JOURNAL OF APPROXIMATION THEORY 26, 204–211 (1979)

The Metric Projection on C^2 Manifolds in Banach Spaces

THEAGENIS ABATZOGLOU

Department of Mathematics, Iowa State University, Ames, Iowa 50011

Communicated by Oved Shisha

Received June 17, 1977

We study the best approximation of a point x in a Banach space B from a C^2 manifold. We derive a formula for the radius of curvature at points on the manifold and we use it to obtain a formula for the Frechet derivative of the metric projection onto the manifold.

1. INTRODUCTION

We consider a C^2 manifold M of dimension $k < \infty$ in a Banach space B. We study the metric projection P_M of B onto M, defined by $P_M(x) = \{ y \mid y \in M, \inf_{m \in M} \| x - m \| = \| x - y \| \}$. In general, $P_M(x)$ need not be a singleton; e.g. it may be empty or consist of more than one point. To avoid such cases we restrict our attention to the set $(\bar{D})^C$, the complement of the closure of D, where $D = \{x \mid x \in B$ and $P_M(x)$ is not a singleton$\}$.

The author has studied P_M for X a Euclidean space and M a closed C^2 manifold, see [1]. R. B. Holmes has proved in [6] that P_M is a C^p map when M is a closed convex set whose boundary is of class C^{p+1} and X is a Hilbert space. In the case that X is finite dimensional, J. R. Rice has established conditions which guarantee that P_M is a singleton at x (and hence continuous there); see [7]. Other authors, [3], [4], [8], use the radius of curvature of the manifold to determine whether P_M is a singleton at x.

Our goal in this paper is to show that P_M is C^1 on $(\bar{D})^C$ and to find the dependence of $P'_M(x)$ on the curvature of M at $P_M(x)$.

2. DEFINITIONS

We assume that the norm of the Banach space is C^2 and write $\nabla \| x \|$, $\nabla^2 \| x \|$ for the first and second Frechet derivatives of the norm at x. $\nabla^2 \| x \|$ is a bilinear function on $B \times B$ and by $\nabla^2 \| x \| (y)^{(2)}$ we shall mean $\nabla^2 \| x \| (y, y)$.

0021-9045/79/070204–08$02.00/0

Copyright © 1979 by Academic Press, Inc.
All rights of reproduction in any form reserved.

With regards to the k-dimensional C^2 manifold M we make the following assumptions:

(1) M is approximatively compact. (This guarantees that P_M is continuous on $(\bar{D})^C$).

(2) Let M be represented locally by a C^2 function f. Then $f'(b)(R^k)$ is a k-dimensional subspace of B for all b in the domain of f, prime denoting Frechet differentiation.

For m in M we can write $m = f(b)$ and then we define the tangent plane of M at m to be $T_m = m + f'(b)(R^k)$.

If x, y are in B, we say that x is orthogonal to y if $(d/dt) \| x + ty \| |_{t=0} = 0$, and "$x$ is orthogonal to a subspace Y of B" means that x is orthogonal to all y in Y. We also define y to be orthogonal to M at m if y is orthogonal to T_m.

3. Radius of Curvature

Consider a unit vector v orthogonal to M at m. Assume that for every μ in M sufficiently close to m, there is a $t > 0$ with

$$\| tv \| = \| m + tv - \mu \|,$$

i.e.,

$$t = \|(m - \mu) + tv \|. \tag{3.1}$$

We define the radius of curvature of M at m in the direction v to be

$$\rho(m, v) = \lim_{\epsilon \downarrow 0} t_\epsilon \text{ where } t_\epsilon = \inf_{0 < \|\mu - m\| < \epsilon} \{t \mid t = \|(m - \mu) + tv \|\}.$$

If the above assumption does not hold, set $\rho(m, v) = \infty$. Let f be a local representation of M around m with $f(a) = m$.

To obtain an explicit formula for $\rho(m, v)$ we further assume that

$$\nabla^2 \| v \| (f'(a)(b))^{(2)} > 0 \qquad \text{for all} \quad b \neq 0 \text{ in } R^k.$$

Also to avoid trivalities we shall assume that there are points μ arbitrarily closed to m such that (3.1) holds. We proceed by using Taylor's expansion in (3.1). We have

$$\begin{aligned} \|(m - \mu) + tv \| = \| tv \| &+ \nabla \| tv \| (m - \mu) \\ &+ \tfrac{1}{2}\nabla^2 \| tv \| (\mu - m)^{(2)} + o(\| \mu - m \|^2). \end{aligned}$$

Since f is a relative homeomorphism around m, if μ is close to m, we can write $\mu = f(c)$ and

$$\mu = f(c) = f(a) + f'(a)(c - a) + \tfrac{1}{2}f''(a)(c - a)^{(2)} + o(\| c - a \|^2).$$

Hence

$$\begin{aligned}
&\|(m-\mu)+tv\| \\
&\quad = t + \nabla\|tv\|\Big(-f'(a)(c-a) - \tfrac{1}{2}f''(a)(c-a)^{(2)} + o(\|c-a\|^2)\Big) \\
&\qquad + \frac{1}{2}\nabla^2\|tv\|\Big(-f'(a)(c-a) - \tfrac{1}{2}f''(a)(c-a)^{(2)} + o(\|c-a\|^2)\Big)^{(2)} \\
&\qquad + o(\|c-a\|^2) \\
&\quad = t - \frac{1}{2}\nabla\|v\|(f''(a)(c-a)^{(2)}) + \frac{1}{2t}\nabla^2\|v\|(f'(a)(c-a))^{(2)} \\
&\qquad + o(\|c-a\|^2).
\end{aligned}$$

We used the fact that tv is orthogonal to M at $m = f(a)$ to conclude that $\nabla\|tv\|(f'(a)(c-a)) = 0$. We also made use of the homogeneity of the norm to replace $\nabla^2\|tv\|$ by $(1/t)\nabla^2\|v\|$.

Now we can express (3.1) as

$$\begin{aligned}
t = t &- \frac{1}{2}\nabla\|v\|(f''(a)(c-a)^{(2)}) + \frac{1}{2t}\nabla^2\|v\|(f'(a)(c-a))^{(2)} \\
&+ o(\|c-a\|^2).
\end{aligned}$$

Next we solve for t and obtain

$$t = \frac{\nabla^2\|v\|(f'(a)(c-a))^{(2)}}{\nabla\|v\|(f''(a)(c-a)^{(2)}) + o(\|c-a\|^2)}$$

Set $w = (c-a)/\|c-a\|$ and let $c \to a$; we obtain

$$\rho(m,v) = \min_{\|w\|=1}\left\{\frac{\nabla^2\|v\|(f'(a)(w))^{(2)}}{\nabla\|v\|(f''(a)(w)^{(2)})}\;\middle|\;\nabla\|v\|(f''(a)(w)^{(2)}) \geqslant 0\right\}$$

Remark. Since we are assuming that $\nabla^2\|v\|(f'(a)(w))^{(2)} > 0$ for all w of norm 1, if $\nabla\|v\|(f''(a)(w)^{(2)}) \leqslant 0$ for all such w, then $\rho(m,v) = \infty$.

EXAMPLE. Let γ map R into $L^p[0,1]$, $p \geqslant 2$, so that $\gamma(t)$ is a function in $L^p[0,1]$ for each t. Assume $\gamma'(t) \neq 0$ and that $\gamma''(t)$ exists. For a given t we can find a function g in $L^p[0,1]$ such that $\|g\|_p = 1$ and g is orthogonal to γ at $\gamma(t)$; hence $\int_0^1 |g(x)|^{p-2} g(x)\,\gamma'(t)(x)\,dx = 0$. If

$$\nabla^2\|g\|_p(\gamma'(t))^{(2)} = (p-1)\int_0^1 |g(x)|^{p-2}\,|\gamma'(t)(x)|^2\,dx \neq 0$$

then

$$\rho(\gamma(t), g) = \frac{(p-1)\int_0^1 |g(x)|^{p-2} |\gamma'(t)|^2 \, dx}{\int_0^1 |g(x)|^{p-2} g(x) \gamma''(t)(x) \, dx}.$$

Observe that when $\int_0^1 |g(x)|^{p-2} g(x) \gamma''(t)(x) \, dx \leqslant 0$, then $\rho(\gamma(t), g) = \infty$ by the previous remark.

LEMMA 3.1. *Let A, B be real, symmetric $k \times k$ matrices such that $\langle Bw, w\rangle = \nabla^2 \| v \| (f'(a)(w))^{(2)}$ and $\langle Aw, w\rangle = \nabla \| v \| (f''(a)(w)^{(2)})$ for all w in R^k. Then*

(a) *$B = f'(a)^T \nabla^2 \| v \| f'(a)$, where $f'(a)^T$ is the adjoint of $f'(a)$.*

(b) *$A = (a_{ij})_{i,j}$ where $a_{ij} = \nabla \| v \| (\partial^2 f / \partial t_i \, \partial t_j)$.*

Proof. (a) We think of $\nabla^2 \| v \|$ as a bilinear map from $R^k \times R^k$ into R. We can write $\langle Bw, w\rangle = \nabla^2 \| v \| (f'(a)(w))^{(2)} = \langle \nabla^2 \| v \| (f'(a)(w)), f'(a)(w)\rangle = \langle f'(a)^T \nabla^2 \| v \| f'(a)(w), w\rangle$. Since B is symmetric, we get $B = f'(a)^T \nabla^2 \| v \| f'(a)$.

(b) $f''(a)(w)^{(2)} = \sum_{i,j} [\partial^2 f(a)/\partial t_i \, \partial t_j] w_i w_j$ where we set $a = (t_1, \dots, t_k)$, $w = (w_1, \dots, w_k)$. Then

$$\langle Aw, w\rangle = \nabla \| v \| (f''(a)(w)^{(2)}) = \nabla \| v \| \left(\sum_{i,j} \frac{\partial^2 f(a)}{\partial t_i \, \partial t_j} w_i w_j \right)$$
$$= \sum_{i,j} \nabla \| v \| \left(\frac{\partial^2 f(a)}{\partial t_i \, \partial t_j} \right) w_i w_j .$$

Since we assume that A is symmetric, we have $A = (a_{ij})_{i,j} = (\nabla \| v \| \times (\partial^2 f(a)/\partial t_i \, \partial t_j))_{i,j}$.

Observe than we can define the radius of curvature (when positive and finite) by $\rho(m, v) = \min_{\|w\|=1} \{\langle Bw, w\rangle / \langle Aw, w\rangle \mid \langle Aw, w\rangle > 0\}$.

4. THE METRIC PROJECTION P_M

Recall that D is the set of x for which $P_M(x)$ is not a singleton. We consider P_M on the set $U = (\bar{D})^C$. Suppose that $f'(a)^T \nabla^2 \| v \| f'(a)$ is positive definite for v in X orthogonal to M at $m = f(a)$. Then by using Theorem 5.1 of [3] one can show that U is not empty.

We now state our theorem concerning differentiability of P_M in U.

THEOREM 4.1. *Let X be a Banach space whose norm is C^2 and let M be an approximatively compact C^2 manifold in X of dimension $k < \infty$. Let x ($\notin M$) in U be such that $f'(a)^T \nabla^2 \| v \| f'(a)$ is positive definite, where $f(a) = P_M(x)$ and $v = [x - P_M(x)]/\| x - P_M(x)\|$. Assume also that $r = \| x - P_M(x)\| <$*

$\rho(P_M(x), v)$. Then P_M is Frechet differentiable at x and $P'_M(x) = f'(a)(B - rA)^{-1} f'(a)^T \nabla^2 \| v \|$ where A, B are as defined in Lemma 3.1.

Proof. Fix y in X and choose t_0 small enough so that $x + ty$ is in U for all $|t| < t_0$. Consider the function F defined by

$$F(t, t_1, \ldots, t_k) = \| x + ty - f(t_1, \ldots, t_k)\|.$$

Observe that F is C^2 in a neighborhood of $(0, \bar{t}_1, \ldots, \bar{t}_k)$ where $f(\bar{t}_1, \ldots, \bar{t}_k) = P_M(x)$ with $(\bar{t}_1, \ldots, \bar{t}_k) = a$.

Now let $G(t, t_1, \ldots, t_k) = (\partial F/\partial t_1, \ldots, \partial F/\partial t_k)$; G is C^1 in a neighborhood of $(0, \bar{t}_1, \ldots, \bar{t}_k)$ and if $P_M(x + ty) = f(t_1, \ldots, t_k)$, then $\partial F/\partial t_i = 0$, $i = 1, \ldots, k$. We proceed by computing the Jacobian matrix J_G of G with respect to $t_1, \ldots, t_k$ at $(\bar{t}_1, \ldots, \bar{t}_k)$. By a straightforward calculation we obtain,

$$\begin{aligned} J_G = \left(\frac{\partial^2 F}{\partial t_i\, \partial t_j}\right)_{i,j} &= \frac{1}{r} f'(a)^T \nabla^2 \| v \| f'(a) - \left(\nabla \| v \| \left(\frac{\partial^2 f(a)}{\partial t_i\, \partial t_j}\right)\right)_{i,j} \\ &= \frac{1}{r} B - A = \frac{1}{r}(B - rA). \end{aligned}$$

Recall that if $0 < \rho < \infty$, then $r = \| x - P_M(x)\| < \rho(P_M(x), v) = \min_{\|w\|=1}\{\langle Bw, w\rangle/\langle Aw, w\rangle \mid \langle Aw, w\rangle > 0\}$. Now if $\langle Aw, w\rangle \leqslant 0$ for all w, then $B - rA$ is positive definite because B is so by hypothesis, while if $\langle Aw, w\rangle > 0$ for some w, then $\langle Bw, w\rangle/\langle Aw, w\rangle \geqslant \rho$ or $\langle Bw, w\rangle \geqslant \rho\langle Aw, w\rangle$ so that $B - rA = (r/\rho)(B - \rho A) + [(\rho - r)/\rho]B$ is positive definite as for each w, $\langle (B - \rho A)w, w\rangle \geqslant 0$; hence $J_G = (1/r)(B - rA)$ is invertible.

By the implicit function theorem, $t_i = t_i(t)$ in a neighborhood of $(0, \bar{t}_1, \ldots, \bar{t}_k)$ and

$$\left(\frac{\partial t_1}{\partial t}, \ldots, \frac{\partial t_k}{\partial t}\right) = -\left(\frac{\partial^2 F}{\partial t_i\, \partial t_j}\right)^{-1}_{i,j} \cdot \left(\frac{\partial^2 F}{\partial t\, \partial t_i}\right)_i.$$

Now $\partial^2 F/\partial t\, \partial t_i = -(1/r) \nabla^2 \| v \| (y, \partial f/\partial t_i)$ and $(\partial^2 F/\partial t\, \partial t_i)_i = -(1/r) \times f'(a)^T \nabla^2 \| v \| (y)$; hence

$$\left(\frac{\partial t_1}{\partial t}, \ldots, \frac{\partial t_k}{\partial t}\right)\Big|_{t=0} = (B - rA)^{-1} f'(a)^T \nabla^2 \| v \|(y).$$

Finally since $P_M(x + ty) = f(t_1(t), \ldots, t_k(t))$, the chain rule gives

$$\frac{d}{dt} P_M(x + ty)|_{t=0} = f'(a)(B - rA)^{-1} f'(a)^T \nabla^2 \| v \|(y)$$

This shows that P_M has directional derivatives. We also know that P_M is a singleton throughout U and, since M is approximatively compact, P_M is continuous on U. Now we consider a neighborhood of x contained in U and write $a = f^{-1}(P_M(x))$. Since f is a local homeomorphism, $f'(a)$, B and A are continuous in x and therefore $f'(a)(B - rA)^{-1} f'(a) \nabla^2 \| v \|$ is also continuous in x. So we finally have

$$P'_M(x) = f'(a)(B - rA)^{-1} f'(a)^T \nabla^2 \| v \| .$$

COROLLARY 4.1. *If X is a Hilbert space, M is a k-dimensional approximatively compact C^2 manifold, x ($\notin M$) is in U and $\| x - P_M(x)\| < \rho(P_M(x), v)$, then*

$$P'_M(x) = f'(a)(B - rA)^{-1} f'(a)^T$$

where $B = f'(a)^T f'(a)$, $A = (a_{ij})_{i,j} = (\langle v, \partial^2 f/\partial t_i\, \partial t_j\rangle)_{i,j}$ and $r = \| x - P_M(x)\|$.

Proof. If y, z belong to a Hilbert space X, then

$$\nabla^2 \| v \| (y, z) = \langle y, z\rangle - \langle v, y\rangle\langle v, z\rangle.$$

Also $v = (x - P_M(x))/\| x - P_M(x)\| \perp \text{Range}(f'(a))$ so that v is in $\text{Ker}(f'(a)^T)$. It follows now easily that

$$P'_M(x) = f'(a)(B - rA)^{-1} f'(a)^T \nabla^2 \| v \| = f'(a)(B - rA)^{-1} f'(a)^T.$$

COROLLARY 4.2. *With the same hypotheses we have*

$$\| P'_M(x)\| = \frac{\rho}{\rho - r} \text{ where } \frac{1}{\rho} = \max_{\|w\|=1} \frac{\langle Aw, w\rangle}{\langle Bw, w\rangle}.$$

Proof. See Corollary 4.1 in [1].

Observe that if M is the sphere $\| z \| = \rho$ in X and x_0 ($\neq 0$) is a point of X whose distance from M is r, then $P_M(x_0) = \rho x_0/\| x_0 \|$ and

$$\| P'_M(x_0)\| = \frac{\rho}{\rho - r} \text{ if } 0 < \| x_0 \| < \rho$$

and

$$\| P'_M(x_0)\| = \frac{-\rho}{-\rho - r} = \frac{\rho}{\rho + r} \text{ if } \| x_0 \| > \rho.$$

EXAMPLE 4.1. Let M be the set $R_m{}^n[0, 1]$ of rational functions: M consists of ratios of the form p/q where p is a real polynomial of degree $\leqslant n$ and q is a real polynomial of degree $\leqslant m$, positive throughout [0, 1] with $q(0) = 1$. We can represent M by a map f from R^{n+m+1} into $L^p[0, 1]$, $p \geqslant 2$:

$$f(a_0, a_1, \ldots, a_n, b_1, \ldots, b_m) = \left(\sum_{i=0}^{n} a_i x^i\right) \Big/ \left(1 + \sum_{j=1}^{m} b_j x^j\right).$$

M is a C^2 manifold and it is also approximatively compact in $L^p[0, 1]$. Therefore each g in $L^p[0, 1]$ has a best approximation from M. Call such a best approximation p_0/q_0; then it is known (see [2]) that p_0/q_0 is a normal element of M. This means that $\dim(p_0 Q + q_0 P) = n + m + 1$ where P and Q are the spaces of real polynomials of degrees $\leqslant n$ and $\leqslant m$, respectively.

We look at the tangent space of M at p_0/q_0 and compute:

$$\frac{\partial(p_0/q_0)}{\partial a_i} = \frac{x^i}{q_0} = \frac{x^i q_0}{q_0^2}, \quad i = 0, 1, \ldots, n,$$

and

$$\frac{\partial(p_0/q_0)}{\partial b_j} = -\frac{x^j p_0}{q_0^2}, \quad j = 1, 2, \ldots, m.$$

It is easy to see that the normality of p_0/q_0 guarantees the linear independence of the $n + m + 1$ partial derivatives, which implies that M is a non-singular C^2 manifold around p_0/q_0.

Also, for $\geqslant 2$ and $g \neq p_0/q_0$ it can easily be shown that, with

$$v = \frac{g - (p_0/q_0)}{\| g - (p_0/q_0)\|_{L_p(0,1)}}$$

$\nabla^2 \| v \|_p$ is positive definite on the tangent space of M at p_0/q_0. So by considerations at the beginning of Section 4, there exist an open set in $L^p[0, 1]$ such that the metric projection P_M into the manifold $M = R_m{}^n[0, 1]$ of rational functions is differentiable with derivative given by Theorem 4.1.

5. On the Differentiability of the Distance Function

R. Holmes proved in [6] that if M is a Chebyshev, convex set in a Banach space X whose norm is C^1, then $d_M(x) = \| x - P_M(x)\|$ is also C^1. We prove the following

THEOREM 5.1. *Let X be a Banach space whose norm is Frechet differentiable and let M be a set in X such that there exists an open subset of X where the metric projection P_M is a singleton and continuous. Then $d_M(x) = \| x - P_M(x)\|$ is C^1 on U, and $\nabla d_M(x) = \nabla \| x - P_M(x)\|$ there.*

Proof. Choose x in V and y in X. Let $t_0 > 0$ be so that $x + ty$ is in V whenever $|t| \leqslant t_0$. Consider $F(t) = \|x + ty - P_M(x + ty)\|$ for $|t| \leqslant t_0$; then $F(t)$ is Lipschitzian and so $F'(t)$ exists a.e. for $|t| < t_0$. For such a fixed t, choose $h \neq 0$ so that $|t| + |h| < t_0$ and set $z = x + ty - P_M(x + ty)$. Then

$$\begin{aligned}\frac{F(t+h)-F(t)}{h} &= \frac{\|z+hy\|-\|z\|}{h}\\ &+ \frac{\|z+hy-P_M(x+ty+hy)+P_M(x+ty)\|-\|z+hy\|}{h}.\end{aligned}$$

Assuming $F'(t)$ exists, we get

$$F'(t) = \lim_{h\to 0}\frac{\|z+hy\|-\|z\|}{h} = \nabla\|z\|(y)$$

because $\|z + hy - P_M(x + ty + hy) + P_M(x + ty)\| - \|z + hy\| = \|x + ty + hy - P_M(x + ty + hy)\| - \|x + ty + hy - P_M(x + ty)\| \leqslant 0$ for small $|h|$.

By hypothesis, $z = x + ty - P_M(x + ty)$ is continuous in t; also since the norm of X is Frechet differentiable we can use Theorem 1 of [5] to conclude that $G(t) = \nabla\|z\|(y)$ is continuous in t. Observe that $F'(t) = G(t)$ a.e. and also that F is absolutely continuous, hence now we can say that $F'(t) = G(t) = \nabla\|z\|(y)$ for all $|t| < t_0$. Finally the continuity of $\nabla\|z\|$ at $z = x - P_M(x)$ implies that $\nabla d_M(x) = \nabla\|x - P_M(x)\|$.

References

1. T. Abatzoglou, The minimum norm projection on C^2-manifolds in R^n, *Trans. Amer. Math. Soc.* **243** (1978), 115–122.
2. E. W. Cheney and A. A. Goldstein, Mean square approximation by generalized rational functions, *Math. Z.* **95** (1967), 232–241.
3. C. K. Chui, E. R. Rozema, P. W. Smith, and J. D. Ward, Metric curvature, folding and unique best approximation, *SIAM J. Math. Anal.* **7** (1976), 436–449.
4. C. K. Chui and P. W. Smith, Unique best nonlinear approximation in Hilbert spaces, *Proc. Amer. Math. Soc.* **49** (1975), 66–70.
5. J. R. Giles, On a characterization of differentiability of the norm of a normed linear space, *J. Austral. Math. Soc.* **12** (1971), 106–114.
6. R. Holmes, Smoothness of certain metric projections on Hilbert space, *Trans. Amer. Math. Soc.* **183** (1973), 87–100.
7. J. R. Rice, "Approximation of Functions II," Addison–Wesley, Reading, Mass., 1969.
8. E. R. Rozema and P. W. Smith, Nonlinear approximation in uniformly smooth Banach spaces, *Trans. Amer. Math. Soc.* **188** (1974), 199–211.