2.1) *satisfy* $\| x_\epsilon - x_0 \| \leqq c \sqrt{\epsilon}$ *as* ϵ *tends to zero.*

Proof. For x_ϵ defined by Eq. (2.1), let x_ϵ' in C_ϵ satisfy $\| x_\epsilon' - x_0 \| \leqq c_{r_0} \epsilon$ for $\epsilon > 0$, where $r_0 = \| x_0 \|$. Since $\| x_\epsilon \| \leqq \| x_\epsilon' \| = \| x_0 + x_\epsilon' - x_0 \| \leqq \| x_0 \| + c_{r_0} \epsilon \equiv r$, we know that there exist y_ϵ' in C_0 satisfying $\| x_\epsilon - y_\epsilon' \| \leqq c_r \epsilon$. Now, by the usual characterization of the point of minimum norm in a convex set, we have $\langle x_\epsilon , x_\epsilon' - x_\epsilon \rangle \geqq 0$ and $\langle x_0 , y_\epsilon' - x_0 \rangle \geqq 0$. Adding these two inequalities and rearranging terms yields

$$0 \leqq \langle x_\epsilon - x_0 , x_0 - x_\epsilon \rangle + \langle x_\epsilon - x_0 , x_\epsilon - y_\epsilon' \rangle + \langle x_\epsilon , x_\epsilon' - x_0 + y_\epsilon' - x_\epsilon \rangle .$$

Therefore $\| x_\epsilon - x_0 \|^2 \leqq \| x_\epsilon - x_0 \| c_r \epsilon + r(2 c_r \epsilon)$; solving this for $\| x_\epsilon - x_0 \|$ then yields $\| x_\epsilon - x_0 \| \leqq \frac{1}{2}\{c_r \epsilon + [(c_r \epsilon)^2 + 8 r c_r \epsilon]^{1/2}\}$ from which the conclusion of the theorem follows. Q.E.D.

This theorem shows that x_ϵ is Hölder continuous at x_0 with exponent at least $\frac{1}{2}$ in its dependence on the "distance" between C_0 and C_ϵ; the example of Section 1 shows this estimate to be sharp. We remark that similar estimates can be obtained for more general uniformly convex norms.

3. LIPSCHITZ CONTINUITY IN THE DEFINING PARAMETERS

As suggested by our example in Section 1, we hope to show that x_0 moves by order ϵ when the inequalities defining C_0 are perturbed by order ϵ. Any convex set C can be defined via $C = \{x \mid \langle l(t), x \rangle \leqq a(t)$ for all t in $T\}$, where T is some index set, $a(t)$ is a scalar, and $l(t)$ is an element of X for each t in T. We wish to consider sets $C_\epsilon = \{x \mid \langle l_\epsilon(t), x \rangle \leqq a_\epsilon(t)$ for all t in $T\}$ where $\| l_\epsilon(t) - l(t) \| \leqq \epsilon$ and $| a_\epsilon(t) - a(t) | \leqq \epsilon$ for all t in T; we wish to identify conditions under which one can prove that $\| x_0 - x_\epsilon \| \leqq c\epsilon$ for some constant c and x_ϵ defined by Eq. (2.1). Thus far in our work we have only succeeded in treating the case in which T is a finite set; we proceed to the study of this case.

We assume now that C_0 has the form

$$C_0 = \{x \mid G_0 x \leqq g_0 , D_0 x = d_0\}, \tag{3.1}$$

where G_0 and D_0 are linear operators, $G_0 \colon X \to \mathbf{R}^{m'}$, $D_0 \colon X \to \mathbf{R}^{r'}$, g_0 is in $\mathbf{R}^{m'}$ and d_0 is in $\mathbf{R}^{r'}$. We use the symbols $\leqq$, $\leq$, $<$ and their analogs $\geqq$, $\geq$, $>$

in the sense of [9]. That is, $b \leqq 0$ if and only if each component of b is nonpositive, $b < 0$ if and only if each component of b is negative, and $b \leq 0$ if and only if $b \leqq 0$ but $b \neq 0$. For convenience, we no longer keep track of the precise dimension of the ranges of various linear operators with finite dimensional ranges; the dimensions are always such that all indicated compositions, order relations, et cetera are well defined. We also use the same symbol $\| \cdot \|$ to denote a variety of norms. For convenience we often will use the symbol c as a generic constant, seldom the same in different occurrences. We denote by $G_0^{(i)}$ (and similarly for other operators with finite dimensional range) the ith component linear functional in G_0, that is, $G_0 x$ is the vector whose ith component is $\langle G_0^{(i)}, x \rangle$. Let X_0 be the subspace of X spanned by the set $\{G_0^{(i)}, D_0^{(j)}$ for all $i, j\}$. It is a simple consequence [7, 9] of duality theory (or of orthogonally decomposing x into components with respect to X_0) that the solution x_0 to Eq. (2.1), that is minimizing $\| x \|^2$ over C_0, must in fact lie in X_0. Thus we get the same solution by considering the strictly finite dimensional problem of minimizing $\| x \|^2$ over

$$C_0' \equiv \{x \mid G_0 x \leqq g_0, D_0 x = d_0, x \text{ in } X_0\}. \tag{3.2}$$

In the setting of Eq. (3.2) we may loosely think of G_0 and D_0 as (being given by) rectangular matrices.

We now consider perturbations of C_0, namely

$$C_\epsilon = \{x \mid G_\epsilon x \leqq g_\epsilon, D_\epsilon x = d_\epsilon\}, \quad \text{with} \quad \| G_0 - G_\epsilon \| \leqq \epsilon, \quad \| g_\epsilon - g_0 \| \leqq \epsilon,$$
$$\| D_\epsilon - D_0 \| \leqq \epsilon, \quad \| d_\epsilon - d_0 \| \leqq \epsilon, \tag{3.3}$$

and with G_ϵ and D_ϵ mapping into the same range spaces as G_0 and D_0, respectively. Once again the point x_ϵ of minimum norm in C_ϵ must lie in X_ϵ, the span of $\{G_\epsilon^{(i)}, D_\epsilon^{(j)}$ for all $i, j\}$, so that we may restrict ourselves to minimizing $\| x \|^2$ over

$$C_\epsilon' = \{x \mid G_\epsilon x \leqq g_\epsilon, D_\epsilon x = d_\epsilon, x \text{ in } X_\epsilon\}. \tag{3.4}$$

By means of Eqs. (3.2) and (3.4) we have reduced our problem to one of considering the effect of perturbations in the equalities and inequalities defining finite dimensional polyhedra; this is precisely the problem discussed at length in [3]. The arguments there show that we can partition G_0 and g_0 into

$$G_0 = \begin{pmatrix} A_0 \\ B_0 \end{pmatrix}, \qquad g_0 = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} \tag{3.5}$$

such that

$$\begin{cases} B_0 x = b_0 \text{ for all } x \text{ in } C_0', \text{ and either } A_0 \text{ is vacuous or} \\ \text{there is } \hat{x} \text{ in } C_0' \text{ with } A_0 \hat{x} \leqq a_0 - h \text{ for some } h > 0. \end{cases} \tag{3.6}$$

It also is shown in [3] that in order for C_0' and C_ϵ' to be close together as ϵ tends to zero one must require that the rank of the matrix representations of $N_\epsilon \equiv \binom{B_\epsilon}{D_\epsilon}$ on X_ϵ (where G_ϵ is partitioned as $\binom{A_\epsilon}{B_\epsilon}$ in parallel with that for G_0) must be constant for $\epsilon \geqq 0$. It is straightforward to see that this hypothesis on the ranks of the matrix representations of $\binom{B_\epsilon}{D_\epsilon}$ on X_ϵ is equivalent to the same hypothesis on the dimension of the ranges of $\binom{B_\epsilon}{D_\epsilon}$ on X since this dimension is equal to the aforementioned rank. Therefore, from Theorem 4.2 of [3] and from Hoffman's Theorem [5, 3] we immediately deduce the following:

PROPOSITION 3.7. *In addition to the general hypotheses in Eqs.* (3.1)–(3.6), *suppose that*

$$\left\{\begin{array}{l}\textit{the dimension of the range of } N_\epsilon \equiv \begin{pmatrix} B_\epsilon \\ D_\epsilon \end{pmatrix} \textit{ on } X \textit{ equals the} \\ \textit{dimension of the range } N_0 \equiv \begin{pmatrix} B_0 \\ D_0 \end{pmatrix} \textit{ on } X \textit{ for all } \epsilon \geqq 0.\end{array}\right. \tag{3.8}$$

Then there exist positive constants c and ϵ_0 depending on C_0' such that:

(1) *to each x_0' in C_0' satisfying $\epsilon(1 + \| x_0' \|) \leqq \epsilon_0$ there corresponds an x_ϵ' in C_ϵ' satisfying $\| x_0' - x_\epsilon' \| \leqq c\epsilon(1 + \| x_0' \|)$; and*

(2) *to each x_ϵ' in C_ϵ' there corresponds an x_0' in C_0' satisfying $\| x_0' - x_\epsilon' \| \leqq c\epsilon(1 + \| x_\epsilon' \|)$.*

Combining the above proposition with Theorem 2.2 and recalling that our minimization problems over C_0 and C_ϵ are equivalent to those over C_0' and C_ϵ', we find the following simple result.

COROLLARY 3.9. *In addition to the general hypotheses in Eqs.* (2.1) *and* (3.1)–(3.6), *assume that Eq.* (3.8) *holds. Then there exist positive constants ϵ_0 and c such that, for $\epsilon \leqq \epsilon_0$, one has $\| x_0 - x_\epsilon \| \leqq c\sqrt{\epsilon}$.*

We remind the reader at this point that the above corollary is not what we set out to prove; rather we had hoped to use the results of [4] on the stability of quadratic programs to prove the stronger result $\| x_\epsilon - x_0 \| \leqq c\epsilon$. In fact we have now reached precisely the point following Proposition 4.6 of [4] at which the result $\| x_\epsilon - x_0 \| \leqq c\sqrt{\epsilon}$ for quadratic programs was strengthened to $\| x_\epsilon - x_0 \| \leqq c\epsilon$. Were it not for the fact that the resulting strong result, Theorem 4.24 of [4], was stated for the case in which X is finite dimensional, we could apply that theorem here without further ado. Fortunately, the analysis leading to that theorem can be repeated almost identically with only minor notational differences such as replacing a matrix transpose like A_0^T by the adjoint operator A_0^*; we do not waste the space here to reproduce those arguments for our more general setting. By thus extending Theorem 4.24 of [4], we immediately obtain our desired result.

THEOREM 3.10. *Let our general hypotheses of Eqs.* (2.1) *and* (3.1)–(3.6) *hold, so that* x_ϵ *minimizes* $\| x \|^2$ *over* $C_\epsilon = \{x \mid G_\epsilon x \leqq g_\epsilon , D_\epsilon x = d_\epsilon\}$, *where* G_ϵ *and* D_ϵ *have finite dimensional ranges and where* G_ϵ *is partitioned into* $G_\epsilon = \binom{A_\epsilon}{B_\epsilon}$ *in parallel with the partition of* G_0 *into* $G_0 = \binom{A_0}{B_0}$ *according to Eq.* 3.6. *Suppose, moreover, that Eq.* (3.8) *holds so that the dimension of the range of* $\binom{B_\epsilon}{D_\epsilon}$ *over X is constant for* $\epsilon \geqq 0$. *Then there exist positive constants c and* ϵ_0 *such that* $\| x_\epsilon - x_0 \| \leqq c\epsilon$ *whenever* $\epsilon \leqq \epsilon_0$ *and* $\epsilon = \max\{\| G_0 - G_\epsilon \|, \| g_0 - g_\epsilon \|, \| D_0 - D_\epsilon \|, \| d_0 - d_\epsilon \|\}$.

This result thus shows that, for polyhedral sets in infinite dimensional Hilbert spaces, metric projections are Lipschitz continuous in their dependence on the equalities and inequalities defining their range sets, under fairly minimal hypotheses. Thus perturbations or inaccuracies in the data defining such approximation problems lead to perturbations or inaccuracies in the solution that are of the same order of magnitude.

REFERENCES

1. PHILIP M. ANSELONE AND JOEL DAVIS, Perturbations of best approximation problems, to appear.
2. JAMES W. DANIEL, "The Approximate Minimization of Functionals," Prentice–Hall, Englewood Cliffs, NJ, 1971.
3. JAMES W. DANIEL, On perturbations in systems of linear inequalities, Center for Numerical Analysis Technical Report CNA-50, The University of Texas, Austin Texas, (1972); also *SIAM J. Numer. Anal.* **10** (1973), 299–307.
4. JAMES W. DANIEL, Stability of the solution of definite quadratic programs, Center for Numerical Analysis Technical Report CNA-58, The University of Texas, Austin, Texas (1972); also *Math. Progr.* **5** (1973), 41–53.
5. A. J. HOFFMAN, On approximate solutions of systems of linear inequalities, *J. Res. Nat. Bur. Standards, Sect. B* **49** (1952), 263–265.
6. JEAN L. JOLY, Une famille de topologies et de convergence sur l'ensemble des fonctionnelles convexes, thesis, Faculté de Sciences de Grenoble, France, 1970.
7. PIERRE-JEAN LAURENT, "Approximation et Optimisation," Hermann, Paris, 1972.
8. PIERRE-JEAN LAURENT AND PHAM DINH TUAN, Global approximation of a compact set by elements of a convex set in a normed space, *Numer. Math.* **15** (1970), 137–150.
9. OLVI L. MANGASARIAN, "Nonlinear Programming," McGraw–Hill, New York, 1969.
10. STEPHEN M. ROBINSON, Normed convex processes, *Trans. Amer. Math. Soc.* **174** (1972), 127–140.
11. STEPHEN M. ROBINSON, Bounds for error in the solution set of a perturbed linear program, *J. Lin. Alg. Appl.* **6** (1973), 69–82.
12. KARL H. USOW, On L^1 approximation I: Computation for continuous dependence, *SIAM J. Numer. Anal.* **4** (1967), 70–88.