TCHEBYCHEFF APPROXIMATION IN A COMPACT METRIC SPACE

BY JOHN R. RICE

Communicated by A. Erdélyi, March 26, 1962

1. **Introduction.** In this note[1] an outline of two theories of Tchebycheff approximation is given for functions defined on a compact metric space. The first[2] of these lacks the elegance of the one variable theory, but is descriptive of the true situation. Many results of the one variable theory have counterparts here. The second theory is developed only for functions defined on finite point sets. A special type of Tchebycheff approximation, the strict approximation, is introduced and the resulting theory is similar to the classical one variable theory. It is, in particular, shown that the strict approximations are unique. These theories allow one to solve the central problem of approximation theory, the computation of best and strict approximations.

2. **Preliminaries.** Spaces and sets in general are denoted by capital script letters $\mathcal{A}$, $\mathcal{B}$, $\cdots$ and elements of such sets are denoted by lower case letters x, y, $\cdots$. Let $\mathcal{B}$ be a compact metric space. The space of real-valued continuous functions defined on $\mathcal{B}$ is denoted by $\mathcal{C}$ and has elements f, g, $\cdots$. The norm in $\mathcal{C}$ is taken to be

$$\|f\| = \max_{x\in\mathcal{B}} |f(x)| .$$

Let $\mathcal{L}$ be an n-dimensional subspace of $\mathcal{C}$ with basis functions $g_i(x)$, $i=1, 2, \cdots, n$.

$$L(A, x) = \sum_{i=1}^{n} a_i g_i(x), \qquad | a_i | < \infty,$$

is an element of $\mathcal{L}$ with parameters $A=(a_1, a_2, \cdots, a_n)$.

The Tchebycheff approximation problem in this context is stated as follows: Given $f(x)$ in $\mathcal{C}$ determine A^* such that

$$\|f(x) - L(A^*, x)\| \leqq \|f(x) - L(A, x)\|$$

for all A. Such a $L(A^*, x)$ is said to be a *best approximation* to $f(x)$ with *deviation* $\|f(x)-L(A^*, x)\|$. The elements of $\mathcal{B}$ where the norm is assumed, i.e.,

[1] Proofs of these results and some related results will be published elsewhere. Preprints are available from the author for some of this material.

[2] A similar viewpoint has been presented recently by Lawson [5].

$$|f(x) - L(A^*, x)| = \|f(x) - L(A^*, x)\|$$

are said to be *extremal points*. An extremal point x_0 is said to be positive or negative according as $f(x_0) - L(A^*, x_0)$ is positive or negative. $\mathcal{P}$ and $\mathcal{N}$ are sets of positive and negative extremal points. These designations may also be used when $L(A^*, x)$ is not a best approximation to $f(x)$. The sets $\mathcal{P}$ and $\mathcal{N}$ are said to be *isolable* if there is an A such that $L(A, x) > 0$ on $\mathcal{P}$ and $L(A, x) < 0$ on $\mathcal{N}$.

It is known that the extremal points play an important role in the determination of best approximations. In order to study Tchebycheff approximation in more than one variable our attention is focused on particular subsets of the extremal point set. A subset $\mathcal{R}$ of extremal points is said to be a *critical point set* if the positive and negative parts, $\mathcal{P}$ and $\mathcal{N}$, of $\mathcal{R}$ are not isolable, but the positive and negative parts of any proper subset of $\mathcal{R}$ are isolable. Critical point sets play a central role in this note.

3. Uniqueness and characterization. The existence of best approximations follows from the fact that $\mathcal{B}$ is a compact metric space.

It is known that best approximations may not be unique. Mairhuber [6] has shown that if best approximations are to be unique for every $f(x)$ then $\mathcal{B}$ must be homeomorphic to a subset of the unit circle. See [1] and [8] for further discussion. Recently Rivlin and Shapiro [7] have shown that there is no possibility of formulating an interesting restricted approximation problem, a possibility suggested by the result of Collatz [3].

There is a unique set associated with this problem, the set of critical point sets. Note that this set is not the set of extremal points. We have

THEOREM 1. *The sets of critical point sets of two distinct best approximations to a given $f(x)$ are identical.*

In the theory of Tchebycheff approximation for one real variable, the concepts of *alternation*, or equioscillation, and *Tchebycheff sets* hold a central position. Alternation is an intrinsic feature of Tchebycheff approximation and the concept of a critical point set is specifically introduced to describe this phenomenon. Critical point sets are intimately related to the concept of *irreducibly inconsistent systems of linear inequalities* [2]. The possibility for the generalization of Tchebycheff sets is not promising, as seen by the example of polynomials.

THEOREM 2. *$L(A^*, x)$ is a best approximation to $f(x)$ if and only if the set of extremal points of $L(A^*, x) - f(x)$ contains a critical point set.*

The following theorem is necessary for the applicability of the method of ascent to the computation of best approximations.

THEOREM 3. *Let $L(A^*, x)$ be a best approximation to $f(x)$ with a critical point set $\mathcal{R}$ of k points. Then $L(A^*, x)$ is a best approximation to $f(x)$ on $\mathcal{R}$ and is characterized as a best approximation with the largest deviation among all best approximations to $f(x)$ on subsets of k points of $\mathcal{B}$.*

4. **Strict approximations.** In this section a new type of best Tchebycheff approximation is defined which is a natural extension of Tchebycheff approximation for functions of several variables.[3] The definition is valid only when $\mathcal{B}$ is a finite point set. There is some machinery to be established.

The n-vector $\mathbf{g}$ is defined for a point $x \in \mathcal{B}$ by

$$\mathbf{g} = \mathbf{g}(x) = (g_1(x), g_2(x), \cdots, g_n(x))$$

and the resulting mapping from $\mathcal{B}$ into $\mathcal{E}_n$ is denoted by G. Square brackets, $[\mathbf{g}_i]$, denote the smallest linear subspace of $\mathcal{E}_n$ containing the vectors $\mathbf{g}_i$. The *dimension* of a set $\mathcal{R}$ in $\mathcal{B}$ with respect to the mapping G is defined as the dimension of the linear subspace $[G(\mathcal{R})]$. A set $\mathcal{R}$ in $\mathcal{B}$ is said to be *nondegenerate* if there is no point $x_0 \in \mathcal{R}$ such that $\mathbf{g}(x_0)$ is not contained in the subspace

$$[\mathbf{g}(x) \mid x \in \mathcal{R}, x \neq x_0].$$

ASSUMPTION. *$\mathcal{B}$ is nondegenerate.*

This assumption is made throughout the remainder of this note. A *critical point set with respect to* $\mathcal{L}_1 \subset \mathcal{L}$ is a set $\mathcal{R}$ whose positive and negative parts cannot be isolated by a member of $\mathcal{L}_1$.

DEFINITION. *A constructive definition is given for a strict approximation. Let $\mathcal{R}_1$ be the set of critical point sets of best approximations to $f(x)$ on $\mathcal{B}$. Define $\mathcal{B}_1$ by*

$$\mathcal{B}_1 = \{x \mid \mathbf{g}(x) \in [G(\mathcal{R}_1)]\}$$

and denote by $\mathcal{L}_1$ the set of best approximations to $f(x)$ on $\mathcal{B}$. Then $\mathcal{L}_2$ is defined as the set of $L(A^*, x)$ such that

$$\|L(A^*, x) - f(x)\|_{\mathcal{B}-\mathcal{B}_1} \leqq \|L(A, x) - f(x)\|_{\mathcal{B}-\mathcal{B}_1}$$

for all $A \in \mathcal{L}_1$. Let $\mathcal{R}_2$ be the set of critical point sets with respect to $\mathcal{L}_1$.

[3] In a private communication Dr. Jean Descloux has given a proof of the following result: Assume that there is no unique best Tchebycheff approximation. Let $L(A_p, x)$ be the best L_p approximation to $f(x)$ and let $L(A^*, x)$ be the strict approximation. Then, if $\mathcal{B}$ is a finite set, $\lim_{p\to\infty} L(A_p, x) = L(A^*, x)$.

This construction is continued until $\mathcal{B}_k = \mathcal{B}$. The elements of $\mathcal{L}_k$ are said to be strict approximations to $f(x)$ on $\mathcal{B}$.

On each of the sets $\mathcal{B} - \mathcal{B}_i$ the strict approximations have a maximum deviation d_i. Let the dimension of $\mathcal{R}_i$ be m_i. The *deviation vector* **d** is the vector whose first m_1 components are d_1, whose next m_2 components are d_2 and so forth. The deviation vectors are *ordered lexicographically*. *A strict critical point set* $\mathcal{S}$ is the union of one critical point set $\mathcal{S}_i$ with respect to $\mathcal{L}_i$ from each of the sets $\mathcal{R}_{i+1}$. Denote the positive and negative parts of $\mathcal{S}_i$ by $\mathcal{P}_i$ and $\mathcal{N}_i$, respectively. Then none of the following systems of equalities and inequalities has a solution:

$$(1)\qquad \begin{cases} L(A_1, x) > 0, & x \in \mathcal{P}_1, \\ L(A_1, x) < 0, & x \in \mathcal{N}_1, \end{cases}$$
$$\begin{cases} L(A_2, x) = 0, & x \in \mathcal{S}_1, \\ L(A_2, x) > 0, & x \in \mathcal{P}_2, \\ L(A_2, x) < 0, & x \in \mathcal{N}_2, \end{cases}$$
$$\cdots$$
$$\begin{cases} L(A_i, x) = 0, & x \in \bigcup_{j=1}^{i-1} \mathcal{S}_j, \\ L(A_i, x) > 0, & x \in \mathcal{P}_i, \\ L(A_i, x) < 0, & x \in \mathcal{N}_i, \end{cases}$$
$$\cdots.$$

Any point x in $\mathcal{B} - \mathcal{B}_{i-1}$ where $|L(A^*, x) - f(x)| = d_i$ is said to be an *extremal* point of the strict approximation, $L(A^*, x)$.

THEOREM 4. *Let $f(x)$ be a function defined on a nondegenerate finite set $\mathcal{B}$. Then*

A. *$f(x)$ possesses a strict approximation $L(A^*, x)$ in $\mathcal{L}$,*

B. *the strict approximation is unique,*

C. *$L(A^*, x)$ is the strict approximation to $f(x)$ if and only if the set of extremal points of $L(A^*, x) - f(x)$ contains a strict critical point set of dimension n.*

5. **Computation.** In this section a method of ascent algorithm of the 1 for 1 exchange type is described. Only a method for strict approximations is given. For the computation of best approximations a simplified version may be used.

Assume that at the kth step one has found a strict approximation

$L(A_k, x)$ to $f(x)$ on a nondegenerate strict critical subset $\mathcal{R}_k$ of $\mathcal{B}$ with deviation vector $\mathbf{d}_k$. Let $\mathcal{S}_{ik}$ be the critical point sets which compose $\mathcal{R}_k$. The next step of the algorithm is then:

1. Find $x_{k+1} \in \mathcal{B}$ such that $\mathbf{g}(x_{k+1})$ lies in

$$[G(\cup \mathcal{S}_{jk}) \mid j = j_1, j_2, \cdots, j_p]$$

and such that

$$| L(A_k, x_{k+1}) - f(x_{k+1}) | = d_0 > \min d_{jk}, \qquad j = j_1, j_2, \cdots, j_p.$$

If no such x_{k+1} exists then $L(A_k, x)$ is the strict approximation to $f(x)$ on $\mathcal{B}$.

2. Determine the strict approximation $L(A_{k+1}, x)$ on $\mathcal{R}_k \cup \{x_{k+1}\}$ and choose $\mathcal{R}_{k+1}$ as a strict critical point set of $L(A_{k+1}, x)$.

LEMMA. $\mathbf{d}_{k+1} > \mathbf{d}_k$.

THEOREM 5. *The sequence* $L(A_k, x)$ *converges to the strict approximation* $L(A^*, x)$ *to* $f(x)$ *on* $\mathcal{B}$.

There are some nontrivial details of the algorithm relative to the second step which remain to be described. The strict approximation on $\mathcal{R}_k \cup \{x_{k+1}\}$ is determined as follows:

Let $\mathcal{S}'$ denote the smallest collection of $\mathcal{S}_{jk}$ upon which $\mathbf{g}(x_{k+1})$ depends. Select those subsets $\mathfrak{I}_j$, $j = 1, 2, \cdots, q$, of $\mathcal{S}'$ for which $d_j < d_0$. The natural ordering of the $\mathcal{S}_{jk}$ is retained and the positive part, negative part and deviation vector component of $\mathfrak{I}_j$ are denoted by $\mathcal{U}_j$, $\mathcal{V}_j$ and d_j, respectively. Consider the system (the point x_{k+1} is added to both $\mathcal{U}_q$ and $\mathcal{V}_q$):

$$
\begin{aligned}
&L(A, x) = 0, && x \in \bigcup_{j=1}^{p} \mathfrak{I}_j, \quad x \notin \mathcal{S}', \\
(2)\qquad &L(A, x) - f(x) < d_p, && x \in \bigcup_{j=p+1}^{q} \mathcal{U}_j, \\
&L(A, x) - f(x) > -d_p, && x \in \bigcup_{j=p+1}^{q} \mathcal{V}_j.
\end{aligned}
$$

Let p^* be the largest integer for which this system of equalities and inequalities is consistent. Once p^* is determined one examines the irreducibly inconsistent subsystems [4] of (2). The subsystem which yields the largest deviation corresponds to the critical point set to be used for the next step of the algorithm. $L(A_{k+1}, x)$ may be explicitly computed from this subsystem.

REFERENCES

1. R. C. Buck, *Linear spaces and approximation theory*, On Numerical Approximation, pp. 11–23, Univ. of Wisconsin Press, Madison, Wis., 1959.

2. W. B. Carver, *Systems of linear inequalities*, Ann. of Math. **23** (1922), 212–220.

3. L. Collatz, *Approximation von Funktionen bei einer und bei mehreren unabhängigen Veränderlichen*, Z. Angew. Math. Mech. **36** (1956), 198–211.

4. Ky Fan, *On systems of linear inequalities*, Linear Inequalities and Related Systems, pp. 99–156, Princeton Univ. Press, Princeton, N. J., 1956.

5. C. L. Lawson, *Contributions to the theory of linear least maximum approximation*, Thesis, University of California, Los Angeles, 1961.

6. J. C. Mairhuber, *On Haar's theorem concerning Chebychëv approximation problems having unique solutions*, Proc. Amer. Math. Soc. **7** (1956), 609–615.

7. T. J. Rivlin and H. S. Shapiro, *Some uniqueness problems in approximation theory*, Comm. Pure Appl. Math. **13** (1960), 35–47.

8. I. J. Schoenberg, *On the question of unicity in the theory of best approximation*, Ann. New York Acad. Sci. **86** (1960), 682–691.

RESEARCH LABORATORIES, GENERAL MOTORS CORPORATION