JOURNAL OF APPROXIMATION THEORY **11**, 338–342 (1974)

Some Remarks on an Alternation Theorem of Geiger

HANS-PETER BLATT

Institut für Angewandte Mathematik I der Universität, Erlangen-Nürnberg 8520 Erlangen, Martensstraße 1, West Germany

Communicated by G. Meinardus

Let I be the interval $[-1, 1]$ and $C(I)$ the linear space of all continuous, real-valued functions defined on I. In $C(I)$ we consider the Tchebycheff norm $\|\cdot\|$ and an n-dimensional subspace $U = \operatorname{span}(u_1, ..., u_n)$. We put

$$V = \left\{ v(a, x) = \frac{u(a, x)}{u(a, -x)} = \frac{\sum_{i=1}^{n} a_i u_i(x)}{\sum_{i=1}^{n} a_i u_i(-x)} \,\middle|\, u(a, x) > 0 \text{ for } x \in I \right\}$$

and suppose $V \neq \phi$. Here a is the vector $(a_1, ..., a_n)$ formed by the coefficients of $u(a, x) = \sum_{i=1}^{n} a_i u_i(x)$. Now we seek for a given $f \in C(I)$ a best approximation with respect to V, i.e., we wish to determine an element $v_0 \in V$ such that

$$\|f - v_0\| \leqslant \|f - v\|$$

for all $v \in V$.

Geiger [5] considered the case in which U is the subspace π_{n-1} of all real polynomials of degree at most $n - 1$. Assuming $u(a, x)$ and $u(a, -x)$ to be relatively prime and defining

$$d(a) = n - 1 - \left[\frac{n - 1 - m}{2}\right],$$

with m = degree of $u(a, x)$, Geiger proved the following characterization theorem:

$v(a, x)$ is a best approximation to f if and only if one of the following conditions holds:

1. 0 is an extremal point of $f(x) - v(a, x)$.
2. There are two extremal points $x_1, x_2 \in I$ of $f(x) - v(a, x)$ such that $x_1 = -x_2$ and $f(x_1) - v(a, x_1) = f(x_2) - v(a, x_2)$.

3. There are $d(a) + 1$ extremal points $x_0, x_1, \ldots, x_{d(a)}$ such that

$$0 < |x_0| < |x_1| < \cdots < |x_{d(a)}|$$

and

$$\operatorname{sgn}(x_i(f(x_i) - v(a, x_i)) = -\operatorname{sgn}(x_{i+1}(f(x_{i+1}) - v(a, x_{i+1}))$$

for $i = 0, 1, \ldots, d(a) - 1$.

We want to characterize best approximations of the first mentioned problem by transforming it to a problem of approximating vector-valued functions and using the characterization of best approximations in normed linear spaces. We can then interpret $d(a)$ as the dimension of the linear space spanned by the gradient functions.

Let $I_0 = [0, 1]$. We can formulate our problem in the following way:

For a given $(f_1, f_2) \in C(I_0) \times C(I_0)$, we seek an element $v(a, x) \in V$ such that

$$\Delta(a) \leqslant \Delta(b)$$

for all $v(b, x) \in V$. Here $f_1(x) = f(x)$, $f_2(x) = f(-x)$ and

$$\Delta(a) = \max \left\{ \| f_1 - v(a, \cdot) \|_0 , \left\| f_2 - \frac{1}{v(a, \cdot)} \right\|_0 \right\},$$

with $\| \cdot \|_0 =$ Tchebycheff norm on I_0.

For characterizing the best approximations we need the gradient of $v(a, x)$ with respect to the parameter a. From

$$\frac{\partial v(a, x)}{\partial a_i} = \frac{1}{u(a, -x)} [u_i(x) - u_i(-x)\, v(a, x)]$$

we get

$$\operatorname{grad} v(a, x) = \frac{1}{u(a, -x)} (u_1(x) - u_1(-x)\, v(a, x), \ldots, u_n(x) - u_n(-x)\, v(a, x)).$$

Let $\langle\, ,\, \rangle$ denote the scalar product of $\mathbb{R}^n$,

$$M_1(a) := \{x \in I_0 \mid |f_1(x) - v(a, x)| = \Delta(a)\},$$

and

$$M_2(a) := \left\{ x \in I_0 \,\middle|\, \left| f_2(x) - \frac{1}{v(a, x)} \right| = \Delta(a) \right\}$$

then the following theorem holds.

THEOREM 1. *Δ achieves its minimum at $a \in \mathbb{R}^n$ if and only if*

$$\min \left\{ \min_{x \in M_1(a)} (f_1(x) - v(a, x))\langle b, \operatorname{grad} v(a, x)\rangle, \right.$$
$$\left. \min_{x \in M_2(a)} \left(\frac{1}{v(a, x)} - f_2(x)\right) \langle b, \operatorname{grad} v(a, x)\rangle \right\} \leqslant 0$$

for every $b \in \mathbb{R}^n$.

Proof. Observing the form of the extremal points of the unit sphere in $(C(I_0) \times C(I_0))^*$ (Bredendiek [1]) and $\operatorname{grad}[1/v(a, x)] = -[1/v(a, x)^2] \operatorname{grad} v(a, x)$ this condition is just the local Kolmogoroff condition, which is always necessary (Brosowski, Wegmann [3]). The set V is asymptotically convex with the parameter function $a(t) = a + t(b - a)$. Hence the set of functions $(v(a, \cdot), 1/v(a, \cdot))$ lying in $V \times V$ is asymptotically convex for every component. From the chosen norm in the product space $C(I_0) \times C(I_0)$ and $a(0) = a$ we get, as did Brosowski [2] that the condition is also sufficient.

Using the separation theorem for convex sets we obtain

THEOREM 2. *Δ achieves its minimum at $a \in \mathbb{R}^n$ if and only if the origin of $\mathbb{R}^n$ lies in the convex hull of $S_1 \cup S_2$ with*

$$S_1 := \{(f_1(x) - v(a, x)) \operatorname{grad} v(a, x) \mid x \in M_1(a)\},$$
$$S_2 := \left\{- \left(f_2(x) - \frac{1}{v(a, x)}\right) \operatorname{grad} v(a, x) \;\middle|\; x \in M_2(a)\right\}.$$

Let $L(a)$ denote the space spanned by the functions

$u_i(x) - u_i(-x)\, v(a, x)$ $(i = 1,..., n)$ and put

$\nu(L(a)) = 1 +$ maximum number of variations in sign possessed by elements of $L(a)$ in $(0, 1]$,

$\eta(L(a)) =$ dimension of a maximal Haar subspace in $L(a)$ over $(0, 1]$.

We say that $f - v(a, \cdot)$ *alternates k times* if k points $x_i \in I$ exist satisfying

1. $0 < |x_1| < |x_2| < \cdots < |x_k|$.
2. $|f(x_i) - v(a, x_i)| = \|f - v(a, \cdot)\|$.
3. $\operatorname{sgn}(x_i(f(x_i) - v(a, x_i)) = -\operatorname{sgn}(x_{i+1}(f(x_{i+1}) - v(a, x_{i+1}))$ for $i = 1,..., k - 1$.

Then we get the following alternation theorem.

THEOREM 3. *If $v(a, x)$ is a best approximation to f with respect to V then one of the following conditions holds*:

α. 0 *is an extremal point of $f(x) - v(a, x)$.*

β. *There are two extremal points $x_1, x_2 \in I$ of $f(x) - v(a, x)$ such that $x_1 = -x_2$ and $f(x_1) - v(a, x_1) = f(x_2) - v(a, x_2)$.*

γ. *$f - v(a, \cdot)$ alternates $1 + \eta(L(a))$ times.*

If one of the conditions (α), (β) holds or if $f - v(a, \cdot)$ alternates $1 + \nu(L(a))$ times then $v(a, x)$ is a best approximation to f.

We omit the proof of this theorem because we can derive it as in the classical theory using the theorems of Carathéodory and Goldstein (compare Cheney [4]). The exceptional cases are immediately seen from Theorem 2. Moreover, it is clear that the best approximation $v(a, x)$ to f is unique if $L(a)$ is a Haar subspace over $(0, 1]$ and $v(a, x)$ is characterized by condition (γ).

Now the aforementioned theorem of Geiger [5] follows from Theorem 3 considering $U = \pi_{n-1}$ and the

LEMMA. *If $U = \pi_{n-1}$ then $L(a)$ is a Haar subspace of dimension $d(a)$ over any interval not containing the point* 0.

Proof. Since

$$\sum_{i=0}^{n-1} \alpha_i(x^i u(a, -x) - (-x)^i u(a, x))$$

is an odd polynomial of degree $\leqslant (n - 1 + m)$, the function

$$\sum_{i=0}^{n-1} \alpha_i(x^i - (-x)^i v(a, x))$$

has at most $d(a) - 1$ zeros in intervals not containing 0. Consider the linear mapping

$$\phi\colon \pi_{n-1} \to L(a)$$

defined by $x^i \to x^i - (-x)^i v(a, x)$ for $i = 0, 1, \ldots, n - 1$. Hence we conclude

$$\dim L(a) = \dim \pi_{n-1} - \dim(\ker \phi).$$

From the equivalent relations

$$p(x) = \sum_{i=0}^{n-1} \alpha_i x^i \in \ker \phi \Leftrightarrow \sum_{i=0}^{n-1} \alpha_i x^i = v(a, x) \sum_{i=0}^{n-1} \alpha_i(-x)^i$$

$$\Leftrightarrow p(x) = u(a, x) \cdot r(x) \quad \text{with even} \quad r(x)$$

we get:

$$\dim(\ker \phi) = \left[\frac{n - 1 - m}{2}\right] + 1.$$

Thus we obtain $\dim L(a) = d(a)$.

References

1. E. Bredendiek, Charakterisierung und Eindeutigkeit bei Simultanapproximation, *Z. Angew. Math. Mech.* **50** (1970), 403–410.
2. B. Brosowski, Nicht-lineare Tschebyscheff-Approximation, *Bibliograph. Institut Mannheim* (1968).
3. B. Brosowski and R. Wegmann, Charakterisierung bester Approximationen in normierten Vektorräumen, *J. Approximation Theory* **3** (1970), 369–397.
4. E. W. Cheney, Approximation by Generalized Rational Functions, from "Approximation of Functions," H. L. Garabedian, Ed., Elsevier, Amsterdam (1965), 101–110.
5. C. Geiger, Über eine Klasse rationaler Tschebyscheff-Approximationen mit Nebenbedingungen, *J. Approximation Theory* **1** (1968), 340–354.