A Note on Non-Linear Approximation Theory

A. A. Goldstein
A NOTE ON NON-LINEAR APPROXIMATION THEORY

by

A. A. Goldstein*
University of Washington, Seattle

Mathematical Note No. 482
Mathematics Research Laboratory
BOEING SCIENTIFIC RESEARCH LABORATORIES
September 1966

*Supported by Air Force Office of Scientific Research and the Boeing Scientific Research Laboratories
ABSTRACT

Recently J. R. Rice [1] initiated a geometrical study of non-linear approximations. In what follows below we offer a small contribution to certain analytical aspects of mean-square non-linear approximations. One result is to exhibit a family of non-linear topological subspaces of the space $C[a,b]$ with the mean-square metric which has a local unique best approximation property.
Recently J. R. Rice [1] initiated a geometrical study of non-linear approximations. In what follows below we offer a small contribution to certain analytical aspects of mean-square non-linear approximations. One result is to exhibit a family of non-linear topological subspaces of the space \( C[a,b] \) with the mean-square metric which has a local unique best approximation property.

Let \( S \) be a subset of a normed linear space \( E \). Let \( \hat{S} \) be any open set containing \( H(S) \), the convex hull of \( S \). Let \( A \) denote a map from \( \hat{S} \) into a real inner product space \( F \). Assume the range of \( A \) is not dense in \( F \). Assume that \( A \) is twice Gateaux (G) differentiable on \( H(S) \). Assume, moreover, that for each point \( x \in H(S) \) these G-derivatives are bounded linear operators, that is \( A'(x) \) is a bounded linear operator from \( E \) to \( F \) and \( A''(x) \) is a bounded linear operator from \( E \) to the space of bounded linear operators from \( E \) to \( F \). Let \( p \) be a point in \( F \) which is not in the closure of the range of \( A \). Set

\[
h(x) = ||A(x)-p||^2,\]

\[
u = \inf \left\{ ||A'(x)k||/||k|| : x \in H(S) \text{ and } x+k \in S \right\},\]

\[
\gamma = \inf \left\{ ||A(z+k)-A(z)||/||k|| : z \text{ and } z+k \in S \right\} \]

\[
C = \sup \left\{ ||A'(x)|| : x \in H(S) \right\}.\]

The notations \( A'(x)k \) and \( (A''(x)k)(u) \) will be employed for first and second differentials, respectively. For simplicity the norms in \( E \) and \( F \) are denoted by the same symbol \( ||\cdot|| \).
We show first that if \( z \in S \) is stationary point of \( h \), that is \( h'(z)k = 0 \) for \( z+k \in S \), and \( h(z) \) is sufficiently small, then \( z \) is global minimizer for \( h \). State otherwise, \( A(z) \) is a best approximation to \( p \) out of the range of \( A \). Moreover, \( z \) is unique.

**Lemma 1.** If \( A \) is twice differentiable on \( \mathcal{H}(S) \), then so is \( h \). Let \( \theta = z + \alpha k \) for any \( \alpha \in (0,1) \), where \( z \) and \( z+k \) are in \( S \). Then

\[
\frac{1}{2}(h''(\theta)k)(k) = [A(z)-p,(A''(\theta)k)(k)] + [A(\theta)-A(z),(A''(\theta)k)(k)] + [A'(\theta)k,A'(\theta)k].
\]

**Proof.** If \( f : E \to F \) is Frechet differentiable and \( g : F \to \mathbb{R} \) is Gateaux differentiable, the composition \( gf \) satisfies the chain-rule \( h'(z)k = g'(f(z))f'(z)k \). (See e.g. [2, p. 659]). If \( f(\theta) = A(\theta) - p \) and \( g(x) = ||x||^2 \), then \( h'(\theta)k = 2[A(\theta)-p,A'(\theta)k] \). By writing

\[
\frac{1}{2t} [h'(\theta+tk)k-h'(\theta)k] = \frac{1}{t} \{ [A(\theta+tk)-A(\theta)+A(\theta)-p, (A'(\theta+tk)-A'(\theta))k+A'(\theta)k]-[A(\theta)p,A'(\theta)k] \},
\]

expanding, and passing to the limit as \( t \to 0 \), we get:

\[
\frac{1}{2}(h''(\theta)k)(k) = [A(\theta)-p,(A''(\theta)k)(k)] + [A'(\theta)k,A'(\theta)k].
\]

Whence the lemma.

**Theorem 1.** Assume that \( C \) and \( B \) are finite and that \( \mu \) and \( \gamma \) are positive. Assume \( z \) is a stationary point for \( h \). If \( h(z) < \frac{\mu^2}{B(1 + \frac{2C}{\gamma})} \), then \( h(x) < h(z) \) for all \( x \neq z \) in \( S \).
Suppose that $z$ is not a unique best approximation. Then for some $k \neq 0$ and $z+k \in S$, $||A(z+k)-p|| \leq h(z)$. Since $||A(z)-p|| = h(z)$ it follows that $2h(z) > ||A(z+k)-A(z)|| > ||k|| \gamma$; whence $||k|| \leq 2h(z)/\gamma$.

Since $h'(z)k = 0$ for all $k \in E$ we have by Taylor's Theorem that $h(z+k) - h(z) = (h''(\theta)k)(k)/2$ for some $\theta$ on the open line segment joining $z$ and $z+k$. It would be a contradiction to show $h(z+k) - h(z) > 0$.

By the above lemma we have:

$$\frac{1}{2}(h''(\theta)k)(k) > -Bh(z) ||k||^2 - CB||k||^3 + \mu^2 ||k||^2,$$

where the inequality $||A(\theta)z-A(z)|| \leq C||k||$ has been employed [2, p. 659]. Since $||k|| \leq 2h(z)/\gamma$, it follows that $(h''(\theta)k)(k)$ is positive whenever $h(z) < \mu^2/(1 + 2C)$. Q.E.D.

**Remark 1.** The hypothesis that $\gamma$ is positive may be replaced by the hypothesis that $S$ is bounded by a sphere of radius $R < \mu^2/2CB$. Then if $h(z) < \mu^2/2B$, the above theorem holds. Observe also that if $A$ is a linear operator, $B = 0$ and $z$ is a best approximation independently of the value of $h(z)$.

**Remark 2.** Assume the hypothesis of Theorem 1. Assume $E$ is finite dimensional and $A$ is defined on $E$. Assume that for some $x_0 \in S$ the level set $D = \{x \in E : h(x) \leq h(x_0)\}$ is a proper subset of $S$. Then $p$ has a best approximation out of the range of $A$. 
Proof. Take $x$ in $D$. Then $h(x_0) \geq u \|x-x_0\| - h(x_0)$. Hence $D$ is bounded. By the continuity of $h$, $D$ is closed. Clearly in seeking the minimum of $h$ we may confine our attention to $D$. Thus $h$ achieves a minimum on $D$. Q.E.D.

Remark 3. Observe that if we assume the hypotheses of Remark 2, and the additional hypotheses that $x_0$ itself is not a minimizing point for $h$, then for every $p \in F$ there exists stationary points $x$ for $h$, that is points $x \in S$ such that $[A(x)-p,A'(x)k] = 0$ for all $x+k \in S$. To see this observe that because $h$ is continuous, $D$ has interior, and therefore a necessary condition that a point $z$ minimize $h$ is that $h'(z) = 0$.

Remark 4. If $S$ is complete and $u$ and $\gamma$ are positive, then the ranges of $A$, and $A'(x)$ are closed for all $x \in H(S)$.

Proof. Let $\{A(x^k)\}$ be any Cauchy sequence in the range of $A$. Since $\|A(x^k)-A(x^s)\| \geq \gamma \|x^k-x^s\|$, the sequence $\{x^k\}$ is also Cauchy with limit, say, $y$. But $\|A(x^k)-A(y)\| \leq C\|x^k-y\|$, showing that $A$ is closed. The same proof works for $A'(x)$.

We now turn our attention to the "reverse problem" of approximation theory: Given a point $y$ on the range of $A$, does there exist a point $p \neq y$ whose best approximation is $y$?

Theorem 2. Assume the hypotheses of Theorem 1. Assume that $S = E$ and $E$ is complete. Fix $x \in S$. There exists a point $p \in F$ such that $A(x)$ is a unique best approximation to $p$ if and only if $A'(x)$ is not onto $F$. 

Proof. We prove first that the not onto property for \( A'(x) \) is necessary. Otherwise, \( ||A(x)-p|| < ||A(y)-p|| \) for all \( y \neq x \) but the range of \( A'(x) \) is onto. Since \([A(x)-p,A'(x)k]\) = 0 for all \( k \in E \), and we may choose \( \bar{k} \) such that \( A'(x)\bar{k} = A(x)-p \neq 0 \), we get \([A(x)-p,A'(x)\bar{k}] > 0\), a contradiction. Conversely, if \( A'(x) \) is not onto, then since the range of \( A'(x) \) is closed, it follows the deficiency of the range of \( A'(x) \) exceeds 0. Let \( O(x) \) denote the orthogonal complement of the range of \( A'(x) \). Clearly if \( p \in O(x) + A(x) \), \( x \) is a stationary point. By Theorem 1, moreover, if \( ||A(x)-p|| < \mu^2/B(1+2C) \), then \( A(x) \) is a unique point on the range of \( A \) closest to \( p \).

Example. Let \( C[a,b] \) denote the space of continuous function on \([a,b]\). In \( C[a,b] \) define an inner product \( \langle f,g \rangle = \int_a^b f(t)g(t)dt \). In \( E_n \) denote by \([x,y]\) the inner product of \( n \)-tuples \( x \) and \( y \). Let \( ||\cdot||_2 \) and \( ||\cdot|| \), respectively, be the norms in \( C[a,b] \) and \( E_n \) which arise from these inner products. Let \( v(t) = (v_1(t),...,v_n(t)) \), \( v_i \in C[a,b] \), \( (i=1,2,...,n) \), be a Haar family on \([a,b]\). This means that for any subset of distinct points \( t_i \), \( (1 \leq i \leq n) \), the matrix \( \{v_j(t_i) : 1 \leq i, j \leq n \} \) has rank \( n \). Thus if multiple roots are counted only once, the generalized polynomial \( p(x,t) = \langle v(t),x \rangle \) has at most \( n-1 \) roots. Let \( P = ||p(x,t)|| \), i.e. \( ||\langle v(\cdot),x \rangle ||_2 \leq P||x|| \) for all \( x \in E_n \).

We now specify the mapping \( A \) discussed above as follows: \( A(x) = f(\langle v(\cdot),x \rangle) \), where \( f \) is the continuous real-valued function specified below. Thus \( A : E_n \to C[a,b] \).
Specifically, let \( f \) be a differentiable, real-valued function defined everywhere on the reals, \( \mathbb{R} \). Assume that \( f'(t) \geq \alpha > 0 \) for all \( t \in \mathbb{R} \), and that \( f'' \) exists and is bounded above on \( \mathbb{R} \), say by \( N \).

**Lemma 2.** The mapping \( A \) satisfies the hypotheses of the above Theorems 1 and 2.

**Proof.** Because \( x \neq 0 \) implies \([v(\cdot),x]\) has at most \( n-1 \) roots, \( ||[v(\cdot),x]||_2 > 0 \) for all \( x \neq 0 \). Thus \( ||[v(\cdot),x]||_2 \geq \beta ||x|| \) for some \( \beta > 0 \) and all \( x \in E_n \). We shall denote the point \( A(x) \) in \( C[a,b] \) by \( A(x,\cdot) \) when the pointwise values \( A(x,t) \) are singled out for attention.

Since \( A'(x,t)k = f'([v(t),x])[v(t),k], \) we get \( ||A'(x,\cdot)k||_2 = ||f'([v(\cdot),x])[v(\cdot),k]||_2 \geq \alpha ||[v(\cdot),k]||_2 \geq \alpha \beta ||k|| = \nu ||k|| \). Similarly using the ordinary mean-value theorem on the real-valued functions \( f([v(t),\cdot]) \) defined on the ray \( \{z = x+\theta k : 0 < \theta < 1 \} \) we get

\[
||A(x+k,\cdot) - A(x,\cdot)||_2/||k|| = ||f([v(\cdot),x+k]) - f([v(\cdot),x])||_2/||k|| =
||f'([v(\cdot),x])k||_2 \geq \alpha ||[v(\cdot),k]||_2 = \nu. \]

Thus \( \gamma = \nu \) in the above Theorem 1.

By Remarks 2 and 3 there exists at least one point \( z \in S \) such that \( [A(z)-p,A'(z)k] = 0 \) for all \( k \in E_n \).

We calculate that \( ||(A''(x)k)(k)|| = \)
\[
||f''([v(\cdot),x])[v(\cdot),k][v(\cdot)k]||_2 \leq \alpha^2 ||k||^2, \]
whence \( \alpha^2 = N \). Let \( M = \max\{f'(r) : r \in \mathbb{R}_0\} \). Then \( ||A'(x)|| \leq MP = C. \)
It remains only to check that \( A \) and \( A'(x) \) are not onto \( F \). The former follows because \( f \) is strictly monotone and continuous. Thus the range of \( A \) is a homeomorphic image of the span of \( \{v_1, \ldots, v_n\} \), which is a finite dimensional subspace of \( C[a,b] \). For the latter we observe that the range of \( A'(x) \) for each \( x \) is spanned by the functions \( f'(v(\cdot), x)v_j \), \( 1 \leq j \leq n \). Thus \( A'(x) \) is a finite dimensional linear subspace. Q.E.D.

Remark 5. We have constructed for each function \( f \) satisfying the above hypotheses a topological subspace of \( C[a,b] \) with the mean-square metric which is homeomorphic to a Haar subspace. This subspace has a local unique approximation property—each point sufficiently close to the subspace has a unique best approximation in the subspace. Moreover, every point on the subspace is the unique best approximation to some nearby point which is not in the subspace.
REFERENCES
