JOURNAL OF APPROXIMATION THEORY 20, 1–9 (1977)

Existence and Convergence of Discrete Nonlinear Best L_2 Approximations

JERRY M. WOLFE

Department of Mathematics, University of Oregon, Eugene, Oregon 97403

Communicated by E. W. Cheney

Received May 8, 1975

In the application of nonlinear approximation theory one is usually constrained to the calculation of best approximations on certain finite subsets of a given domain. Two basic questions immediately arise: (1) Does a best approximation exist on such a finite set? And (2) If best approximations are calculated on a sequence of finite sets that "fill out" the domain (in some sense) then do the calculated approximations converge to a best approximation over the whole domain?

In the first section of this paper we study these two questions in the context of nonlinear approximation of continuous functions on finite subsets of the interval $[-1, 1]$ in the least-squares sense. In the second section we consider the rate of convergence of discrete approximations to continuous ones. The results obtained will apply to many types of rational approximations, to exponential approximation, and more generally to most of the so-called Γ-families of Hobby and Rice [1]. The setting for our analysis is as follows. Let $f \in C[-1, 1]$, $S \subset E^N$ be open and let $A : S \to C[-1, 1]$ be such that

The map $(x, t) \to A(x)(t)$ defines an analytic function (of $N + 1$ variables) on $S \times [-1, 1]$. (1)

Note that (1) implies that A has continuous Fréchet derivatives of all orders on S with respect to the uniform norm on $C[-1, 1]$. We now present the following two examples to illustrate condition (1) above.

EXAMPLE 1. Let $S = E^{2N} = \{(a_1, ..., a_N, \lambda_1, ..., \lambda_N) \mid a_i, \lambda_i \in E\ i = 1, ..., N\}$ and define $A : S \to C[-1, 1]$ by $A(a_1, ..., a_N, \lambda_1, ..., \lambda_N)(t) = A(a, \lambda)(t) = a_1 e^{\lambda_1 t} + \cdots + a_N e^{\lambda_N t}$. Then clearly A is an analytic function of the a_i's, λ_i's, and t, so (1) is satisfied.

EXAMPLE 2. Let $S = \{(a_0, ..., a_n, b_1, ..., b_m) \in E^{m+n+1} \mid 1 + b_1 t + \cdots + b_m t^m > 0$ for all $t \in [-1, 1]\}$ and define $A(a_0, ..., a_n, b_1, ..., b_m)(t) = A(a, b)$

ISSN 0021-9045

$(t) = (a_0 + a_1 t + \cdots + a_n t^n)/(1 + b_1 t + \cdots + b_m t^m)$. Again it is clear from this formula that the map is analytic in the a_i's, b_i's, and t so that the ordinary rational functions also satisfy (1).

Before proceeding, a word on notation might be helpful. If ϕ is a map defined on the open set U of a normed linear space X with values in a normed linear space Y and if $x \in U$ then $\phi^{(j)}(x)$ will denote the jth Frechet derivative of ϕ at x for j a positive integer. When it is necessary to evaluate this multilinear operator at some set of j values $\{h_1, \ldots, h_j\}$ we will use the notation $\phi^j(x)(h_1, \ldots, h_j)$. Also for typographical clarity, when the inverse of the derivative $\phi'(x)$ is needed, we will sometimes use the notation $\phi'_{-1}(x)$ rather than the more cumbersome $\phi'(x)^{-1}$.

The problem then is to minimize the functional $\psi(x) = \int_{-1}^{1} [A(x)(t) - f(t)]^2 \, dt = [A(x) - f, A(x) - f]$ over S. For each $g \in C[-1, 1]$ the integral is approximated by a quadrature formula of the form $\sum_{j=0}^{M} \alpha_{jM} g(t_{jM})$ where α_{jM} and t_{jM} are fixed, $j = 0, \ldots, M$, and $\{t_{jM}\} \subset [-1, 1]$. The discrete problem then is to minimize

$$\sum_{j=0}^{M} \alpha_{jM}[A(x)(t_{jM}) - f(t_{jM})]^2 = [A(x) - f, A(x) - f]_M$$

over S.

Our analysis will be carried out by studying the functions arising in the following simple lemma.

LEMMA 1. *Assume the setting above. Then a necessary condition that x be a local minimum of $\psi(x)(\psi_M(x))$ is that*

$$F_i(x) = \left[A(x) - f, \frac{\partial A}{\partial x_i}(x)\right] = 0$$

$$\left(F_{iM}(x) = \left[A(x) - f, \frac{\partial A}{\partial x_i}(x)\right]_M = 0\right), \qquad i = 1, \ldots, N.$$

Proof. Since $A(x)$ is Fréchet differentiable on S, $\psi(x)$ is clearly differentiable and so at a local minimum x, $\partial\psi/\partial x_i(x) = 0$, $i = 1, \ldots, N$ (the same is clearly true of $\psi_M(x)$) and calculating we have $\partial\psi/\partial x_i(x) = 2[A(x) - f, \partial A/\partial x_i(x)]$ and $\partial\psi_M/\partial x_i(x) = 2[A(x) - f, \partial A/\partial x_i(x)]_M$. ∎

Let U be a bounded open convex subset of E^N and define X as $\{\sigma : U \to E^N \mid \sigma'(x)$ exists and is continuous on $\overline{U}\}$. Then X is a real linear vector space and becomes a normed linear space if we define $N_1(\sigma) = \sup_{x \in \overline{U}} \|\sigma(x)\| + \sup_{x \in \overline{U}} \|\sigma'(x)\|$, where $\|\cdot\|$ is some vector norm on E^N and the derivative norm is the induced operator norm. The basic existence result of this paper is based on the following fundamental lemma.

LEMMA 2. *Let $F_0 \in X$ be such that $F_0(x_0) = 0$ and $F_0'(x_0)$ is nonsingular where $x_0 \in U$. Then there is a ball B about x_0 in U and a $\delta > 0$ such that if $N_1(F - F_0) < \delta$ then there is a unique $x(F) \in B$ such that $F(x(F)) = 0$. In fact, the map $F \to x(F)$ is Fréchet differentiable.*

Proof. Define $\Omega : X \times U \to E^N$ by $\Omega(\sigma, x) = \sigma(x)$. Then $\partial\Omega/\partial\sigma(\sigma, x)$ exists and is given by the relationship $\partial\Omega/\partial\sigma(\sigma, x)(\tau) = \tau(x)$ for each $\tau \in X$. (Here, of course, $\partial\Omega/\partial\sigma$ means the Fréchet derivative of ψ with respect to σ.) Also, $\partial\Omega/\partial x\,(\sigma, x) = \sigma'(x)$ for each $x \in U$ and $\sigma \in X$. Thus, $\Omega(F_0, x_0) = 0$ and $\partial\Omega/\partial x\,(F_0, x_0)$ is nonsingular by hypothesis. Also the map $(F, x) \to \partial\Omega/\partial x\,(F, x) = F'(x)$ is continuous on $X \times U$ since if $N_1(F_r - F) + |x_r - x| \to 0$ then

$$\left|\frac{\partial\Omega}{\partial x}(F_v, x_v) - \frac{\partial\Omega}{\partial x}(F, x)\right| = |F_r'(x_r) - F'(x)|$$

$$\leqslant |F_r'(x_r) - F'(x_r)| + |F'(x_r) - F'(x)|$$

$$\leqslant N_1(F_r - F) + |F'(x_r) - F'(x)| \to 0$$

$$\text{as} \quad r \to \infty.$$

Thus, the implicit function theorem [2, p. 230] applies and so there exists a ball B_0 about F_0 (of radius δ, say) and a ball B_1 about x_0 in U and a differentiable map $x : B_0 \to B_1$ such that $x(F_0) = x_0$ and for each $F \in B_1$ we have $F(x(F)) = 0$. Moreover, $x'(F, \tau) = -(\partial\Omega/\partial x)^{-1}(F, x(F))(\partial\psi/\partial\sigma(F, x(F))(\tau)) = -F'_{-1}(x(F))(\tau(x(F)))$. In particular, $x'(F_0)(\tau) = -F'_{-1}(x_0)(\tau(x_0))$ for each $\tau \in X$. ∎

Remark 1. It is simple to show that the mapping Ω of Lemma 1 is in fact continuously djfferentiable so that the map $F \to x(F)$ also has this property. See also Lemma 3 of Section 2.

In order to apply Lemma 2 to the discrete approximation problem we make one further mild assumption about the quadrature formulas employed. This condition is satisfied by all the standard methods for numerical integration (see [5, p. 343]).

ASSUMPTION. For each $g \in C[-1, 1]$ the quadrature formulas are such that

$$\left|\int_{-1}^{1} g(t)\,dt - \sum_{j=0}^{M} \alpha_{jM}\, g(t_{jM})\right| \leqslant CW(g, \Delta_M)$$

where C is some constant independent of g, $\Delta_M = \max_{0 \leqslant j \leqslant M-1} |t_{j+1,M} - t_{jM}|$, and $W(g, \cdot)$ is the modulus of continuity of g.

THEOREM 1. *Suppose $x_0 \in S$ is such that the map*

$$F_0(x) = \left(\left[A(x) - f, \frac{\partial A}{\partial x_1}(x)\right], \ldots, \left[A(x) - f, \frac{\partial A}{\partial x_N}(x)\right]\right)^T$$

satisfies $F_0(x_0) = 0$ and such that $F_0'(x_0)$ is nonsingular. Also assume that $\Delta_M \to 0$ as $M \to \infty$. Then there is a ball B about x_0 and an M_0 such that for all $M \geqslant M_0$ there exists a unique $x_M \in B$ such that $F_M(x_M) = 0$. Moreover, $x_M \to x_0$ as $M \to \infty$.

Proof. Using Lemma 2 it is sufficient to show that $N_1(F_M - F) \to 0$ as $M \to \infty$, where in this case the set U used to define the space X of Lemma 2 can be chosen as an open ball centered at x_0. Now

$$F'(x) = \left(\left[A(x) - f, \frac{\partial^2 A}{\partial x_i\, \partial x_j}(x)\right] + \left[\frac{\partial A}{\partial x_i}(x), \frac{\partial A}{\partial x_j}(x)\right]\right)_{1 \leqslant i,\ j \leqslant N}$$

and

$$F_M'(x) = \left(\left[A(x) - f, \frac{\partial^2 A}{\partial x_i\, \partial x_j}\right]_M + \left[\frac{\partial A}{\partial x_i}(x), \frac{\partial A}{\partial x_j}(x)\right]_M\right)_{1 \leqslant i,\ j \leqslant N.}$$

To show $N_1(F_M - F_0) \to 0$, it is clearly sufficient to show that

$$\sup_{x \in \bar{U}} \max_i \left|\left[A(x) - f, \frac{\partial A}{\partial x_i}(x)\right] - \left[A(x) - f, \frac{\partial A}{\partial x_i}(x)\right]_M\right| \to 0 \qquad (*)$$

and

$$\sup_{x \in \bar{U}} \max_{i,j} \left|\left[A(x) - f, \frac{\partial^2 A}{\partial x_i\, \partial x_j}\right] - \left[A(x) - f, \frac{\partial^2 A}{\partial x_i\, \partial x_j}(x)\right]_M + \left[\frac{\partial A}{\partial x_i}(x), \frac{\partial A}{\partial x_j}(x)\right] - \left[\frac{\partial A}{\partial x_i}(x), \frac{\partial A}{\partial x_j}(x)\right]_M\right| \to 0$$

$$\text{as} \quad M \to \infty. \qquad (**)$$

But,

$$\begin{aligned}
&\left|\left[A(x) - f, \frac{\partial A}{\partial x_i}(x)\right] - \left[A(x) - f, \frac{\partial A}{\partial x_i}(x)\right]_M\right| \\
&\quad \leqslant CW\left((A(x) - f)\frac{\partial A}{\partial x_i}(x), \Delta_M\right) \\
&\quad \leqslant CW\left(A(x)\frac{\partial A}{\partial x_i}(x), \Delta_M\right) + CW\left(f\frac{\partial A}{\partial x_i}(x), \Delta_M\right) \\
&\quad \leqslant C\left(\|A(x)\|_\infty W\left(\frac{\partial A}{\partial x_i}(x), \Delta_M\right) + \left\|\frac{\partial A}{\partial x_i}(x)\right\|_\infty W(A(x), \Delta_M)\right) \\
&\qquad + \|f\|_\infty W\left(\frac{\partial A}{\partial x_i}(x), \Delta_M\right) + \left\|\frac{\partial A}{\partial x_i}\right\|_\infty W(f, \Delta_M) \\
&\quad \leqslant C_1 \Delta_M + C_2 W(f, \Delta_M).
\end{aligned}$$

where C_1 and C_2 are independent of i and x for $x \in \bar{U}$. In obtaining the last inequality we have used assumption (1) to obtain uniform Lipschitz constants independent of $x \in \bar{U}$ for $A(x)$ and $\partial A/\partial x_i (x)$, $i = 1,..., N$. Thus (*) $\to 0$ and $M \to \infty$. Proceeding in a similar way (using assumption (1)) we find also that there are constants C_3 and C_4 independent of i, j and $x \in \bar{U}$ such that

$$\Big|\Big[A(x) - f, \frac{\partial^2 A}{\partial x_i\, \partial x_j}(x)\Big] + \Big[\frac{\partial A}{\partial x_i}(x), \frac{\partial A}{\partial x_j}(x)\Big] - \Big[A(x) - f, \frac{\partial^2 A}{\partial x_i\, \partial x_j}(x)\Big]_M - \Big[\frac{\partial A}{\partial x_i}(x), \frac{\partial A}{\partial x_j}(x)\Big]_M\Big| \leqslant C_3 \Delta_M + C_4 W(f, \Delta_M)$$

for all $i \leqslant 1, j \leqslant N$. Since $\Delta_M \to 0$ as $M \to \infty$, we conclude $N_1(F - F_M) \to 0$ and so by Lemma 2 the conclusion of the theorem is valid. ▮

COROLLARY 1. *Assume the hypotheses of Theorem* 1 *hold and in addition that* $F_0'(x_0)$ *is positive definite and that* A^{-1} *exists and is continuous on a relative neighborhood of* $A(x_0)$. *Then* $A(x_0)$ *is a local best approximation to* f *and each* $A(x_M)$ *(for* M *sufficiently large) is a local best discrete approximation to* f.

Proof. Clearly x_0 is a local minimum of the functional $\psi(x) = [A(x) - f, A(x) - f]$ and the continuity of A^{-1} at $A(x_0)$ implies that $A(x_0)$ is a local best approximation to f. From the convergence of x_M to x_0 and F_M to F_0 in the norm topology of X we have that $F_M'(x_M)$ is positive definite for M sufficiently large that x_M is a local minimum of $\psi_M(x) = [A(x) - f, A(x) - f]_M$; the continuity of A^{-1} on a neighborhood of $A(x_0)$ yields the desired conclusion. ▮

Remark 2. In special cases, such as ordinary rational approximation [3] or certain types of Γ families [4], it is easy to show that if $A(x_0)$ is a unique (global) best approximation for the continuous problem, then for M sufficiently large, $A(x_M)$ is in fact a (global) best discrete approximation. For example, in the rational function case, a discrete best approximation exists for M sufficiently large and the sequence of these best approximants converge uniformly to the unique continuous best approximation $A(x_0)$ [3]. Since these discrete best approximations satisfy the equations of Lemma 1, we imply from the uniqueness part of Theorem 1 that for M sufficiently large $A(x_M)$ is a (global) discrete best approximation to f.

CONVERGENCE OF BEST APPROXIMATIONS

It seems reasonable that the accuracy of the quadrature formulas used to approximate the integral should affect the rate at which the parameters of the discrete approximations converge to the best approximation parameters for the continuous problem. In this section we shall demonstrate that this is indeed true provided that the function approximated is sufficiently smooth. To simplify the analysis we shall assume that the data points t_{jM}, $j = 0, \dots, M$ are all equally spaced.

Suppose $x_0 \in S$ is such that $F_0(x_0) = 0$ and $F_0'(x_0)$ is invertible and let U be an open ball centered at x_0. For $j = 1, 2, \dots$ let $X_j = \{\sigma : U \to E^N \mid \sigma^{(j)}(x)$ exists and is continuous on $\bar{U}\}$ and define a norm N_j on X_j by

$$N_j(\sigma) = \sum_{l=0}^{j} \sup_{x \in \bar{U}} |\sigma^{(l)}(x)|.$$

LEMMA 3. *The map $\Omega(F, x) = F(x)$ defined on $X_j \times \bar{U} \to E^N$ is j times continuously differentiable.*

Proof. Since the Fréchet derivative $\Omega^{(j)}(F, x)$ will be a jth-order multilinear operator on $X_j \times E^N$ into E^N, it is sufficient to describe it by its action on an arbitrary j-tuple of points in $X_j \times E^N$, say $((\tau_1, h_1), (\tau_2, h_2), \dots, (\tau_j, h_j))$. One finds by a direct and simple calculation that the directional derivative (which we do not distinguish notationally from the Fréchet derivative) exists in the given direction and is given by the formula

$$\begin{aligned}\Omega^{(j)}(F, x)&((\tau_1, h_1), \dots, (\tau_j, h_j)) \\ &= F^{(j)}(x)(h_1, \dots, h_j) + \tau_1^{(j-1)}(x)(\hat{h}_1, \dots, h_j) + \tau_2^{(j-1)}(x)(h_1, \hat{h}_2, \dots, h_j) \\ &\quad + \cdots + \tau_j^{(j-1)}(x)(h_1, \dots, \hat{h}_j),\end{aligned}$$

where $\hat{h}_l$ denotes that the lth term is to be omitted. From this formula and the fact that $F, \tau_1, \dots, \tau_j$ have continuous jth-order Fréchet derivatives on $\bar{U}$, it follows easily that $\Omega(F, x)$ has j Fréchet derivatives on $X_j \times \bar{U}$ and that if $N_j(F_r - F) + |X_r - X| \to 0$, then $\|\Omega^{(j)}(F_r, X_r) - \Omega^{(j)}(F, x)\| \to 0$. ∎

The next corollary is an immediate consequence of the above and the implicit function theorem.

COROLLARY 2. *The map $x(\cdot)$ (the existence of which is guaranteed by Theorem 1) defined implicitly by $\Omega(F, x) = 0$ is j times continuously differentiable on its domain of definition.*

Let $h = 2/M$ and $I_h(g) = \sum_{j=0}^{M} a_{jM} g(t_{jM})$ and assume that $\int_{-1}^{1} g(x)\,dx - I_h(g) = C_1 h^P$ if $g \in C^{(P)}[-1, 1]$, where C_1 is of the form $\alpha g^{(P)}(\xi)$ for some $\xi \in (-1, 1)$. Also assume that there is an $m \geq 1$ such that if $g \in C^{(m+P)}[-1, 1]$, then the above error is of the form $C_1' h^P + O(h^{m+P})$ where C_1' is independent of h.

EXAMPLE. Consider the composite trapezoid rule given by $I_h(g) = (h/2)[g(1) + g(-1)] + h\sum_{j=1}^{M-1} g(t_{jM})$. Then it is well known [5] that $P = 2$ and that if $g \in C^{(4)}[-1, 1]$ then

$$\int_{-1}^{1} g(x)\,dx - I_h(g) = \frac{[g'(-1) - g'(1)]}{12} h^2 + \frac{g^{(4)}(\eta)\,h^4}{360},$$

where $-1 < \eta < 1$. Thus, in this case, $C_1' = (g'(-1) - g'(1))/12$, $m = 2$, and the term $g^{(4)}(\eta)\,h^4/360$ is clearly of the form $O(h^4) = O(h^{m+P})$.

LEMMA 4. *For each $x \in S$, $F_0(x) - F_h(x) = h^P K_1 + O(h^{P+m})$, if $f \in C^{(m+P)}[-1, 1]$ where $K_1 \in E^N$ is independent of h, m and P are as above, and $O(h^{P+m})$ denotes a vector in E^N such that $\| O(h^{m+P})\|/h^{m+P} \leqslant C < \infty$ for all h sufficiently small and positive. (Here $\|\cdot\|$ is an arbitrary norm on E^N.)*

Proof. $(F_0(x) - F_h(x))_i = [A(x) - f, \partial A/\partial x_i(x)] - [A(x) - f, \partial A/\partial x_i(x)]_M$, $1 \leqslant i \leqslant N$, and by assumption each such component of $F_0(x) - F_h(x)$ is of the form $k_i h^P + O(h^{m+P})$ where k_i is a real constant independent of h for $i = 1, 2,..., N$. Thus, $F_0(x) = F_h(x) = h^P K_1 + O(h^{m+P})$, where $K_1 = (k_1, ..., k_N)^T$ is independent of h. ∎

LEMMA 5. *Assume $f \in C^{(P)}[-1, 1]$ and let q be an arbitrary nonnegative integer. Then as $h \to 0$ $F_h \to F$ in the topology of X_q and in fact $N_q(F_h - F) = O(h^P)$.*

Proof. The proof follows from the observation that for any j and any nonnegative integers $i_1, ..., i_N$ we have that

$$\frac{d^j}{dt^j}\left(\frac{\partial^L A}{\partial x_1^{i_1} \cdots \partial x_N^{i_N}}(x)\right)(t)$$

is uniformly bounded on $\bar{U} \times [-1, 1]$ by assumption 1, where $L = i_1 + \cdots + i_N$. Then, for example, in bounding $\sup_{x\in\bar{U}} \| F_0'(x) - F_h'(x)\|$ we must bound the entries of the matrix

$$L(h) = \left(\left[A(x) - f, \frac{\partial^2 A}{\partial x_i\,\partial x_j}(x)\right] + \left[\frac{\partial A}{\partial x_i}(x), \frac{\partial A}{\partial x_j}(x)\right]\right.$$
$$\left. - \left[A(x) - f, \frac{\partial^2 A}{\partial x_i\,\partial x_j}(x)\right]_M - \left[\frac{\partial A}{\partial x_i}(x), \frac{\partial A}{\partial x_j}(x)\right]_M\right),$$

where $1 \leqslant i, j \leqslant N$. But since $f \in C^{(P)}[-1, 1]$, it is evident that the magnitude of each entry of $L(h)$ has an upper bound of the form $C_{ij}h^P$ where C_{ij} is independent of h since the respective Pth derivatives (with respect to t) are uniformly bounded on $\bar{U} \times [-1, 1]$. Proceeding in the same way with the other derivatives (the elementary but lengthy details we will not record here) the bound $N_q(F_h - F) = O(h^P)$ is obtained. ∎

We now have the following theorem which is the main result of this section.

THEOREM 2. *Let $h = 2/M$ and let the quadrature rule $I_h(g) = \sum_{j=0}^{M} \alpha_{jM} g(t_{jM})$ and the positive integers m and P be as before. Assume $f \in C^{(m+P)}[-1, 1]$. Then for all h sufficiently small, $x(F_h) = x(F_0) + h^P C_1 + O(h^L)$ where C_1 is an element of E^N that is independent of h and $L = \min(2P, m + P)$.*

Proof. Let $q \geq 2$ be arbitrary. By Lemma 5, $F_h \to F_0$ in the topology of X_q and by Corollary 2, $x(\cdot)$ is q times continuously differentiable on some ball in X_q centered at F_0. Thus in particular, $x(F_h) = x(F_0) + x'(F_0)(F_h - F_0) + O(N_q(F_h - F_0)^2) = x(F_0) + F_{-1}^1(x_0)(F_h(x_0) - F_0(x_0)) + O(N_q(F_h - F_0)^2)$ (where $F_{-1}^1(x_0) = (F_0^1(x_0))^{-1}) = x(F_0) + F_{-1}^1(x_0)(K_1 h^P + O(h^{m+P})) + O(h^{2P}) = x(F_0) + h^P F_{-1}^1(x_0)\, K_1 + F_{-1}^1(x_0)\, O(h^{m+P}) + O(h^{2P}) = x(F_0) + h^P C_1 + O(h^L)$ where $C_1 = F_{-1}^1(x_0)\, K_1$. Note also that in the above $O(h^{m+P})$ and $O(h^L)$ are vectors in E^N.

Remark 3. The standard way of discretizing the continuous problem is to minimize $\sum_{j=0}^{M} [A(x)(t_{jM}) - f(t_{jM})]^2 (1/2M)$, where the t_{jM}'s are equally spaced. This, however, corresponds to a quadrature formula of the type above with $p = 1$ (actually it is the so-called rectangle rule with one end point added), while if one uses the composite trapezoid rule, say, then $p = 2$. In addition, other standard methods could be employed such as Simpson's rule, still higher-order Newton–Cotes formulas, or Romberg integration. Theorem 2 shows that in such cases Richardson extrapolation could be used to accelerate the convergence of the discrete problem coefficients to the coefficients of the solution to the continuous problem.

Another obvious discretization method that could be used would be the Gaussian quadrature rules. Here the data would not be equally spaced, in general, but the high precision of such formulas should make them especially useful. Clearly, a considerable amount of testing is needed to determine the extent to which the results obtained here can be used to lessen the work needed to solve practical problems. We hope to report elsewhere on the results of such experiments.

Remark 4. In the last section of this paper we have frequently assumed that the linear map $F'(x)$ was positive definite at the local best approximation being considered. If this condition is dropped, then existence of discrete best approximations can still be proved for sufficiently dense discrete subsets (in the l_2 case, say) for such families as the ordinary rational functions and the exponential family [3, 4]. Also, the convergence can be shown to be uniform over the entire interval if the best approximation is unique. However, no information about rate of convergence is obtained once the nonsingularity hypothesis on $F'(x)$ is dropped.

REFERENCES

1. C. R. HOBBY AND J. R. RICE, Approximation from a curve of functions, *Arch. Rat. Mech. Anal.* **24** (1967), 91–106.
2. L. H. LOOMIS AND S. STERNBERG, "Advanced Calculus," Addison-Wesley, Reading, Mass., 1968.
3. J. M. WOLFE, Discrete rational L_p-approximation. *Math. Comput.*, in press.
4. H. L. LOEB AND J. M. WOLFE, Discrete nonlinear approximation, *J. Approximation Theory* **4** (1973), 265–385.
5. E. ISAACSON AND H. B. KELLER, "Analysis of Numerical Methods," Wiley, New York, 1966.