JOURNAL OF APPROXIMATION THEORY 2, 79–84 (1969)

Nonlinear Chebychev Approximation by Vector-Norms

ALEX BACOPOULOS

Michigan State University, East Lansing, Michigan 48823

1.

Two well-known nonlinear approximating families in the context of Chebychev approximation are the unisolvent functions and the rational functions. In 1961 John Rice [5] defined the notion of unisolvence of variable degree. He developed what is perhaps today the most general theory of nonlinear Chebychev-type approximation on an interval, which includes both unisolvent and rational approximations.

In this paper, a given function is approximated by unisolvent functions of variable degree, simultaneously with respect to *several* weight functions. A notion of vector-alteration is defined, which permits a characterization of best approximations along the lines of the standard Chebychev theory, and which generalizes the above results of Rice.

2.

Let $f(x)$ be a continuous function to be approximated on $[a,b]$; let P be a non-empty subset of Euclidean n-space E_n, let $\{F(A,x): A \in P\}$ be the class of approximating functions, and let $\{w_s(x)\}$ be k continuous, positive (weight) functions on $[a,b]$, $s = 1, 2, \ldots, k$. For any function $g(x)$ define:

$$\|g(\cdot)\| = \sup\{|g(x)| : a \leqslant x \leqslant b\}.$$

Define a vector-valued function:

$$G(A) = (\|w_1(\cdot)[f(\cdot) - F(A,\cdot)]\|, \|w_2(\cdot)[f(\cdot) - F(A,\cdot)]\|, \ldots, \|w_k(\cdot)[f(\cdot) - F(A,\cdot)]\|).$$

A is said to be a *better-than-or-equal* approximation to B, if and only if,

$$\|w_s(\cdot)[f(\cdot) - F(A,\cdot)]\| \leqslant \|w_s(\cdot)[f(\cdot) - F(B,\cdot)]\|$$

for each s, $s = 1, 2, \ldots, k$. We shall denote this by $G(A) \leqslant \cdot G(B)$. A point p of a subset S of E_k will be called a *minimal point* of S, if and only if, there is no $q \neq p$, $q \in S$, with the property $q \leqslant p$. If $A^* \in P$, then A^* (or $F(A^*,x)$) is said to be a *best approximation*, if and only if, $G(A^*)$ is a minimal point of the set

$\{G(A):A \in P\}$. Observe that if A^* is a best approximation to f with respect of any one of the k weight functions, then A^* is a best approximation because of its uniqueness. The partial ordering $\leqslant\cdot$ having been defined, the symbols $\cdot\geqslant$, $<\cdot$, $\cdot>$ are understood as usual.

The *Problem of Chebychev Approximation by Vector-Norms* may be stated as follows: Examine for existence, uniqueness and characterization the A in $P = E_n$ which are best approximations. The following concepts will be used to restrict the class of approximating functions.

The real-valued function $F(A,x)$ is defined for x in $[a,b]$, and A in E_n. It will be assumed to be continuous in x and A. $F(A,x)$ is said to be *solvent of degree m at $A^* \in E_n$*, if given a sequence $\{x_j : a \leqslant x_1 < x_2 < \ldots < x_m \leqslant b\}$ and $\epsilon > 0$, there exists a $\delta(A^*, \epsilon, x_3, x_2, \ldots, x_m) > 0$ such that $|y_j - F(A^*, x_j)| < \delta$ implies that there is a solution, $A \in E_n$, to $F(A, x_j) = y_j$, $j = 1, 2, \ldots, m$, with $\max_{a \leqslant x \leqslant b} |F(A,x) - F(A^*,x)| < \epsilon$. A family of functions $\{F(A,x): A \in E_n\}$ is said to satisfy *the density condition*, if and only if, given $A \in E_n$ and any $\epsilon > 0$, there exist vectors B and C in E_n such that

$$F(A,\cdot) - \epsilon < F(B,\cdot) < F(A,\cdot) < F(C,\cdot) < F(A,\cdot) + \epsilon.$$

$F(A,x)$ is said to be *unisolvent of degree m* at $A^* \in E_n$, if

(i) $F(A,x)$ is solvent of degree m at A^*; (ii) $F(A,x)$ is not solvent of degree $m+1$ at A^*; (iii) for any $A \neq A^*$, $F(A^*,x) - F(A,x)$ has at most $m-1$ zeros on $[a,b]$.

Denote by $V(A,x)$ the vector-valued function

$$(w_1(x)[f(x) - F(A,x)],\ w_2(x)[f(x) - F(A,x)],\ \ldots,\ w_k(x)[f(x) - F(A,x)]).$$

Given an A in E_n, a point x_0 of $[a,b]$ will be called a *positive vector-extremum* of $V(A,x)$, if for some s, $1 \leqslant s \leqslant k$,

$$w_s(x_0)[f(x_0) - F(A,x_0)] = \|w_s(f - F(A,\cdot))\|;$$

similarly, x_0 is called a *negative vector-extremum*, if for some s, $1 \leqslant s \leqslant k$,

$$w_s(x_0)[f(x_0) - F(A,x_0)] = -\|w_s(f - F(A,\cdot))\|.$$

The "error curve" $V(A,x)$ is said to *vector-alternate n* times on $[a,b]$, if there are $n+1$ points $x_1 < x_2 < \ldots < x_{n+1}$ on $[a,b]$ such that $x_1, x_2, \ldots, x_{n+1}$ are, alternatively, positive and negative vector-extrema of $V(A,x)$.

Given K, a subset of E_n, we denote by $M(K)$ the minimal set of $\{G(A): A \in K\}$, i.e., $M(K) = \{G(A): A$ is a best approximation in $K\}$. For notational convenience we write M in place of $M(E_n)$.

3.

The existence of best approximations here is understood in the context of the above partial order, $\leqslant\cdot$. The set of infima of the descending chains of

$\{G(A):A \in E_n\}$ has, in general, the cardinality of the continuum. Therefore, although there may exist best $(\leqslant\cdot)$ approximations, e.g., each B_s for which

$$\|w_s(\cdot)(f(\cdot)-F(B_s,\cdot))\| = \inf_{A\in P} \|w_s(\cdot)(f(\cdot)-F(A,\cdot))\|, \text{ for some } s,$$

it is *a priori* possible that there may also exist descending chains in $\{(G(A):A \in E_n)\}$ whose infima are not attained. It is also possible that one existing best approximation is unique while another is not (see Theorem 4). In what follows some conditions for existence are given. It will be shown later that each best approximation is unique.

THEOREM 1. *Let $\{F(A,x):A \in E_n\}$ be a unisolvent family of functions of degree n on $[a,b]$. Then if μ is the infimum of any chain in $\{G(A):A \in E_n\}$, there exists $A^* \in E_n$ such that $G(A^*)=\mu$.*

This theorem is a direct generalization of Theorem 5 of [*6*]. Its proof will be omitted.

THEOREM 2. *Let $\{F(A,x):A \in E_n\}$ be a unisolvent linear (in A) family of functions of degree n on $[a,b]$, and let $k=2$. Then, the minimal set M is a Jordan arc, if and only if, $B_1 \neq B_2$. If $B_1 = B_2$, M is a point.*

Proof. If $B_1 = B_2$, it is clear that M is a point. Assume then that $B_1 \neq B_2$ and note that $G(B_1) \neq G(B_2)$. We first prove the theorem for the linear approximating class $\{F(A,x):A \in P\}$, where P is a compact, convex subset of E_n containing B_1 and B_2. The unisolvence will be invoked at the end of the proof.

Denote by L the straight line segment joining $G(B_1)$ and $G(B_2)$. A homeomorphism will be exhibited which maps L onto $M(P)$. For each point p of L let l_p denote the straight line of slope 1 passing through p. Denote by $N(x)$ the subset of E_2 defined by $\{y:y \leqslant\cdot x\}$. The subsets of $N(x)$: $\{y:y\leqslant\cdot x$; second coordinate of $y=$ second coordinate of $x\}$ and $\{y:y\leqslant\cdot x$; first coordinate of $y=$ first coordinate of $x\}$ will be called, respectively, upper face and right face of $N(x)$. Given a point p of L, let $x_0(p)$ be the point on l_p of smallest coordinates such that $N(x_0(p)) \cap \{G(A):A \in P\} \neq 0$. It will be shown that $x_0(p) \in M(P)$. The existence of $x_0(p)$ follows from the fact that $\{G(A):A \in P\}$ is compact and connected. To show that $x_0(p) \in M(P)$, assume that for some $p \in L$, $x_0(p) \notin M(P)$. Without loss of generality it will be assumed that p is neither $G(B_1)$ nor $G(B_2)$. Then there is a point z of $\{G(A):A \in P\}$ which belongs to either the upper face or the right face of $N(x_0(p))$, say the right face. Let $z=(z_1,z_2)$ and $x_0(p)=(x_1,x_2)$. Since $p \neq G(B_1)$, for some $\epsilon > 0$ there exists a smallest nonnegative y_2 satisfying

$$N(x_1-\epsilon,\ x_2+y_2) \cap \{G(A):A \in P\} \neq 0.$$

Let $w=(w_1,w_2)$ be some point of this set. From what has been said it follows that $x_1=z_1$, $x_2>z_2$, and that w belongs to the upper face of $N(x_1-\epsilon,x_2+y_2)$.

Now pick a positive α so that $\alpha < (x_2 - z_2)/(x_2 + y_2 - z_2)$, and let $A_z \in G^{-1}(z)$ and $A_w \in G^{-1}(w)$. We have

$$G(\alpha A_z + (1-\alpha) A_w) \leqslant \cdot\, \alpha z + (1-\alpha) w$$

which implies that there is a point (u_1, u_2) in $N(x_0(p)) \cap \{G(A): A \in P\}$ with $u_1 < x_1$, in contradiction to the way $x_0(p)$ was defined. Therefore $x_0(p) \in M(P)$. Define h by $h(p) = x_0(p)$ for $p \in L$. The continuity of h follows immediately, so a homeomorphism $h: L \to M(P)$ has been exhibited.

Finally, let h_i be homeomorphisms as defined above from L onto $M(S_i)$, where S_i are closed disks with the properties that S_1 contains B_1 and B_2, $S_i \subset S_{i+1}$ and $\bigcup_{i=1}^{\infty} S_i = E_n$. Now, for each $p \in L$, the points $\{h_i(p)\}$ are bounded from "below" and are decreasing on l_p. Let $H(p) = \lim_{i\to\infty} h_i(p)$. It is now clear that the function H on the set L is a homeomorphism onto the set $\{H(p): p \in L\}$, i.e., onto the set of infima of all chains in $\{G(A): A \in E_n\}$. By Theorem 1, $\{H(p): p \in L\} = M$, which concludes the proof.

4.

In this section the characterization and uniqueness of each best approximation is given. Note that, in general, there will be many "unique best approximations".

THEOREM 3. *Let $\{F(A,x): A \in E_n\}$ be a unisolvent family of functions on $[a,b]$ of variable degree. Let this family satisfy the density condition, and denote, for each $A \in E_n$, by $n(A)$ the degree of unisolvence at A. Then $F(A^*,x)$ is a best approximation to a given continuous function $f(x)$ on $[a,b]$, if and only if, $V(A^*,x)$ vector-alternates at least $n(A^*)$ times on $[a,b]$.*

Proof. Assume that $F(A^*,x)$ is a best approximation to $f(x)$ and that $V(A^*,x)$ vector-alternates r times, $r < n(A^*)$. The proof depends on the density condition of $\{F(A,x): A \in E_n\}$ which guarantees that $r \geqslant 1$. It consists of showing that there is a $B \in E_n$ satisfying $G(B) < \cdot\, G(A^*)$. We distinguish several cases depending on whether a and/or b are vector-extrema.

Assume that a is not a vector-extremum of $V(A^*,x)$. For simplicity of notation we shall write n instead of $n(A^*)$. Divide the interval $[a,b]$ into $r+1$ subintervals by the points $x_0 < x_{n-r} < x_{n-r+1} < \ldots < x_n$ where $x_0 = a$, $x_n = b$ and the rest of the points x_j are chosen so that (i) the approximation $F(A^*,x)$ interpolates $f(x)$ at x_j, $j = n-r, n-r+1, \ldots, (n-1)$; (ii) $V(A^*,x)$ vector-alternates once in any two adjacent subintervals, while it does not vector-alternate in any one subinterval. Let

$$D_i = \|w_i(\cdot)(f(\cdot) - F(A^*, \cdot))\|, \qquad i = 1, 2, \ldots, k.$$

By the continuity of F, there is a $\delta > 0$ so that for some $\theta > 0$,

$$|w_s(x)(f(x) - F(A^*, x))| < D_s - \theta, \qquad s = 1, 2, \ldots, k,$$

whenever $x \in [a, a + \delta]$. Now pick $n - r - 1$ points in $[a, a + \delta]$, say $x_1 < x_2 < \ldots < x_{n-r-1}$, and define:

$$M_{ij} = \max_{x_j \leqslant x \leqslant x_{j+1}} (w_i(x)(f(x) - F(A^*, x)))$$

$$m_{ij} = \min_{x_j \leqslant x \leqslant x_{j+1}} (w_i(x)(f(x) - F(A^*, x)))$$

$$d_{ij} = D_i - \frac{M_{ij} - m_{ij}}{2},$$

for $j = n - r - 1, n - r, \ldots, n - 1$; $i = 1, 2, \ldots, k$. Choose some positive vector-extremum x' in $[a, b]$. By the solvency of degree n at A^*, it follows that given any $\epsilon > 0$, there is a $B \in E_n$ with the properties:

(i) $F(B, x_i) = F(A^*, x_i)$, $\qquad i = 1, 2, \ldots, (n - 1)$,

(ii) $F(B, x') = F(A^*, x') + \delta_1$, $\qquad$ for some $\delta_1 \leqslant \epsilon/2$,

(iii) $\|w_i(\cdot)(F(B, \cdot) - F(A^*, \cdot))\| < \epsilon$, $\qquad i = 1, 2, \ldots, k$.

In particular, pick ϵ to be the smallest among $\theta/2$ and d_{ij}, $j = n - r - 1, n - r, \ldots, n - 1$; $i = 1, 2, \ldots, k$. It then follows that (i), (ii) and (iii), together with the unisolvence of $F(A, x)$ imply that $G(B) <\cdot\, G(A^*)$, a contradiction.

In case b is not a vector-extremum and in the case that a and b are both vector-extrema, the proofs are similar to the above.

Conversely, assume that $F(A^*, x)$ vector-alternates at least n times and that there is a $B \in E_n$, $B \neq A$, satisfying $G(B) <\cdot\, G(A^*)$. Then it follows that $F(B, x) - F(A^*, x)$ has at least n zeros on $[a, b]$, contradicting the unisolvence of degree n of $F(A, x)$ at A^*. This establishes the characterization of each best approximation.

THEOREM 4. *Each best approximation of theorem* 3 *is unique*, i.e., *given* $\mu \in M$, *there is only one* $A^* \in E_n$ *such that* $G(A^*) = \mu$.

Proof. Given $\mu \in M$, assume that there are two vectors A_1, $A_2(\neq A_1)$ with the property that $G(A_1) = G(A_2) = \mu$ and with respective degrees of unisolvence $n(A_1)$ and $n(A_2)$. By Theorem 3, the functions $V(A_1, x)$ and $V(A_2, x)$ have, respectively, at least $n(A_1) + 1$ and $n(A_2) + 1$ vector-extrema of alternating sign in $[a, b]$. Let $\{x_i : x_i < x_{i+1}\}$ be $n(A_1) + 1$ such vector-extrema of $V(A_1, x)$. The expression $F(A_2, x_i) - F(A_1, x_i)$ will be alternately nonnegative and nonpositive as i ranges from 1 to $n(A_1) + 1$. Now, as in the standard case of $k = 1$ ([4], p. 62), the unisolvence of degree $n(A_1)$ of $F(A_1, x)$ implies that $A_1 = A_2$, a contradiction.

5. Remarks

The necessity of the density condition for the type of proof of Theorem 3 given, has been explained in a recent paper by Dunham [1].

A simple example which illustrates Theorems 2, 3 and 4 is the following: Let $f(x) = x$ be approximated by constants $\{a\}$, let $k = 2$, $w_1 \equiv 1$, and

$$w_2 = \begin{cases} \frac{\delta - \epsilon}{\delta} x + \epsilon, & 0 \leqslant x \leqslant \delta, \\ x, & \delta \leqslant x \leqslant 1. \end{cases}$$

For small $\delta > 0$ and $\epsilon > 0$, it is easy to verify that the best approximations consist of each a staisfying $\frac{1}{2} \leqslant a \leqslant -2 + 2\sqrt{2}$, and that the error of each best approximation exhibits vector-alternation. It is also seen that M here is the straight line segment joining the points $G(\frac{1}{2}) = (\frac{1}{2}, \frac{1}{2})$ and $G(-2 + 2\sqrt{2}) = (-2 + 2\sqrt{2}, 3 - 2\sqrt{2})$.

Finally, the results of this paper can be generalized if, instead of using the standard weight functions $w_i(x)$, we use generalized weight functions $W_i(x,y)$, in the sense of Moursund [2]. It is a straightforward matter to verify that all the above theorems remain valid if $\|W_i[\cdot, f(\cdot) - F(A, \cdot)]\|$ is used instead of $\|w_i(\cdot)(f(\cdot) - F(A, \cdot))\|$ for each i, $i = 1, 2, \ldots, k$. An interesting example of the use of a generalized weight function is given in [3].

6. Acknowledgment

This article forms part of the author's doctoral dissertation, written under the direction of Professor Preston C. Hammer, and submitted to the Graduate School of the University of Wisconsin. The author wishes to thank Dr. Hammer for his valuable assistance and encouragement.

References

1. C. B. Dunham, Necessity of alternation. *Canad. Math. Bull.* **10** (1968), 743–744.
2. D. G. Moursund, Chebychev approximation using a generalized weight function. *J. SIAM Numer. Anal.* **3** (1966), 435–450.
3. D. G. Moursund and G. D. Taylor, Optimal starting values for the Newton-Raphson calculation of inverses of certain functions. *J. SIAM Numer. Anal.* To appear.
4. J. R. Rice, "The Approximation of Functions," Vol. 1. Addison-Wesley, Reading, 1964.
5. J. R. Rice, Tchebycheff approximation by functions unisolvent of variable degree. *Trans. Am. Math. Soc.* **99** (1961), 298–302.
6. L. Tornheim, On n-parameter families of functions and associated convex functions. *Trans. Am. Math. Soc.* **69** (1950), 457–467.