JOURNAL OF APPROXIMATION THEORY 26, 124–131 (1979)

Chebyshev Approximation with Sums of Logarithmic Functions

ECKARD SCHMIDT†

Department of Mathematics, University of Calgary, Calgary, Alberta, Canada

Communicated by E. W. Cheney

Received March 29, 1974

1. INTRODUCTION

In [1], Rice suggests the investigation of approximation by functions of the form

$$\sum_{i=1}^{n} a_i \log(1 + t_i x), \qquad a_i \in \mathbb{R}, \qquad x \in [-1, 1], \qquad t_i \in (-1, 1), \quad (1.1)$$

with emphasis on varisolvence.

We give an example showing that the system

$$\{\log(1 + t_i x), \qquad i = 1,..., n\}, \quad (1.2)$$

is not a Haar system in general. Thus, its value for interpolation and uniform approximation is rather limited. It turns out that the presence of the constants plays a crucial role. Dunham [2] was able to show that the system

$$\{1, \log(1 + t_i x), \qquad i = 1,..., n\} \quad (1.3)$$

is a Haar system. On that basis, he derived characterization criteria for best approximants in the family

$$\left\{\nu \in C[0, \alpha], \nu(x) = a_0 + \sum_{i=1}^{n} a_i \log(1 + t_i x), a_i \in \mathbb{R}, t_i \in \mathbb{R}, t_i x > -1\right\}. \quad (1.4)$$

However, for the family (1.1), the existence of a best approximant cannot be guaranteed.

We expand the above family to ensure existence, and we give characterization criteria that contain Dunham's as special cases. The generalized family contains, besides the logarithmic functions, plynomials and certain rational functions. Its properties are similar to those of exponentials, including a compactness property [7]. It differs from the latter as one nonlinear parameter is not free.

† Deceased.

0021-9045/79/060124–08$02.00/0

2. The Approximating Family

First, we present the example mentioned above.

Example 2.1. Let $p(x) := a_1 \log(1 + t_1 x) + a_2 \log(1 + t_2 x)$. Let $a_2 = -2a_1 \neq 0$, $x_1 = 0$, $x_2 = t_2 = 1/3$, $t_1 = 1/9$. Then, $p(x) \not\equiv 0$ and $p(x_1) = p(x_2) = 0$.

Hence, (1.2) is not a Haar system in general. This implies that a definition of the degree of varisolvency has to use both number and value of the nonlinear parameters t_i. This author prefers a concept of a degree that uses only the number of the t_i, counting multiplicities.

To expand (1.3) properly to ensure existence, we use the framework of γ-polynomials; see, e.g., [9]. Given a kernel function $\gamma(x, t)$, one considers the set of functions, so-called γ-polynomials.

$$\left\{ \nu \mid \nu(x) = \sum_{i=1}^{l} \sum_{j=0}^{m_i} a_{ij} [\gamma(x, t_i)]^{(j)} \right\}, \tag{2.1}$$

with

$$[\gamma(x_i, t_i)]^{(j)} := (\partial^j / \partial t^j)\, \gamma(x, t)|_{t=j_i}. \tag{2.2}$$

To generate the logarithmic functions and to include the constant functions, we define

$$\gamma(x, t) := 1 + \log(1 + tx). \tag{2.3}$$

We will use this kernel under the constraint that one of the nonlinear parameters t_i be zero.

Clearly, any function of the form indicated in (1.4) can be represented as

$$\nu(x) = \sum_{i=1}^{n+1} b_i \gamma(x, t_i), \qquad \text{with at least one} \quad t_i = 0$$

and vice versa. We consider, without loss of generality, the intervals $X = [0, 1]$, $T = (-1, \infty)$, and endow $C[0, 1]$ with the uniform norm:

$$\| u \| := \sup_{x \in [0,1]} | u(x) |.$$

Using definitions (2.2) and (2.3), we define the following approximating family

$$V_n := \left\{ h \in C[0, 1],\, h(x) = \sum_{i=1}^{l} \sum_{j=0}^{m_i} a_{ij} [\gamma(x, t_i)]^{(j)}, \right.$$

$$\left. a_{ij} \in \mathbb{R},\, t_i \in T, \text{ distinct},\, t_1 = 0,\, \sum_{i=1}^{l} (m_i + 1) \leqslant n \right\}. \tag{2.4}$$

Thus, a typical element of V_n is a function

$$h(x) = p(x) + \sum_{i=2}^{l} a_{i0} \log(1 + t_i x) + \sum_{i=2}^{l} \sum_{j=1}^{m_i} a_{ij} \frac{x^j}{(1 + t_i x)^j}, \tag{2.5}$$

with p a polynomial of degree at most m_1 and $\sum_{i=1}^{l} (m_i + 1) \leqslant n$.

An important subset of V_n is obtained by letting $m_i = 0$ for $i = 1,..., l$. We define

$$V_n{}^0 := \left\{ h \in V_n, h(x) = a_1 + \sum_{i=2}^{n} a_i \log(1 + t_i x) \right\}. \tag{2.6}$$

The need for including the polynomials or, in other words, specifying one of the nonlinear parameters to be zero can be illustrated by the following example.

EXAMPLE 2.2. Let

$$p(x) = \sum_{i=1}^{2} a_i [1 + \log(1 + t_i x)], \tag{2.7}$$

with $a_1 = 1$, $a_2 = -2$.
Define

$$\lambda_1 := \left(\frac{e - 1}{2e}\right)^{1/2}, \qquad \lambda_2 := e(1 + \lambda_1)^2 - 1. \tag{2.8}$$

Then, it is easy to see that

$$1 + k\lambda_1 = e(1 + k\lambda_2)^2, \qquad \text{for } k = 1, 2 \tag{2.9}$$

and

$$[1 + \log(1 + k\lambda_1)] - 2[1 + \log(1 + k\lambda_2)] = 0. \tag{2.10}$$

Now, let x be arbitrary but fixed $0 < x \leqslant \frac{1}{2}$ and let $x_1 = x$, $x_2 = 2x$. Define

$$t_i := \lambda_i / x, \qquad i = 1, 2. \tag{2.11}$$

Then, $p(x) \not\equiv 0$, $p(x_1) = p(x_2) = 0$, $x_i \in X$, $t_i \in T$, $i = 1, 2$. Here, $t_i \neq 0$ for $i = 1, 2$, and we conclude that the constraint in the definition of V_n that one t_i be zero cannot be omitted. As is easily seen, one can describe V_n by characterizing the derivatives of its elements as follows

$$\begin{aligned} V_n &= h \in C(X), \\ h' &= p/q, \qquad \text{with } p \text{ a real polynomial of degree at most } \quad n - 2, \\ q(x) &= \left\{ \prod_{i=1}^{n-1} (1 + s_i x), s_i \in \mathbb{R}, q(x) > 0 \text{ on } X \right\}. \end{aligned} \tag{2.12}$$

Here, the s_i are not necessarily distinct or nonzero. A few further definitions will be useful.

DEFINITION 2.3. A function $h \in V_n$ is said to be in canonical form if it has a representation (2.5) with m_1 being the degree of the polynomial p and $a_{im_i} \neq 0$ for $i = 2,..., l$.

DEFINITION 2.4. Let a function $h \in V_n$ be in canonical form. Then, the number

$$k(h) := \sum_{i=1}^{l} (m_i + 1) \tag{2.13}$$

is referred to as the degree of h and $l = l(h)$ is called the length of h. An element of V_n has length 1 if and only if it is a polynomial. The degree is well-defined since the functions involved are linearly independent as is seen from Theorem 3.1. Furthermore, it should be noted that in accordance with the restriction $t_1 = 0$, the term $p(x)$ in (2.5) contributes at least 1 to the defining sum in (2.13) whether or not $p(x) \equiv 0$.

3. UNIQUENESS AND CHARACTERIZATION OF BEST APPROXIMATIONS

Example 2.1 shows that the kernel $\gamma(t, x)$ as defined in (2.1) is not strictly sign regular of any order $n \geqslant 2$ and the results of Braess [4] on Descartes families are not applicable here. However, as is seen from Theorem 3.1, the defining determinantal inequalities (2.5) in [4] are valid under the restriction that one of the t_i is zero. In this section, we follow standard arguments using the local Haar condition [5].

THEOREM 3.1. *Any nontrivial function in V_n has at most $n - 1$ zeros in X, counting multiplicities.*

Proof. Let $h \not\equiv 0$ be in V_n. Then, from the equivalent representation (2.12), it follows by Rolle's theorem that h can have at most $n - 1$ (distinct) zeros. Since h satisfies a differential equation of the form

$$u^{(n)}(x) + p_{n-1}(x)\, u^{(n-1)}(x) + \cdots p_0(x)\, u(x) = 0$$

with continuous coefficients, this implies (see, for example, [6, Chap. 3]) that h can have at most $n - 1$ zeros, counting multiplicities. An important consequence is the following

COROLLARY 3.2. *The difference of a function $h \in V_n$ with degree $k(h)$ and* *any other function in V_n has at most $n + k(h) - 2$ zeros in X, counting multiplicities.*

Proof. Let h^1 and h^2 be two arbitrary functions in V_n and assume corresponding superscripts in their respective canonical representations. Then,

$$k(h^1 + h^2) \leqslant k(h^1) + k(h^2) - 1 - \min\{m_1{}^1, m_1{}^2\}, \\ \leqslant k(h^1) + k(h^2) - 1.$$

DEFINITION 3.3. A function $g \in C(X)$ will be said to have an alternant of length m if there are m points x_i in X such that $x_1 < x_2 < \cdots < x_m$, $|g(x_i)| = \max_{x \in X} |g(x)|$ and $g(x_i) = -g(x_{i+1})$ for $i = 1,\ldots, m - 1$.

Now, we can state the main result of this section.

THEOREM 3.4. *Let $f \in C(X)$ and $h \in V_n$.*

(i) *If $f - h$ has an alternant of length $n + k(h)$, then h is the unique best approximation to f in V_n.*

(ii) *If h is a best approximation to f in V_n, then $f - h$ has an alternant of length $n + l(h)$.*

(iii) *There is at most one best approximation to f in $V_n{}^0$. An element $h \in V_n{}^0$ is the best approximation if and only if $f - h$ has an alternant of length $n + l(h)$.*

Proof. Statement (i) follows by standard arguments from Corollary 3.2. Let now h^* be a best approximation to f in V_n, $k^* = k(h^*)$ and $l^* = l(h^*)$. We can write h^* as

$$h^*(x) = \sum_{j=0}^{m_1{}^*} a_j x^j + \sum_{i=2}^{l^*} \sum_{j=0}^{m_i{}^*} b_{ij}[1 + \log(1 + t_i x)]^{(j)} \\ + \sum_{i=k^*+1}^{n} c_i[1 + \log(1 + s_i x)], \tag{3.1}$$

with $b_{im_i{}^*} \neq 0$, $c_i = 0$, and certain fixed $s_i \neq 0$, distinct and distinct from the t_i. Let $U(h^*)$ be the set of all functions in V_n that can be represented in the form of the right-hand side of (3.1) and satisfying the above conditions, except $c_i \neq 0$. Then, for an arbitrary function $g \in U(h^*)$, the corresponding gradient space is spanned by the functions

$$\begin{array}{ll} 1, x,\ldots, x^{m_1{}^*}, & \\ 1 + \log(1 + t_i x), & i = 2,\ldots, l^*, \\ 1 + \log(1 + s_i x), & i = k^* + 1,\ldots, n, \\ x^j/(1 + t_i x)^j, & i = 2,\ldots, l^*, \quad j = 1,\ldots, m_i{}^* + 1 \end{array} \tag{3.2}$$

and is contained in V_{n+l^*-1}.

Since $l^* \leqslant k^*$, Theorem 3.1 implies that $U(h^*)$ satisfies the Haar condition locally with the dimension of the gradient space being $n + l^* - 1$. Obviously, h^* is also a best approximation to f in $U(h^*)$ and thus, (ii) follows by [5, Theorem 12]. Statement (iii) is implied by (i) and (ii) since $l(h) = k(h)$ if $h \in V_n^0$.

4. A Compactness Result

In this section, we show that bounded subsets of V_n are compact in the topology of compact convergence, as is the case with exponential sums [7]. As a tool we use the following lemma, which generalizes a result in [8].

LEMMA 4.1. *Let $M \subset C^2[a, b]$ be a uniformly bounded set of functions with*

$$\| g \| := \sup_{x \in [a,b]} | g(x)| \leqslant K < \infty, \qquad \text{for all} \quad g \in M \tag{4.1}$$

and let there exist a uniform bound for the number of zeros of g''. Then, there exists an infinite subset $\hat{M}$ of M, a finite subset Z of $[a, b]$ and a $\delta_0 > 0$ such that with

$$I_\delta := \{x, \mid x - z \mid > \delta, \qquad z \in Z \cup \{a\} \cup \{b\}\}, \tag{4.2}$$

the inequality

$$| g'(x)| \leqslant 2K/\delta \tag{4.3}$$

holds for all $g \in \hat{M}$, all $x \in I_\delta$ with any δ, $0 < \delta \leqslant \delta_0$.

Proof. We can select a sequence $\{g_m\} \subset M$ for which the zeros of g_m', g_m'' (both assumed to be not identically zero) converge as $m \to \infty$. Let Z be the set of the corresponding limits. The remainder of the proof, which we omit, is a straightforward argument using the monotonicity of $| g_m' |$ on appropriate subintervals.

THEOREM 4.1. *Any set*

$$V_{n,K} := \{h \in V_n , \| h \| \leqslant K < \infty\}, \tag{4.4}$$

contains a sequence that converges uniformly on any compact subinterval of (a, b) to an element of V_n. If d, $0 \leqslant d \leqslant n - 1$, nonlinear parameters converge to the boundary of T, then the limit function is in V_{n-d}.

Proof. Since for each $h \in V_n$, the first and second derivative has at most $n - 2$ and $2n - 4$ zeros, respectively, there exists, by Lemma 4.1, a subsequence $\{h_m\} \subset V_{n,K}$ and a nonempty closed interval Y such that the h_m'

are uniformly bounded on Y. By (2.12), we have $h_m' = p_m/q_m$, where p_m is a polynomial of degree at most $n - 2$ and

$$q_m(x) = \prod_{i=1}^{n-1} (1 + s_i^m x), \tag{4.5}$$

with real s_i^m. Upon normalizing $\| q_m' \|_Y = 1$, we see that there exists a subsequence and polynomials p and q such that $p_m \to p$ and $q_m \to q$, uniformly. This implies that the coefficients and zeros converge. Since all zeros of q_m are real and lie outside [0, 1], there are no zeros of q in (0, 1). Now, let $c = \lim_{m\to\infty} h_m(\frac{1}{2})$ and $r = P/q = \bar{p}/\bar{q}$, where $\bar{p}$ and $\bar{q}$ have no common factors. We define

$$h(x) := c + \int_{1/2}^{x} r(t)\, dt. \tag{4.6}$$

Then, for any δ, $0 < \delta < \delta_0$, $h_m \to h$, uniformly on $[\delta, 1 - \delta]$. Now, we show that h is in V_n. First, we observe that h has no singularity at $x = 0$ or $x = 1$, because if it had one, there would exist an x in the open interval (0, 1) such that

$$\int_{1/2}^{x} r(t)\, dt > 2K,$$

which would imply that $|h(x)| > K$, contradicting the hypothesis $\| h_m \| \leqslant K$. As a result of the normalization, $\bar{q}(x)$ must be of the form

$$\bar{q}(x) = A \cdot x^j \prod_{i=1}^{\nu} (1 + t_i x),$$

with A a real constant. Lemma 4.1 implies that j can be at most 1. However, j must be zero, because otherwise, upon integration, we would get a term $C \cdot \log(x)$, contributing a nonremovable singularity for h at $x = 0$, contrary to the above observation. A similar argument for the left-hand end of the interval shows that $\bar{q}(x)$ has no factor $(1 - x)$. Thus, we have proved that h is in V_n using (2.12). Furthermore, we have shown that if d is the number of nonlinear parameters converging to ∞ or -1, then d linear factors, x^j, $(1 - x)^i$, $j + i = d$ are common to p and q and h' has a representation $h' = \bar{p}/\bar{q}$ with $\partial\bar{p} \leqslant n - 2 - d$, $\partial\bar{q} \leqslant n - 1 - d$, from which we conclude $h \in V_{n-d}$.

5. Existence

Not every function in $C(X)$ has a best approximation in V_n^0 for $n \geqslant 2$, as is clear from $(1/t) \cdot \log(1 + tx) \to x$ as $t \to 0$. However, the results of the previous section imply that the extension of $V_n{}^0$ is appropriate to ensure existence of a best approximation.

THEOREM 5.1. *Let $f \in C[0, 1]$. There exists at least one best Chebyshevian approximation to f in V_n.*

Proof. Let $\{h_m\} \subset V_n$ be a minimizing sequence, i.e.,

$$\lim_{m\to\infty} \| h_m - f \| = \inf_{h \in V_n} \| h - f \|,$$

the norm being the supremum norm over [0, 1]. This sequence is uniformly bounded and by Theorem 4.1, contains a subsequence that converges uniformly on every compact subinterval of (0, 1), hence, pointwise on (0, 1), to an element h^* in V_n. Since f is continuous, a standard argument (see, e.g., [3]) then shows that h^* is a best approximation to f in V_n.

REFERENCES

1. J. R. RICE, "The approximation of Functions," Vol. 2, Addison-Wesley, Reading, Mass. 1969.
2. C. B. DUNHAM, Chebyshev Approximation by Logarithmic Families, *Z. Angew. Math. Mech.* **53** (1973), 352–353.
3. C. B. DUNHAM, Existence and Continuity of the Chebyshev Operator, *SIAM Review* **4** (1968), 444–446.
4. D. BRAESS, Chebyshev Approximation by γ-Polynomials. I, *J. Approximation Theory* **9** (1973), 20–43.
5. G. MEINARDUS AND D. SCHWEDT, Nichtlineare Approximationen, *Arch. Rational Mech. Anal.* **17** (1964), 297–326.
6. W. A. COPPEL, "Disconjugacy," Lecture Notes in Mathematics, Springer, New York, 1971.
7. E. SCHMIDT, Zur Kompaktheit bei Exponentialsummen, *J. Approximation Theory* **3** (1970), 445–454.
8. H. WERNER, Der Existenzsatz fur das Tschebyscheffsche Approximations-problem mit Exponentialsummen, *ISNM* **12** (1969), 133–143.
9. C. R. HOBBY AND J. R. RICE, Approximation from a curve of functions, *Arch. Rational Mech. Anal.* **24** **91**–106.