THE MINIMUM NORM PROJECTION
ON $C^2$-MANIFOLDS IN $R^n$

BY
THEAGENIS J. ABATZOGLOU

Abstract. We study the notion of best approximation from a point $x \in R^n$
to a $C^2$-manifold. Using the concept of radius of curvature, introduced by J.
R. Rice, we obtain a formula for the Fréchet derivative of the minimum
norm projection (best approximation) of $x \in R^n$ into the manifold. We also
compute the norm of this derivative in terms of the radius of curvature.

1. Introduction. Let $x \in R^n$ and $M$ be a $C^2$ manifold of dimension $k$,
k < $n$. The minimum norm projection, whenever it is defined, is the map $P_M$
which takes $x$ into the element of $M$ closest to $x$, i.e. $\min_{m \in M} \|x - m\| = \|x
- P_M(x)\|$, where the norm is the Euclidean one. Define $A = \{y | y \in R^n,$
$P_M(y)$ is multivalued$\}$. Let $U = (A)^C$.

J. R. Rice has studied existence of the map $P_M$ in terms of the radius of
curvature and established continuity of $P_M$ on general grounds. In [1], [2], [4]
and [5], we have an investigation of the existence of $P_M$ in a Banach space
setting. The results use the notion of curvature.

In this paper we will examine the existence and differentiability properties
of $P_M$ around point $x$, in relation to the radius of curvature of $M$ at $m$, where
$m = P_M(x)$.

2. Definitions. Let $f$ be a local representation of the manifold $M$ around $m$.
We assume the following:
(1) $f$ is an open map in its domain of definition, i.e. some open set in $R^k$.
(2) $f$ is $C^2$.
(3) $f'(a) \cdot R^k = R^k$.

Furthermore, assuming $f(a) = m$, we define the tangent plane of $M$ at $m$ to
be $T_m = m + f'(a) \cdot R^k$.

A vector $v = y - m$ is orthogonal to $M$ at $m$ if $v$ is orthogonal to $T_m$.

3. Radius of curvature. Consider a vector $v = (y - m)/\|y - m\|$
orthonormal to $T_m$.

We then consider the ray $m + tv, t \geq 0$, and points $\mu \in M$ close to $m$ such
that $\|(m + tv) - m\| = \|(m + tv) - \mu\|$ holds for some $0 < t < \infty$. We
solve for $t$ and obtain
\[ t^2 = \left\| (m - \mu) + t\nu \right\|^2 = \|m - \mu\|^2 + 2t\langle \nu, m - \mu \rangle + t^2, \]
\[ t = \|m - \mu\|^2 / 2\langle \nu, m - \mu \rangle. \tag{3.1} \]

We now define the radius of curvature of \( M \) at \( m \) in the direction \( \nu \) to be
\[ \rho(m, \nu) = \liminf_{\mu \to m} \left\{ \frac{\|m - \mu\|^2}{2\langle \nu, m - \mu \rangle} \mid \langle \nu, m - \mu \rangle > 0 \right\}. \]

**Remark.** If \( \langle \nu, \mu - m \rangle < 0 \) for all \( \mu \) near \( m \) in \( M \) then we define \( \rho(m, \nu) = \infty \).

For Hilbert spaces, this definition is equivalent to that given in [1], [2], [4] and [5].

Since \( M \) is representable by a \( C^2 \) homeomorphism \( f \) near \( m \), for \( \mu \) sufficiently close to \( m \) we can write \( \mu = f(b) \), and
\[ \mu - m = f(b) - f(a) = f'(a)(b - a) + \frac{1}{2} f''(a)(b - a)^2 + o(\|b - a\|^2). \]

Expressing (3.1) in terms of \( a, b, f \) we obtain
\[ t = \frac{\left\| f'(a)(b - a) + \frac{1}{2} f''(a)(b - a)^2 \right\|^2 + o(\|b - a\|^2)}{\langle \nu, f''(a)(b - a)^2 \rangle + o(\|b - a\|^2)}. \]

We divide numerator and denominator by \( \|b - a\|^2 \) and get
\[ \rho(m, \nu) = \min_{\|w\| = 1} \left\{ \frac{\left\| f'(a)(w) \right\|^2}{\langle \nu, f''(a)(w)^2 \rangle} \mid \langle \nu, f''(a)(w)^2 \rangle > 0 \right\}. \]

**Example.** Let \( M \) be the unit sphere in \( \mathbb{R}^3 \) with parametric representation
\[ f(\theta, \phi) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta). \]

Let
\[ x = r(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta), \quad r < 1, \]
and
\[ \nu = \frac{x - f(\theta, \phi)}{\|x - f(\theta, \phi)\|} = -(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta). \]

Assume also that \( w = (w_1, w_2) \); then
\[ \left\| f'(\theta, \phi)(w) \right\|^2 = w_1^2 + \sin^2 \theta w_2^2 \]
and
\[ \langle \nu, f''(\theta, \phi)(w)^2 \rangle = w_1^2 + \sin^2 \theta w_2^2 \to \rho(m, \nu) = 1 \]
in this case.
4. The minimum norm projection \( P_M \). We now examine the existence of \( P_M' \) in terms of the radius of curvature. Recall that by definition \( A \) is the set where \( P_M \) is multivalued. We have

**Lemma 4.1.** Let \( x \in A^c \); then

\[
x \in (\overline{A})^c = U \iff \| x - P_M (x) \| < \rho \left( P_M (x), \frac{x - P_M (x)}{\| x - P_M (x) \|} \right).
\]

**Proof.** Assume \( x \in (\overline{A})^c \); then by Lemma 3.2 in \([1]\),

\[
\| x - P_M (x) \| < \rho (P_M (x), v) \quad \text{where} \quad v = \frac{(x - P_M (x))/\| x - P_M (x) \|}.
\]

Let \( t_0 = \sup \{ t | P_M (m + tv) = m = P_M (x) \} \); then by Theorem 4.8 in \([3]\), \( m + t_0 v \notin (\overline{A})^c \). Combining these 2 results we obtain \( \| x - P_M (x) \| < \rho (P_M (x), v) \); then by Theorems 11–15 in \([4]\) we have

\[
x \in (\overline{A})^c = U.
\]

We continue with a technical lemma:

**Lemma 4.2.** Let \( A = (a_{ij}), B = (b_{ij}) \) be \( k \times k \) matrices where

\[
a_y = \begin{pmatrix} v, \frac{\partial f(a)}{\partial t_i} \frac{\partial f(a)}{\partial t_j} \\ \end{pmatrix}, \quad b_y = \begin{pmatrix} \frac{\partial f(a)}{\partial t_i} \frac{\partial f(a)}{\partial t_j} \end{pmatrix},
\]

\[
v = \frac{x - f(t_1, \ldots, t_k)}{\| x - f(t_1, \ldots, t_k) \|} \quad \text{and} \quad a = (t_1, \ldots, t_k).
\]

Then we have:

(a) \[
\| f'(a)(w) \|_2^2 = \langle Bw, w \rangle,
\]

(b) \[
\langle v, f''(a)(w) \rangle = \langle Aw, w \rangle,
\]

where \( w = (w_1, \ldots, w_k) \).

**Proof.**

(a) \[
f'(a) = \begin{bmatrix} \frac{\partial f}{\partial t_1} & \cdots & \frac{\partial f}{\partial t_i} & \cdots & \frac{\partial f}{\partial t_k} \end{bmatrix}
\]

thinking of \( \partial f/\partial t_i \) as a column vector. Also,

\[
\| f'(a)(w) \|_2^2 = \langle f'(a)(w), f'(a)(w) \rangle = \langle (f'(a))^T f'(a) w, w \rangle.
\]

But
(f'(a))^T f'(a) = \begin{bmatrix}
\frac{\partial f}{\partial t_1} \\
\vdots \\
\frac{\partial f}{\partial t_j} \\
\vdots \\
\frac{\partial f}{\partial t_k}
\end{bmatrix}
\begin{bmatrix}
\frac{\partial f}{\partial t_1} & \cdots & \frac{\partial f}{\partial t_j} & \cdots & \frac{\partial f}{\partial t_k}
\end{bmatrix}
= \left(\frac{\partial f}{\partial t_i}, \frac{\partial f}{\partial t_j}\right)_{ij};

so (a) is proved.

(b) Write f as (f_1, \ldots, f_n) and w = (w_1, \ldots, w_k); then

\[ f''(a)(w) = \left(\sum_{ij} \frac{\partial^2 f}{\partial t_i \partial t_j} w_i w_j\right), \]

so that

\[ \langle v, f''(a)(w) \rangle = \sum_{ij} \left(\frac{\partial^2 f}{\partial t_i \partial t_j} \right) w_i w_j = \langle Av, w \rangle. \]

We now state our main theorem.

**Theorem 4.1.** Let M be a closed C^2 manifold in R^n of dimension k < n. Let x \in U = (A)c; then P_M is Fréchet differentiable at x and the derivative is given by the formula

\[ P'_M(x) = f'(a)(B - rA)^{-1} f'(a)^T, \]

where f is the parametric representation of M, f(a) = P_M(x), r = \|x - P_M(x)\| and A, B are as defined in the previous lemma.

**Proof.** Let x \in U, y \in R^n, and t_0 \exists x + ty \in U for |t| < t_0. Consider a C^2 representation of M around P_M(x), say f: V \to M, where V is open in R^k and a function

\[ F(t, t_1, t_2, \ldots, t_k) = \frac{1}{2} \|x + ty - f(t_1, \ldots, t_k)\|^2. \]

F is obviously C^2.

Let

\[ G(t, t_1, \ldots, t_k) = (\partial F/\partial t_1, \ldots, \partial F/\partial t_k). \]

If \( P_M(x + ty) = f(t_1, \ldots, t_k), \) then \( \partial F/\partial t_i = 0, i = 1, \ldots, k. \) Say \( P_M(x) = f(t_1, \ldots, \tilde{t}_k); \) then we know that G is C^1 in a neighborhood of (0, \tilde{t}_1, \ldots, \tilde{t}_k).

We now investigate the invertibility of the Jacobian matrix of G with respect to \( t_1, \ldots, t_k \) at the point (\tilde{t}_1, \ldots, \tilde{t}_k).
By computation,
\[
\frac{\partial^2 F}{\partial t_i \partial t_j} = -\left( x - f(\bar{t}_1, \ldots, \bar{t}_k), \frac{\partial^2 f}{\partial t_i \partial t_j} \right) + \left( \frac{\partial f}{\partial t_i}, \frac{\partial f}{\partial t_j} \right).
\]

Set
\[
v = \frac{x - f(\bar{t}_1, \ldots, \bar{t}_k)}{\| x - f(\bar{t}_1, \ldots, \bar{t}_k) \|} \quad \text{and} \quad r = \| x - f(\bar{t}_1, \ldots, \bar{t}_k) \|;
\]

then, according to the previous lemma,
\[J_G = B - rA.
\]

Recall that
\[
rho = \rho(m, v) = \min \{ \langle Bw, w \rangle / \langle Aw, w \rangle | \langle Aw, w \rangle > 0 \},
\]

and, by Lemma 4.1, \(0 < r < \rho < \infty\).

If \(\langle Aw, w \rangle < 0\) \(\forall w \in \| w \| = 1\), then \(B - rA\) is invertible, being positive definite.

On the other hand,
\[
\langle (B - rA)w, w \rangle = \langle Bw, w \rangle - r \langle Aw, w \rangle
\]

for all \(w\) in \(R^k \equiv \| w \| = 1\). Therefore \((B - rA)^{-1}\) exists.

By the implicit function theorem, \(t_i = t_i(t)\) in a neighborhood of \((0, \bar{t}_1, \ldots, \bar{t}_k)\), and
\[
\left( \frac{\partial t_1}{\partial t}, \ldots, \frac{\partial t_k}{\partial t} \right) = \left( \frac{\partial^2 F}{\partial t_i \partial t_j} \right)_{ij}^{-1} \left( \frac{\partial^2 F}{\partial t_i \partial t_j} \right)_{ij},
\]

where \(\partial^2 F / \partial t_i \partial t_j = -\langle y, \partial f / \partial t_i \rangle\). Since \(P_M(x + \psi) = f(t_1(t), \ldots, t_k(t))\), using the chain rule we obtain
\[
\frac{d}{dt} P_M(x + \psi) \bigg|_{t=0} = f'(a) \left[ (B - rA)^{-1} \right] f'(a)^T(y).
\]

This shows \(P_M'\) has directional derivatives; also the assumption that \(M\) is closed implies the continuity of \(P_M\) on \(U\). Also \(f'(a)(B - rA)^{-1}f'(a)\) depends continuously on \(a = (t_1, \ldots, t_k)\) and therefore varies continuously with \(x\) so that \(f'(a)(B - rA)^{-1}f'(a)^T\) is the Fréchet derivative of \(P_M\).

We now compute the norm of \(P_M'(x)\).

**Corollary 4.1.** Let \(P_M'(x) = f'(a)(B - rA)^{-1}f'(a)^T; then**
\[ \|P_M'(x)\| = \rho / (\rho - r) = -1 / (1 - r/\rho), \]

where
\[ 1/\rho = \max_{\|w\|=1} \langle Aw, w \rangle / \langle Bw, w \rangle = 1/\rho(m, v). \]

**Proof.** \( P_M'(x) \) is selfadjoint and semipositive definite as we showed in the proof of Theorem 4.1. The rank of \( f'(a)^T \) is \( k = \text{rank of } f'(a) \) by definition. Then we have the rank of \( f'(a)(B - rA)^{-1}f'(a)^T = k \). Choose \( k \) mutually orthogonal eigenvectors, \( \{v_1, \ldots, v_k\} \), of \( P_M'(x) \) such that
\[ \|P_M'(x)\| = \lambda_i \quad \text{where } P_M'(x)(v_i) = \lambda_i v_i \]
for any \( i, 1 < i < k \).

\[ f'(a)(B - rA)^{-1}f'(a)^T(v_i) = \lambda_i v_i; \quad (4.1) \]

therefore
\[ f'(a)^Tf'(a)(B - rA)^{-1}(f'(a)^Tv_i) = \lambda_i (f'(a)^Tv_i). \]

It is clear from (4.1) that \( \{f'(a)^Tv_i\}_{i=1}^k \) is a linearly independent set. Therefore, if \( w \) is an eigenvector of \( f'(a)^Tf'(a)(B - rA)^{-1} \), then we can write
\[ w = \sum_{i=1}^k a_i f'(a)^T(v_i) \quad \text{and } f'(a)^Tf'(a)(B - rA)^{-1}w = \lambda w. \]

So
\[ f'(a)^Tf'(a)(B - rA)^{-1}\sum_{i=1}^k a_i f'(a)^T(v_i) = \sum_{i=1}^k a_i \lambda_i f'(a)^T(v_i). \]

Thus
\[ \sum_{i=1}^k a_i \lambda_i f'(a)^T(v_i) = \lambda \sum_{i=1}^k a_i f'(a)^T(v_i). \]

Hence
\[ a_i \lambda_i = \lambda a_i, \quad i = 1, \ldots, k, \quad \text{and } \lambda_i = \lambda \quad \text{for } a_i \neq 0. \]

This shows that the maximum eigenvalue of \( f'(a)^Tf'(a)(B - rA)^{-1} \) is \( \lambda_1 = \|P_M'(x)\| \). Recall that, by definition, \( f'(a)^Tf'(a) = B \). Thus
\[ f'(a)^Tf'(a)(B - rA)^{-1} = B(B - rA)^{-1} = (I - rAB^{-1})^{-1}, \]

which implies
\[ 1/\lambda_1 = \text{smallest eigenvalue of } I - rAB^{-1}, \]

so
\[ 1/\lambda_1 = 1 - r(\text{largest eigenvalue of } AB^{-1}), \]

and
Consider now \( \max_{\|w\| = 1} \langle Aw, w \rangle / \langle Bw, w \rangle = 1 / \rho \). Therefore,
\[
\langle Aw, w \rangle < \frac{1}{\rho} \langle Bw, w \rangle, \quad \text{so} \quad \left\langle \left( \frac{1}{\rho} B - A \right) w, w \right\rangle > 0).
\]
Since \( B / \rho - A \) is selfadjoint, the min is attained at an eigenvector, and since this min \( = 0 \), \( (B / \rho - A)w_0 = 0 \), and \( w_0 / \rho = B^{-1} Aw_0 \). We claim \( 1 / \rho \) is the largest eigenvalue of \( B^{-1} A \); if not, set \( 1 / \rho' > 1 / \rho \), \( \exists \, w_1 / \rho' = B^{-1} Aw_1 \). Then \( (B / \rho' - A)w_1 = 0 \) and
\[
\left\langle \left( \frac{1}{\rho} B - A \right) w, w \right\rangle = \left\langle \left( \frac{1}{\rho} B - A \right) w, w \right\rangle + \left\langle \left( \frac{1}{\rho'} - \frac{1}{\rho} \right) Bw, w \right\rangle > 0
\]
\( \forall w \, \exists \|w\| = 1 \), a contradiction.

5. Examples.

Example 5.1. Let \( M \) be the sphere of radius \( \rho \) in \( R^3 \). Let \( x \in R^3 \, \exists \|x\| = d \). Set \( r = |\rho - d| \). Then, by the previous corollary,
\[
\|P'_M(x)\| = \frac{\rho_0}{\rho_0 - r} \quad \text{where} \quad \frac{1}{\rho_0} = \max_{\|w\| = 1} \frac{\langle Aw, w \rangle}{\langle Bw, w \rangle}.
\]
By simple computations we obtain
\[
\rho = \rho_0 \quad \text{if} \ d < \rho,
= -\rho_0 \quad \text{if} \ d > \rho,
\]
so that
\[
\|P'_M(x)\| = \begin{cases} \rho / (\rho - r) & \text{if} \ d < \rho, \\ \rho / (\rho + r) & \text{if} \ d > \rho. \end{cases}
\]
This result suggests the following observation: If \( M \) is a \( C^2 \) manifold with radius of curvature \( \rho \) at the point \( m = P_M(x) \), and if \( \|x - P_M(x)\| = r \), then by the previous corollary, \( \|P'_M(x)\| = \rho / (\rho - r) \), which is exactly the same estimate for a sphere of radius \( \rho \) and point \( x \) whose distance from the sphere is \( r \).

Example 5.2. Let \( M \) be \( f(x, y) = (x, y, xy) \) and \( x = (0, 0, r) \) where \( 0 < r < 1 \). Then, from
\[
\|(0, 0, r) - (x, y, xy)\|_2^2 = x^2 + y^2 + (xy - r)^2
= x^2 + y^2 - 2rxy + x^2y^2 + r^2
= (rx - y)^2 + (1 - r^2)x^2 + x^2y^2 + r^2,
\]
it is clear that \( P_M(x) = (0, 0, 0) \). Computing,
Also \( v = (0, 0, 1) \) and
\[
\langle v, \partial^2 f / \partial x^2 \rangle = \langle v, \partial^2 f / \partial y^2 \rangle = 0,
\]
and
\[
\langle v, \partial^2 f / \partial x \partial y \rangle = \langle v, \partial^2 f / \partial y \partial x \rangle = 1,
\]
so that \( A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \). Therefore,
\[
P_M'(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} 1 & -r \\ -r & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \frac{1}{1 - r^2} \begin{pmatrix} 1 & r & 0 \\ r & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},
\]
and
\[
\|P_M'(x)\| = \frac{\rho}{\rho - r}, \quad \text{where} \quad \frac{1}{\rho} = \max_{\|w\|=1} \frac{\langle Aw, w \rangle}{\langle Bw, w \rangle} = 1.
\]
Thus \( \|P_M'(x)\| = 1/(1 - r) \).

**EXAMPLE 5.3.** Let
\[
M = \begin{cases} 
(x, 0), & x < 0, \\
(x, 1 - \sqrt{1 - x^2}), & x > 0, x < 1;
\end{cases}
\]
then for \( |\epsilon| < \frac{1}{2} \) we have
\[
P_M(\epsilon, \frac{1}{2}) = (\epsilon, 0) \quad \text{if} \ \epsilon < 0,
\]
\[
= \left( \frac{2\epsilon}{\sqrt{1 + 4\epsilon^2}}, 1 - \frac{1}{\sqrt{1 + 4\epsilon^2}} \right) \quad \text{if} \ \epsilon > 0.
\]
But \( dP_M(\epsilon, \frac{1}{2})/d\epsilon \big|_{\epsilon=0} \) does not exist. Our manifold \( M \) is \( C_1 \) at \( (0, 0) \) but not \( C^2 \).

**REFERENCES**


**DEPARTMENT OF MATHEMATICS, IOWA STATE UNIVERSITY, AMES, IOWA 50011**