JOURNAL OF APPROXIMATION THEORY 6, 291–300 (1972)

Nonlinear Tchebycheff Approximation with Constraints*

G. A. GISLASON

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

AND

H. L. LOEB

Department of Mathematics, University of Alaska, Fairbanks, Alaska

Communicated by E. W. Cheney

Received June 16, 1970

DEDICATED TO PROFESSOR J. L. WALSH ON THE OCCASION OF HIS 75TH BIRTHDAY

INTRODUCTION

In this paper we discuss a problem of nonlinear Tchebycheff approximation in which the approximating functions are required to satisfy nonlinear side conditions. In our study we remove a hypothesis employed by Hoffmann [5]. (In this connection note our Lemma 2.) Thus, in one application we are able to apply our results to osculatory interpolation using such families as exponentials [6]. Indeed, the analysis of Perrie [1] shows that our results also apply to ordinary rational functions.

BASIC RESULTS

Let Ω be an open set in real n-space, R^n, where for each $A = (a_1, ..., a_n)$ in R^n we define

$$\| A \| = \max_{1 \leqslant i \leqslant n} | a_i |.$$

To each $A \in \Omega$ we assign a continuous real-valued function $F(A, x)$, $x \in [0, 1]$, such that the partial derivatives

$$\partial F(A, x)/\partial a_i, \qquad i = 1, ..., n$$

exist and are continuous in A and x. We let

$$W(A) = \left\{ \sum_{i=1}^{n} c_i \frac{\partial F(A, x)}{\partial a_i} : c_i \text{ real} \right\}$$

* Supported in part by N.S.F. Grant GP-18609.

and denote by $d(A)$ the dimension of $W(A)$ as a subspace of $C[0, 1]$; we assume $d(A) > 0$. Finally, set

$$V = \{F(A, x)\colon A \in \Omega\}.$$

When appropriate we will let $f^{(j)}(x)$ and $F^{(j)}(A, x)$ denote, respectively, the j-th derivative of $f(x)$ and $F(A, x)$ with respect to x.

Approximation will always be in the Tchebycheff norm:

$$\|f\| = \sup_{x\in[0,1]} |f(x)| \qquad \text{for} \quad f \in C[0, 1].$$

We consider the problem of approximating a function $g \in C[0, 1]$ in the Tchebycheff norm by functions $F(A, x) \in V$ which satisfy certain side conditions

$$N_i(A) = 0, \qquad i = 1,..., k.$$

Let $V' = \{F(A, x) \in V\colon N_i(A) = 0,\ i = 1,..., k\}$ and assume V' is non-empty. A function $F(A^*, x) \in V'$ is said to be a best approximation to g if and only if

$$\| g - F(A^*)\| = \inf\{\| g - F(A)\|\colon F(A, x) \in V'\}.$$

The following assumptions are made on the family V and the side conditions $N_1 ,..., N_k$.

(1) For each $A \in \Omega$, $W(A)$ is a Haar subspace of $C[0, 1]$ of dimension $d(A)$; that is, every nonzero element in $W(A)$ has at most $d(A) - 1$ zeros, where $d(A) > k$.

(2) For each $A \in \Omega$ and some basis for $W(A)$, which we assume without loss of generality is

$$\{\partial F(A, x)/\partial a_1 ,..., \partial F(A, x)/\partial a_{d(A)}\},$$

the matrix

$$\left(\frac{\partial N_i(A)}{\partial a_j} : \begin{matrix} i = 1,..., k \\ j = 1,..., d(A)\end{matrix}\right)$$

has rank k. Such a basis will be called a canonical basis.

LEMMA 1. *Let $F(A^*, x) \in V$. Then there exist $q = d(A^*) - k$ points $0 \leqslant x_1 < \cdots < x_q \leqslant 1$ such that for each $\epsilon > 0$ there is a $\delta > 0$, depending on A^*, x_i, and ϵ, for which $0 \leqslant x_1' < \cdots < x_q' \leqslant 1$ and*

$$\max_{\substack{1\leqslant i\leqslant q \\ 1\leqslant j\leqslant k}} \{1x_i' - x_i |, | e_{i+j} |\} \leqslant \delta$$

imply that a vector $A' \in \Omega$ can be found which satisfies

$$F(A', x_i') = F(A^*, x_i') + e_i, \qquad i = 1,\dots, q,$$
$$N_j(A') = e_{q+j}, \qquad j = 1,\dots, k,$$
$$A' = (a_1',\dots, a_d', a_{d+1}^*,\dots, a_n^*),$$

and

$$\| A' - A^* \| \leqslant \epsilon.$$

Here, $d = d(A^*)$, $A^* = (a_1^*,\dots, a_n^*)$.

(Note that we assume that the first d components of the gradient vector for $F(A^*, x)$ form a canonical basis for $W(A^*)$.)

Proof. Let $0 \leqslant x_1 < \cdots < x_d \leqslant 1$. From (1) and (2) we may assume without loss of generality that

$$\det \left(\frac{\partial F(A^*, x_i)}{\partial a_j} : \begin{matrix} i = 1,\dots, d \\ j = 1,\dots, d \end{matrix}\right) \neq 0,$$

and

$$\operatorname{rank} \left(\frac{\partial N_i(A^*)}{\partial a_j} : \begin{matrix} i = 1,\dots, k \\ j = 1,\dots, d \end{matrix}\right) = k.$$

Hence, there exists $q = d - k$ of the x_i, which we label as $0 \leqslant x_1 < \cdots < x^q \leqslant 1$, such that

$$\det \begin{pmatrix} \dfrac{\partial F(A^*, x_i)}{\partial a_j} : \begin{matrix} i = 1,\dots, q \\ j = 1,\dots, d \end{matrix} \\ \cdots \\ \dfrac{\partial N_i(A^*)}{\partial a_j} : \begin{matrix} i = 1,\dots, k \\ j = 1,\dots, d \end{matrix} \end{pmatrix} \neq 0. \tag{3}$$

The result now follows by applying the Implicit Function Theorem to the functions:

$$f_i(a_1,\dots, a_d, x_1',\dots, x_q', e_1,\dots, e_d).$$
$$\equiv F(a_1,\dots, a_d, a_{d+1}^*,\dots, a_n^*, x_i') - e_i, \; i = 1,\dots, q,$$
$$f_{q+j}(a_1,\dots, a_d, x_1',\dots, x_q', e_1,\dots, e_d)$$
$$\equiv N(a_1,\dots, a_d, a_{d+1}^*,\dots, a_n^*) - e_{q+j}, \; j = 1,\dots, k,$$

where $A = (a_1,\dots, a_d, a_{d+1}^*,\dots, a_n^*)$. ∎

Hereafter, for the function in $C[0, 1]$ denoted by $g(x)$ and the function in V denoted by $F(A^*, x)$ we agree that

$$\sigma(x) = \operatorname{sgn}(g(x) - F(A^*, x)).$$

$$X = \{x \in [0, 1] : |g(x) - F(A^*, x)| = \| g - F(A^*)\|\},$$

$$\tilde{N}(A^*) = \{C \in R^n : C = (c_1, ..., c_d, 0, ..., 0)\}$$

$$\text{and} \qquad \left\{\sum_{s=1}^{d} c_s \frac{\partial N_i(A^*)}{\partial a_s} = 0,\ i = 1, ..., k\right\},$$

$$N(A^*) = \left\{\sum_{s=1}^{d} c_s \frac{\partial F(A^*, x)}{\partial a_s} : C \in \tilde{N}(A^*)\right\}.$$

Further, $d = d(A^*)$ and the first d components of the gradient vector form a canonical base. To avoid trivialities we assume $g \notin V'$.

LEMMA 2. *Assume that $F(A^*, x) \in V'$. Then for each $C \in \tilde{N}(A^*)$ with $\| C \| = 1$ there exists a sequence $(A^{(\nu)})_{\nu=1}^{\infty} \subset \Omega - \{A^*\}$ such that*

$$\| A^{(\nu)} - A^* \| \to 0$$

$$F(A^{(\nu)}, x) \in V', \qquad \nu = 1, 2, ...$$

and

$$\frac{A^{(\nu)} - A^*}{\| A^{(\nu)} - A^* \|} \to C.$$

Proof. Let $C \in \tilde{N}(A^*)$ be as in the hypotheses and define

$$h(x) \equiv \sum_{i=1}^{d} c_i \frac{\partial F(A^*, x)}{\partial a_i} \qquad x \in [0, 1].$$

Clearly, $h \neq 0$. By (1) there exist points $0 \leqslant x_1 < \cdots < x_d \leqslant 1$ such that $h(x_i) \neq 0$ for $1 \leqslant i \leqslant d$. Now, (3) holds for $q = d(A^*) - k$ of the x_i, say, $0 \leqslant x_1 < \cdots < x^q \leqslant 1$. Let

$$e_i \equiv h(x_i) \neq 0, \qquad i = 1, ..., q.$$

Let $(t^{(\nu)})_{\nu=1}^{\infty}$ be a sequence of positive numbers such that $t^{(\nu)} \to 0$. For some subsequence of $(t^{(\nu)})_{\nu=1}^{\infty}$, which we do not relabel, there exists by Lemma 1 a sequence $(A^{(\nu)})_{\nu=1}^{\infty} \subset \Omega$ with the properties that

$$F(A^{(\nu)}, x_i) = F(A^*, x_i) + t^{(\nu)} e_i, \qquad i = 1, ..., q,$$

$$N_j(A^{(\nu)}) = 0, \qquad j = 1, ..., k,$$

$$A^{(\nu)} = (a_1^{(\nu)}, ..., a_d^{(\nu)}, a_{d+1}^*, ..., a_n^*),$$

and

$$\| A^{(\nu)} - A^* \| \to 0.$$

By the Mean Value Theorem,

$$F(A^{(\nu)}, x_i) - F(A^*, x_i) = \sum_{s=1}^{d} (a_s^{(\nu)} - a_s^*) \frac{\partial F(A_i^{(\nu)}, x_i)}{\partial a_s}$$

$$= t^{(\nu)} e_i, \qquad i = 1, \ldots, q, \tag{4}$$

$$0 = N_j(A^{(\nu)}) - N_j(A^*)$$

$$= \sum_{s=1}^{d} (a_s^{(\nu)} - a_s^*) \frac{\partial N_j(A_j^{(\nu)})}{\partial a_s}, \qquad j = 1, \ldots, k, \tag{5}$$

where the vectors $A_i^{(\nu)}$ and $A_j^{(\nu)}$ are on the line between A^* and $A^{(\nu)}$. Since $A^{(\nu)} \neq A^*$ for each ν, we may assume that

$$\frac{A^{(\nu)} - A^*}{\| A^{(\nu)} - A^* \|} \to \hat{C} \in R^n,$$

where $\hat{C} = (\hat{c}_1, \ldots, \hat{c}_d, 0, \ldots, 0)$ and $\| \hat{C} \| = 1$. Dividing by $\| A^{(\nu)} - A^* \|$ in (4) and (5) and letting $\nu \to \infty$ implies that

$$\sum_{s=1}^{d} \hat{c}_s \frac{\partial F(A^*, x_i)}{\partial a_s} = e_i \cdot \lim_{\nu \to \infty} \frac{t^{(\nu)}}{\| A^{(\nu)} - A^* \|}, \qquad i = 1, \ldots, q, \tag{6}$$

and

$$\sum_{s=1}^{d} \hat{c}_s \frac{\partial N_j(A^*)}{\partial a_s} = 0, \qquad j = 1, \ldots, k. \tag{7}$$

By (3), (7), and the structure of $\hat{C}$ it follows that for at least one index i the left (and right) side of (6) is not zero. Hence

$$\alpha = \lim_{\nu \to \infty} [t^{(\nu)} / \| A^{(\nu)} - A^* \|] > 0.$$

It follows that $\hat{C} = \alpha C$. Since $\| \hat{C} \| = \| C \| = 1$ we must have $\alpha = 1$ and $\hat{C} = C$. ∎

LEMMA 3. *Assume that $F(A^*, x)$ is a best approximation to $g(x)$ and that $k < d(A^*)$. Then every $h \in N(A^*)$ satisfies*

$$\min_{x \in X} \sigma(x)\, h(x) \leqslant 0.$$

Proof. Suppose to the contrary that there exists some nonzero $h^* \in N(A^*)$ for which

$$\min_{x \in X} \sigma(x)\, h^*(x) > 0.$$

Clearly, we may assume that $\| C^* \| = 1$, where $C^* \in \tilde{N}(A^*)$ and

$$h^*(x) \equiv \sum_{s=1}^{d} c_s{}^* \frac{\partial F(A^*, x)}{\partial a_s}.$$

Let

$$\epsilon = \min_{x \in X} \sigma(x)\, h^*(x) > 0.$$

By continuity and the fact that Ω is open there exists a $\delta > 0$ such that

$$\| C - C^* \| \leqslant \delta \qquad \text{and} \qquad \| A - A^* \| \leqslant \delta$$

imply that $A \in \Omega$ and

$$\min_{x \in X} \sigma(x) \sum_{i=1}^{d} c_i \frac{\partial F(A, x)}{\partial a_i} \geqslant \epsilon/2.$$

Without loss of generality we may assume that $\| g - F(A^*)\| = 1$. For $\delta > 0$ as above, let

$$Z = \left\{ x \in [0, 1] : (g(x) - F(A^*, x)) \sum_{i=1}^{d} c_i \frac{\partial F(A, x)}{\partial a_i} \leqslant \epsilon/3 \right\},$$

whenever $\| C - C^* \| \leqslant \delta$ and $\| A - A^* \| \leqslant \delta$.

Clearly, $X \cap Z = \varnothing$. Hence, $\theta = \sup_{x \in Z} | g(x) - F(A^*, x)| < \| g - F(A^*)\|$. Let

$$E = \sup \left\{ \left| \sum_{i=1}^{n} c_i \frac{\partial F(A, x)}{\partial a_i} \right| : \| C - C^* \| \leqslant \delta,\ \| A - A^* \| \leqslant \delta,\ x \in [0, 1] \right\}.$$

Now, let $(A^{(\nu)})_{\nu=1}^{\infty}$ be a sequence corresponding to C^* as in Lemma 2. Choose ν so large that $\| A^{(\nu)} - A^* \| \leqslant \delta$, $\| C^{(\nu)} - C^* \| \leqslant \delta$ and $\| A^{(\nu)} - A^* \| < \min\{(\| g - F(A^*)\| - \theta)/E, \epsilon/3E^2\}$.

Consider a fixed $x \in [0, 1]$. We have by the Mean Value Theorem that

$$\begin{aligned} F(A^{(\nu)}, x) - F(A^*, x) &= \sum_{r=1}^{d} (a_r^{(\nu)} - a_r{}^*) \frac{\partial F(A^{(\nu)}(x), x)}{\partial a_r} \\ &= \| A^{(\nu)} - A^* \| \cdot h_\nu(x), \end{aligned} \tag{8}$$

where

$$C^{(\nu)} = (c_1^{(\nu)},\ldots, c_d^{(\nu)}, 0,\ldots, 0) = (A^{(\nu)} - A^*)/\| A^{(\nu)} - A^* \|,$$

$$h_\nu(x) \equiv \sum_{r=1}^{d} c_r^{(\nu)} \frac{\partial F(A^{(\nu)}(x), x)}{\partial a_r}$$

and $A^{(\nu)}(x)$ is on the line between A^* and $A^{(\nu)}$.

Assume that $x \in Z$. From (8) and the definitions we have

$$\begin{aligned} | g(x) - F(A^{(\nu)}, x)| &\leqslant | g(x) - F(A^*, x)| + \| A^{(\nu)} - A^* \| \cdot | h_\nu(x)| \\ &< \theta + (\| g - F(A^*)\| - \theta). \end{aligned}$$

Hence, $| g(x) - F(A^*, x)| < \| g - F(A^*)\|$ for $x \in Z$.

Now, suppose that $x \notin Z$. Then by (8) it follows that

$$\begin{aligned} | g(x) - F(A^{(\nu)}, x)|^2 &= | g(x) - F(A^*, x) - \| A^{(\nu)} - A^* \| \cdot h_\nu(x)|^2 \\ &= | g(x) - F(A^*, x)|^2 - 2 \| A^{(\nu)} - A^* \| \cdot (g(x) \\ &\quad - F(A^*, x)) \cdot h_\nu(x) + \| A^{(\nu)} - A^* \|^2 \cdot | h_\nu(x)|^2 \\ &\leqslant \| g - F(A^*)\|^2 - 2\| A^{(\nu)} - A^* \| \cdot (\epsilon/3) \\ &\quad + \| A^{(\nu)} - A^* \| \cdot (\epsilon/3). \end{aligned}$$

Hence, $| g(x) - F(A^{(\nu)}, x)| < \| g - F(A^*)\|$ for $x \notin Z$. By continuity and compactness it follows that $\| g - F(A^{(\nu)})\| < \| g - F(A^*)\|$. ▮

THEOREM 4. *Same hypothesis as Lemma* 3. *Then the origin of real q-space, where* $q = d(A^*) - k = \dim N(A^*)$, *lies in the convex hull of the set* $\{\sigma(x)\hat{x}\colon x \in X\}$ *where* $\hat{x} = (h_1(x),\ldots, h_q(x))$ *and* $\{h_1,\ldots, h_q\}$ *is a basis for* $N(A^*)$.

Proof. This is merely a restatement of Lemma 3. A proof is given by Cheney, [2].

APPLICATIONS

We now give some applications of these results.

Let $0 \leqslant y_1 < \cdots < y_p \leqslant 1$ be p distinct points and $m_1,\ldots, m_p$ be positive integers with $m \equiv \max m_i$. Assume that V satisfies the following condition:

(9) For each $A \in \Omega$ $w(A)$ is an extended Haar subspace of $C^{m-1}[0, 1]$ of order m; that is, each nonzero element of $w(A)$ has at most $d(A) - 1$ zeros counting multiplicities up to order m.

We now consiqer the problem of approximating a given function $g \in C^{m-1}[0, 1]$ by functions $F(A, x)$ in V which satisfy the side conditions

$$N_{ij}(A) \equiv F^{(j)}(A, y_i) - g^{(j)}(y_i) \qquad \begin{matrix} i = 1,\dots, p \\ j = 0,\dots, m_i - 1 \end{matrix} \qquad (10)$$
$$= 0.$$

It is easy to see that (1) and (2) hold in this case. Also, for each $A \in \Omega$, we have

$$N(A) = \left\{ h \in w(A) : h^{(j)}(y_i) = 0 \qquad \begin{matrix} i = 1,\dots, p \\ j = 0,\dots, m_i - 1 \end{matrix} \right\}.$$

THEOREM 5. *If $F(A^*, x)$ is a best approximation to g and*

$$q = d(A^*) - \sum_{i=1}^{p} m_i \geqslant 1$$

then there exist $q + 1$ points $0 \leqslant x_1 < \cdots < x_{q+1} \leqslant 1$ such that

$$\sigma(x_{r+1}) = (-1)^{1+k(r)}\sigma(x_r), \qquad r = 1,\dots, q,$$

and

$$| g(x_r) - F(A^*, x_r)| = \| g - F(A^*)\| \qquad r = 1,\dots, q + 1,$$

where $k(r)$ denotes the sum of the multiplicities m_i for which $x_r < y_i < x_{r+1}$.

Proof. The conclusion follows from Theorem 4 and Theorem 3.1 of the results of Loeb *et al.* [3]. ▌

Now suppose that V has the following properties.

(11) For each $A \in \Omega$, $w(A)$ is an extended Haar subspace of $C^m[0, 1]$ of order $m + 1$.

(12) If $A, A' \in \Omega$, then $F(A, x) - F(A', x)$ can have at most $d(A) - 1$ zeros counting multiplicities up to order $m + 1$, or else $F(A, x) \equiv F(A', x)$.

THEOREM 6. *$F(A^*, x)$ is a best approximation to g if and only if there exist $q + 1$ points $0 \leqslant x_1 < \cdots < x_{q+1} \leqslant 1$, where $q = d(A^*) - \sum_{i=1}^{p} m_i$, such that*

$$\sigma(x_{r+1}) = (-1)^{1+k(r)}\, \sigma(x_r), \qquad r = 1,\dots, q,$$

and

$$| g(x_r) - F(A^*, x_r)| = \| g - F(A^*)\|, \qquad r = 1,\dots, q + 1.$$

In addition, there is at most one best approximation.

Proof. Clearly, we may assume that $q < 0$, since otherwise, by (12), there could be at most one function $F(A, x) \in V$ which satisfies (10). The "uniqueness" and the sufficiency of alternation are easy consequences of (12). Theorem 5 implies the necessity of alternation. ∎

We now consider another application. Suppose

(13) If $F(A, x)$ and $F(A^*, x)$ are in V' and are not identical then they can agree on at most $q - 1$ points where $q \equiv d(A^*) - k$ points.

(14) For $A^* \in \Omega$ and any q distinct points, $0 \leqslant x_1 < x_2 < \cdots < x_q \leqslant 1$, (3) is satisfied.

THEOREM 7. *$F(A^*, x)$ is the best approximation to $g \in C[0, 1]$ if and only if there exist $q + 1$ points, $0 \leqslant x_1 < x_2 < \cdots < x_{q+1} \leqslant 1$, where $q = d(A^*) - k$ such that*

$$| g(x_1) - F(A^*, x_1)| = \| g - F(A^*)\|,$$

$$g(x_i) - F(A^*, x_i) = -(g(x_{i+1}) - F(A^*, x_{i+1})), \qquad i = 1,..., q.$$

Furthermore, there is at most one best approximation.

Proof. We outline the proof. The necessity of the alternations follows from the fact that $N(A^*)$ is a Haar Subspace of dimension $d(A^*) - k$, Theorem 4, and standard results concerning the convex hull. (For example, see [2, p. 64]). The sufficiency of the alternations follows from (13). "Uniqueness" is a consequence of (13), (14), and "Implicit Function Theorem". ∎

We note that the hypotheses of Theorems 6 and 7 could be merged to give other variants.

For our last application we examine the behavior of the best approximation operator.

Again, let V satisfy (11) and (12). The proofs of the next three results require only slight modification of the proofs given by Barrar and Loeb [4] for the corresponding results without interpolation.

We assume that $F(A^*, x)$ is the best approximation to $g \in C^{m-1}[0, 1]$ which satisfies (10), and that

$$d(A^*) = \max_{A \in \Omega} d(A).$$

Also, we let $T(f)$ denote the best approximation to $f \in C^{m-1}[0, 1]$ from V which satisfies (10), if the best approximation exists. Let $V(g)$ be the set of elements of V which satisfy (10).

THEOREM 7. *There exists a number* $\alpha > 0$ *such that if* $F(A, x) \in V(g)$ *then*

$$\| g - F(A)\| \geq \| g - F(A^*)\| + \alpha \| F(A) - F(A^*)\|. \quad \blacksquare$$

THEOREM 8. *There is a number* $\lambda > 0$ *such that if* $f \in C^{m-1}[0, 1]$,

$$f^{(j)}(y_i) = g^{(j)}(y_i), \qquad \begin{matrix} i = 1,..., p, \\ j = 0,..., m_i - 1, \end{matrix} \tag{15}$$

and $T(f)$ *exists, then*

$$\| T(g) - T(f)\| \leq \lambda \| g - f\|.$$

THEOREM 9. *There is a* $\delta > 0$ *such that for any* $f \in C^{m-1}[0, 1]$ *which satisfies* (15) *and has the property*

$$\|f - g\| \leq \delta,$$

$T(f)$ *exists.*

REFERENCES

1. A. L. PERRIE, "Uniform Rational Approximation with Osculatory Interpolation," *J. Comput. System Sci.*, to appear.
2. E. W. CHENEY, "Introduction to Approximation Theory," McGraw-Hill, New York, 1966.
3. H. L. LOEB, D. G. MOURSUND, L. L. SCHUMAKER, AND G. D. TAYLOR, "Uniform Generalized Weight Function Polynomial Approximation with Interpolation," *SIAM J. Numer. Anal.* **6** (1969), 284–293.
4. R. B. BARRAR AND H. L. LOEB, On the continuity of the nonlinear Tchebycheff operator, *Pacific J. Math.* **32** (1970), 593–601.
5. K. H. HOFFMANN, Zur Theorie der nichlinearen Tschebyscheff-Approximation mit Nebenbedingungen, *Numer. Math.* **14** (1969), 24–41.
6. S. KARLIN, "Total Positivity," Vol. 1, Stanford University Press, Stanford, Ca, 1968.