JOURNAL OF APPROXIMATION THEORY 2, 433–447 (1969)

Approximations of Functions of Several Variables: Product Chebychev Approximations, I

STANLEY E. WEINSTEIN

University of Utah, Salt Lake City, Utah 84112

1. INTRODUCTION

Let $\mathscr{F}$, $\|\cdot\|$ be a real normed linear space, and let $\phi_1, \phi_2, \ldots, \phi_n$ be n linearly independent elements of $\mathscr{F}$. For each $\alpha = (a_1, a_2, \ldots, a_n) \in E_n$, the (real) Euclidean n-space, let $P_\alpha = a_1\phi_1 + a_2\phi_2 + \ldots + a_n\phi_n$.

For each given $f \in \mathscr{F}$, the problem of finding an element $P_{\alpha*}$ such that

$$\|F - P_{\alpha*}\| \leqq \|F - P_\alpha\|, \qquad \text{for all } \alpha \in E_n,$$

is called the best approximation (b.a.) problem, and a corresponding solution is called a best approximation to F.

The existence of a solution to the b.a. problem is well known (see e.g. [3]). However, the question of uniqueness is more complex. If the norm in question is strictly convex, then each $F \in \mathscr{F}$ has a unique b.a. The L_p norms, $1 < p < \infty$, are strictly convex; however, the important L_1 and L_∞ norms are not.

Haar [5] proved the following theorem, which helped to answer the question of uniqueness for the uniform norm.

THEOREM 1.1 (Haar). *Let the compact set $D \subset E_k$ contain at least n distinct points. A necessary and sufficient condition that a unique solution exist to the best uniform approximation problem for each given $F \in C(D)$, is that the functions $\{\phi_i\}$ satisfy the following property:*

(1) $P_\alpha = \sum_{i=1}^n a_i\phi_i$ *vanishes in at most $n-1$ distinct points of D, unless $a_1 = a_2 = \ldots = a_n = 0$.*

Property (1) *is equivalent to:*

$$(2) \quad \begin{vmatrix} \phi_1(x_1) & \phi_2(x_1) & \ldots & \phi_n(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \ldots & \phi_n(x_2) \\ \vdots & \vdots & & \vdots \\ \phi_1(x_n) & \phi_2(x_n) & \ldots & \phi_n(x_n) \end{vmatrix} \neq 0$$

for every set $x_1, x_2, \ldots, x_n$ of distinct points of D.

DEFINITION 1.2. A family of real-valued functions $\phi_1, \phi_2, \ldots, \phi_n$ is called a Chebyshev system on D, if and only if they satisfy property (1) (or, equivalently, property (2)) of Theorem 1.1.

Remark. If $D^* \subset D$ and $\{\phi_i\}$ is a Chebyshev system on D, then it is also a Chebyshev system on D^*.

Remark. If $\{\phi_i\}$ is a Chebyshev system on D (containing at least n distinct points), then it is a linearly independent system on D.

Mairhuber [7] established necessary and sufficient conditions for a set D to serve as the domain of definition of a Chebyshev system.

THEOREM 1.3 (Mairhuber). *A compact set $D \subset E_k$, containing at least n points, may serve as the domain of definition of a nontrivial ($n \geqq 2$) Chebyshev system of real-valued continuous functions $\phi_1, \phi_2, \ldots, \phi_n$, if and only if D is homeomorphic to a closed subset of the circumference of a circle.*

Therefore, with the exception of the two cases cited in the above theorem, the best uniform approximation problem has a non-unique solution for some $F \in C(D)$.

Chebyshev sets also play a significant role in the question of uniqueness of a best L_1 approximation (see e.g. [11]). However, the primary goal of this paper is to define and investigate a Chebyshev-like approximation which possesses the property of uniqueness for all $F \in C(D)$. We shall call this approximation the product Chebyshev approximation. We simultaneously define product approximations for all of the L_p norms, $1 \leqq p \leqq \infty$.

The primary virtue of uniqueness is one of communication. We can speak of "the" best uniform or "the" least squares approximation to F. Secondly, uniqueness will often facilitate the algorithm (s) used to find an approximation.

2. THE PRODUCT L_p APPROXIMATIONS TO A CONTINUOUS FUNCTION

Let D denote the rectangle $[a,b] \times [c,d]$, $a < b$, $c < d$, and let $F \in C(D)$. For each $y \in [c,d]$, define the continuous function F_y on $[a,b]$, by

$$F_y(\cdot) = F(\cdot, y).$$

Let $\phi_1, \phi_2, \ldots, \phi_n$ be a Chebyshev system of continuous real-valued functions on $[a,b]$. Let $\|\cdot\|$ be any one of the L_p norms on $[a,b]$, $1 \leqq p \leqq \infty$. Let $\langle \phi_1, \phi_2, \ldots, \phi_n \rangle$ denote the linear space spanned by $\phi_1, \phi_2, \ldots, \phi_n$.

For each $y \in [c,d]$, there exists (even when $p = 1$) a unique "polynomial"

$$P_{\alpha(y)} = \sum_{j=1}^{n} a_j(y) \phi_j, \qquad \alpha(y) = (a_1(y), a_2(y), \ldots, a_n(y)) \in E_n,$$

which is the best $\|\cdot\|$ approximation to F_y on $[a,b]$.

We will show that the functions $a_j(\cdot), j=1,2,\ldots,n$, are continuous on $[c,d]$. Hence, if $\psi_1,\psi_2,\ldots,\psi_m$ is a system of real-valued continuous functions on $[c,d]$, then each a_j possesses, among all linear combinations of the ψ's, a b.a. on $[c,d]$.

For each $y \in [c,d]$, let

$$\rho(y) = \min_{\alpha\in E_n} \|F_y - P_\alpha\| = \|F_y - P_{\alpha(y)}\|.$$

THEOREM 2.1. *The function $\rho(\cdot)$ is continuous on $[c,d]$.*

Proof. Given any $\epsilon > 0$, by the uniform continuity of F on D, there exists a $\delta = \delta(\epsilon) > 0$ such that

$$|y - y_1| < \delta, \qquad y \text{ and } y_1 \in [c,d] \Rightarrow$$

$$\begin{aligned}\rho(y) - \rho(y_1) &= \|F_y - P_{\alpha(y)}\| - \|F_{y_1} - P_{\alpha(y_1)}\| \\ &\leqq \|F_y - P_{\alpha(y_1)}\| - \|F_{y_1} - P_{\alpha(y_1)}\| \\ &\leqq \|F_y - F_{y_1}\| < \epsilon.\end{aligned}$$

Similarly, $\rho(y_1) - \rho(y) \leqq \|F_{y_1} - F_y\| = \|F_y - F_{y_1}\| < \epsilon$.

Throughout this paper σ will denote the usual Euclidean metric on the space in question.

THEOREM 2.2. *Let $X, \|\cdot\|$ be a normed linear space and let $F, \phi_1,\phi_2,\ldots,\phi_n \in X$. For each $\alpha = (a_1,a_2,\ldots,a_n) \in E_n$ let $P_\alpha = \sum_1^n a_i\phi_i$. Let $\rho = \inf_{\alpha\in E_n}\|F - P_\alpha\|$ and*

$$A^* = \{\alpha \in E_n : \|F - P_\alpha\| = \rho\}.$$

Given any $\epsilon > 0$, there exists a $\delta = \delta(\epsilon) > 0$ such that

$$\|F - P_\alpha\| \leqq \rho + \delta \Rightarrow \sigma(\alpha; A^*) = \inf_{\alpha^*\in A^*} \sigma(\alpha;\alpha^*) < \epsilon.$$

Proof. A^* is a nonempty compact set. The function R defined by $R(\alpha) = \|F - P_\alpha\|$ is continuous on E_n.

Assume that there exists a sequence $\{\alpha_n\}$ such that $R(\alpha_n) \leqq \rho + (1/n)$ and $\sigma(\alpha_n; A^*) \geqq \epsilon$. Let $\|F\| = M$. Then,

$$\begin{aligned}\|P_{\alpha_n}\| &\leqq \|F\| + \|F - P_{\alpha_n}\| \\ &\leqq M + \rho + \frac{1}{n} \leqq M + \rho + 1, \qquad \text{for } n = 1,2,\ldots.\end{aligned}$$

Therefore, the sequence $\{\alpha_i\colon i = 1,2,\ldots\}$ is bounded (see, e.g., [6], p. 16). Let α be any limit point of $\{\alpha_i\}$. Then $R(\alpha) \leqq \rho$, which implies $\alpha \in A^*$. However, $\sigma(\alpha_n; A^*) \geqq \epsilon$ for $n = 1,2,\ldots$, implying $0 = \sigma(\alpha; A^*) \geqq \epsilon$.

THEOREM 2.3. *The function $\alpha(\cdot)$, defined at the beginning of this section, is continuous on $[c,d]$.*

Proof. Let y_1 be a point of $[c,d]$ and let $\epsilon > 0$. Let $\delta > 0$ be such that $\|F_{y_1} - P_\alpha\| \leqslant \rho(y_1) + \delta \Rightarrow \sigma(\alpha;\alpha(y_1)) < \epsilon$. By the uniform continuity of F on D and of $\rho(\cdot)$ on $[c,d]$, there exists a $\delta_1 > 0$ such that if $y \in [c,d]$ and $|y - y_1| < \delta_1$, then $\|F_{y_1} - F_y\| \leqslant \delta/2$ and $|\rho(y) - \rho(y_1)| \leqslant \delta/2$. For such a y, we have

$$\begin{aligned}\|F_{y_1} - P_{\alpha(y)}\| &\leqq \|F_{y_1} - F_y\| + \|F_y - P_{\alpha(y)}\| \\ &\leqslant \frac{\delta}{2} + \rho(y) \\ &\leqslant \rho(y_1) + \delta,\end{aligned}$$

and, therefore, $\sigma(\alpha(y);\alpha(y_1)) < \epsilon$.

COROLLARY. *The functions $a_j(\cdot), j = 1,2,\ldots,n$, as defined at the beginning of this section, are continuous on $[c,d]$.*

Let $\psi_1, \psi_2, \ldots, \psi_m$ be a Chebyshev system of continuous real-valued functions on $[c,d]$.

DEFINITION 2.4. Let

$$Q_{\alpha_j} = \sum_{i=1}^{m} a_{ij}\psi_i, \qquad \alpha_j = (a_{1j}, a_{2j}, \ldots, a_{mj}) \in E_m,$$

be the unique best L_p approximation to $a_j(\cdot)$, for $j = 1,2,\ldots,n$.

The function

$$T_A = \sum_{j=1}^{n} Q_{\alpha_j}\phi_j = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}\psi_i\phi_j$$

is called the product L_p approximation to F on D relative to the variable y.

We can similarly define the product L_p approximation to F on D relative to the variable x. The following example illustrates that these two approximations may be distinct.

EXAMPLE 1. Let $D = [-1,1] \times [-1,1]$,

$$F(x,y) \equiv \begin{cases} -2yx^2 + y, & -1 \leqq y \leqq 0, \\ 2yx, & 0 < y \leqq 1, \end{cases}$$

and let $n = m = 1$, $\phi_1(x) \equiv 1$, $\psi_1(y) \equiv 1$.

Then the best uniform approximation to $F_y(\cdot)$ on $[-1,1]$ is $P_{\alpha(y)}(\cdot) \equiv 0$, and the product L_∞ approximation to F on D relative to y is $T_A(x,y) \equiv 0$.

Similarly we find that the product L_∞ approximation to F on D relative to x is $T_B(x,y) \equiv (2-\sqrt{2})/4$.

Product L_∞ approximations will be called product Chebyshev (P.C.) approximations. Throughout the remainder of this paper, P.C. approximations will mean P.C. approximations relative to the variable y.

Remark. If T_A is the product L_p approximation to F on D, and λ is any real number, then λT_A is the product L_p approximation to λF on D.

Remark. If $P_{\alpha^*}(x) \equiv \sum_{i=1}^{n} a_i^* \phi_i(x)$ is the best L_p approximation to $F(x)$ on $[a,b]$, and $Q_{\beta^*}(y) \equiv \sum_{j=1}^{m} b_j^* \psi_j(y)$ is the best L_p approximation to $G(y)$ on $[c,d]$, then $P_{\alpha^*}(x)\, Q_{\beta^*}(y)$ is the product L_p approximation to $H(x,y) \equiv F(x)G(y)$ on $D=[a,b]\times[c,d]$.

Remark. If $\phi_1(x)$ and $\psi_1(y)$ are nonzero constants, and if $P_{\alpha^*}(x)$ and $Q_{\beta^*}(y)$ are as in the last Remark, then $P_{\alpha^*}(x)+Q_{\beta^*}(y)$ is the product L_p approximation to $H(x,y)\equiv F(x)+G(y)$ on $D=[a,b]\times[c,d]$.

The following counter-example illustrates the necessity of the inclusion of a nonzero constant in each of the systems $\{\phi_j\}$, $\{\psi_i\}$ in this Remark.

EXAMPLE 2. Let $D=[1,2]\times[1,2]$ and let

$$H(x,y)\equiv x+y, \qquad n=1, m=2, \qquad \phi_1(x)\equiv x, \qquad \psi_1(y)\equiv 1, \qquad \psi_2(y)\equiv y.$$

Then $\{\phi_i\}$ is a trivial Chebyshev system on $[1,2]$, since $a_1 x=0$ can have no solutions in $[1,2]$, unless $a_1=0$.

For each $y\in[1,2]$, $P_{\alpha(y)}(x)\equiv(1+\frac{2}{3}y)x$, and therefore the P.C. approximation to H on D is $T_A(x,y)\equiv(1+\frac{2}{3}y)x$.

The best uniform approximation to $F(x)\equiv x$ on $[1,2]$ is $P_{\alpha^*}(x)\equiv x$, and the best uniform approximation to $G(y)\equiv y$ on $[1,2]$ is $Q_{\beta^*}(y)=y$. However,

$$P_{\alpha^*}(x)+Q_{\beta^*}(y)\equiv x+y \not\equiv T_A(x,y).$$

Similarly, if H and D are as above and $n=2$, $m=1$, $\phi_1(x)\equiv 1$, $\phi_2(x)\equiv x$, $\psi_1(y)\equiv y$, then

$$T_A(x,y)\equiv(1+\tfrac{2}{3}x)y \not\equiv x+y.$$

3. Computation of the Product Chebyshev Approximation

Throughout the remainder of this paper, our attention is restricted to the uniform norm.

DEFINITION 3.1. Let $D \subset E_k$ and let G map D into E_1. A point $x_0 \in D$ is called a positive (negative) extremal or an $e+$ ($e-$) point for G if $G(x_0) = \sup_D |G(x)|$ ($G(x_0) = -\sup_D |G(x)|$).

The best uniform approximation is characterized by the following theorem.

THEOREM 3.2. *Let $\{\phi_i\}$ be a Chebyshev system of continuous functions on an interval $D = [a,b]$ $(-\infty < a < b < \infty)$. Then P_α is the best uniform approximation to a continuous F on D if and only if there are $n+1$ points $x_1 < x_2 < \ldots < x_{n+1}$ in D, which are alternately $e+$ and $e-$ points for $F - P_\alpha$.*

Note that this characteristic point set need not be unique.

The linear system

$$P_\alpha(x_i) + (-1)^i \rho = F(x_i), \qquad i = 1,2,\ldots,n+1,$$

can be solved for $\alpha = (a_1, a_2, \ldots, a_n)$ and ρ. Therefore, by finding a characteristic point set, one can obtain the b.a.

One iterative procedure which seeks to find the b.a. by finding such a point set, is the Remez exchange algorithm. The convergence of this procedure is outlined in Remez [*10*] and proved in Novodvorskii and Pinsker [*8*]. Verdinger [*17*] shows that if F is differentiable, then the rate of convergence is quadratic.

We can now describe the first product Chebyshev algorithm.

Algorithm 1. (1) Choose some finite point set $Y \subset [c,d]$.

(2) For each $y \in Y$, use the Remez exchange algorithm to find $P_{\alpha(y)}$, the best uniform approximation in $\langle \phi_1, \phi_2, \ldots, \phi_n \rangle$ to F_y on $[a,b]$.

(3) For each $j = 1,2,\ldots,n$, use the Remez exchange algorithm to find $\hat{Q}_{\alpha j}$, the best uniform approximation in $\langle \psi_1, \psi_2, \ldots, \psi_m \rangle$ to $a_j(\cdot)$ on Y.

The function $\hat{T}_A = \sum_{j=1}^n \hat{Q}_{\alpha j} \phi_j$ is the P.C. approximation to F on $\hat{D} = [a,b] \times Y$.

We wish to know how close are $\hat{T}_A$ and T_A, the P.C. approximation to F on $D = [a,b] \times [c,d]$.

THEOREM 3.3 (Rivlin and Cheney [*13*]). *Let M be a finite-dimensional subspace of $C(D)$ and let F be an element of $C(D)$ which has a unique best approximation $P_{\alpha*}$ in M. For any $Y \subset D$ let $P_{\alpha Y}$ denote a* b.a. *to F from M on the set Y. Then as*

$$\delta_Y = \max_{x \in D} \sigma(x;\, Y) = \max_{x \in D} \inf_{y \in Y} \sigma(x; y) \to 0,$$

$$P_{\alpha_Y} \to P_{\alpha*} \textit{ uniformly.}$$

Given any $\epsilon > 0$, by Theorem 3.3 we can choose $Y \subset [c, d]$ such that δ_Y is sufficiently small to have

$$\max_{y \in [c,d]} |\hat{Q}_{\alpha_j}(y) - Q_{\alpha_j}(y)| < \epsilon \left(n \max_{x \in [a,b]} |\phi_j(x)| \right)^{-1}, \qquad \text{for } j = 1, 2, \ldots, n.$$

Hence,

$$\begin{aligned} \|\hat{T}_A - T_A\|_\infty &= \left\| \sum_{j=1}^{n} (\hat{Q}_{\alpha_j} - Q_{\alpha_j}) \phi_j \right\|_\infty \\ &\leqq \sum_{j=1}^{n} \max_{y \in [c,d]} |\hat{Q}_{\alpha_j}(y) - Q_{\alpha_j}(y)| \max_{x \in [a,b]} |\phi_j(x)| \\ &< \sum_{j=1}^{n} \epsilon \left(n \max_{x \in [a,b]} |\phi_j(x)| \right)^{-1} \max_{x \in [a,b]} |\phi_j(x)| \\ &= \epsilon. \end{aligned}$$

Therefore, if Y is chosen such that δ_Y is sufficiently small, then $\hat{T}_A$ provides a good estimate for T_A.

A Fortran V program for Algorithm 1, using double precision arithmetic, was written for the Univac 1108 computer with the following results (Table 1).

In each example, $D = [-1, 1] \times [-1, 1]$,

$$Y = \{-1.00, -.98, -.96, \ldots, 1.00\}, \qquad \hat{D} = [-1, 1] \times Y,$$

$$\phi_1(x) \equiv 1, \phi_2(x) \equiv x, \ldots, \phi_n(x) \equiv x^{n-1},$$

$$\psi_1(y) \equiv 1, \psi_2(y) \equiv y, \ldots, \psi_m(y) = y^{m-1},$$

and Norm denotes the number

$$\|F - T_A\|_\infty = \max_D |F(x, y) - T_A(x, y)|.$$

For economy of space, Norm is given only for several of the larger choices of n and m.

In Algorithm 1, $P_{\alpha(y)}$, the b.a. to F_y, must be found for each $y \in Y$. By the continuity of $\alpha(\cdot)$, for y sufficiently close to y_1, $P_{\alpha(y)}$ will provide a good initial guess for $P_{\alpha(y_1)}$. More directly connected with the Remez exchange algorithm, we shall show that a set of characteristic extremals for $F_{y_1} - P_{\alpha(y_1)}$ will often provide a good initial guess for a characteristic point set for $F_y - P_{\alpha(y)}$.

DEFINITION 3.4. Let R be the continuous function defined on D by

$$R(x, y) = F_y(x) - P_{\alpha(y)}(x).$$

Hence,

$$\rho(y) = \max_{x \in [a,b]} |R(x, y)|.$$

TABLE 1

$F(x,y) = \dfrac{1}{x+y+3}$

(a) $n = m = 3$, $T_A(x,y) = .33162 - .11665x + .03721x^2 - .11658y + .11235xy - .07081x^2y + .03726y^2 - .07085xy^2 + .06057x^2y^2$, Norm = .46090–01.

(b) $n = m = 4$, Norm = .12637–01.

(c) $n = m = 5$, Norm = .34346–02.

(d) $n = m = 6$, Norm = .91519–03.

$F(x,y) = \dfrac{1}{x+y+10}$, $n = m = 6$, Norm = .19451–07.

$F(x,y) = \sqrt{x+y+3}$

(a) $n = m = 2$, $T_A(x,y) = 1.70563 + .30073x + .30076y - .05701xy$, Norm = .47135–01.

(b) $n = m = 3$, Norm = .64898–02.

(c) $n = m = 4$, Norm = .11031–02.

(d) $n = m = 5$, Norm = .20864–03.

(e) $n = m = 6$, Norm = .42162–04.

$F(x,y) = \sqrt{x+y+10}$, $n = m = 6$, Norm = .17133–07.

$F(x,y) = e^{xy}$

(a) $n = m = 3$, $T_A(x,y) = 1.00140 - .00687x^2 + 1.09731xy - 1.09603y^2 + .55404x^2y^2$, Norm = .83359–01.

(b) $n = m = 4$, Norm = .10830–01.

(c) $n = m = 5$, Norm = .10982–02.

(d) $n = m = 6$, Norm = .91866–04.

$F(x,y) = e^{2xy}$, $n = m = 6$, Norm = .67907–02.

$F(x,y) = \sin(xy)$

(a) $n = m = 2$ and $n = m = 3$, $T_A(x,y) = .91010xy$, Norm = .68634–01.

(b) $n = m = 4$ and $n = m = 5$, $T_A(x,y) = 1.00077xy - .00312x^3y - .00313xy^3 - .15402x^3y^3$, Norm = .96828–03.

(c) $n = m = 6$ and $n = m = 7$, $T_A(x,y) = 1.00000xy + .00002x^3y - .00004x^5y + .00002xy^3 - .16681x^3y^3 + .00029x^5y^3 - .00004xy^5 + .00029x^3y^5 + .00774x^5y^5$, Norm = .59600–05.

$F(x,y) = \sin(2xy)$, $n = m = 6$ and $n = m = 7$, $T_A(x,y) = 1.99977xy + .00195x^3y - .00416x^5y + .00195xy^3 - 1.34949x^3y^3 + .03403x^5y^3 - .00416xy^5 + .03403x^3y^5 + .19606x^5y^5$, Norm = .69531–03.

DEFINITION 3.5. For each $y \in [c,d]$, let

$$E(y) = \{x \in [a,b] : |R(x,y)| = \rho(y)\},$$

the set of all extremals for $F_y - P_{\alpha(y)}$.

$E(y)$ is a non-empty compact subset of $[a,b]$.

DEFINITION 3.6. For each $y \in [c,d]$, and each $\epsilon > 0$, set

$$E(y)_\epsilon = \{x \in [a,b] : |x - e| < \epsilon \quad \text{for some} \quad e \in E(y)\}.$$

THEOREM 3.7. *Given an* $\epsilon > 0$ *and a* $y_1 \in [c,d]$, *there exists a* $\delta = \delta(\epsilon, y_1) > 0$ *such that* $y \in [c,d]$ *and* $|y - y_1| < \delta \Rightarrow E(y) \subset E(y_1)_\epsilon$.

Proof. If $E(y_1)_\epsilon \supset [a,b]$, then the theorem is true trivially. If not, let

$$M = \max_{x \in [a,b] - E(y_1)_\epsilon} |R(x, y_1)| < \rho(y_1) = \rho_1.$$

By the uniform continuity of R on D, there exists a

$$\delta = \delta\left(\frac{\rho_1 - M}{2}\right) > 0,$$

such that

$$|R(x,y) - R(x,y_1)| < \frac{\rho_1 - M}{2},$$

for all $x \in [a,b]$ and all $y \in [c,d]$ with $|y - y_1| < \delta$. Then

$$\max_{x \in [a,b] - E(y_1)_\epsilon} |R(x,y)| < M + \frac{\rho_1 - M}{2} = \frac{M + \rho_1}{2}.$$

Also,

$$x \in E(y_1), \qquad y \in [c,d], \qquad |y - y_1| < \delta$$

$$\Rightarrow |R(x,y)| > |R(x,y_1)| - \frac{\rho_1 - M}{2} = \rho_1 - \frac{\rho_1 - M}{2} = \frac{M + \rho_1}{2},$$

$$\Rightarrow \rho(y) > \frac{M + \rho_1}{2}.$$

Therefore, $x \in [a,b] - E(y_1)_\epsilon \Rightarrow x \notin E(y)$.

COROLLARY. *If* $\phi_1(x) \equiv 1$, $\phi_2(x) \equiv x, \ldots, \phi_n(x) \equiv x^{n-1}$ *and if*

$$\frac{\partial^{n+1}}{\partial x^{n+1}} F(x,y)$$

exists and is $\neq 0$ *throughout* D, *then there are* $n+1$ *continuous functions* $x_1(\cdot), x_2(\cdot), \ldots, x_{n+1}(\cdot)$, *such that for each* $y \in [c,d]$, $x_1(y) < x_2(y) < \ldots < x_{n+1}(y)$ *is the unique set of* $n+1$ *alternating extremals for* $F_y - P_{\alpha(y)}$.

Proof. Since

$$\frac{\partial^{n+1}}{\partial x^{n+1}} F(x,y) = \frac{d^{n+1}}{dx^{n+1}} F_y(x) \neq 0,$$

$F_y - P_{\alpha(y)}$ has exactly $n+1$ extremals for each $y \in [c,d]$. Theorem 3.7 completes the proof.

In general, if $F_y - P_{\alpha(y)}$ has exactly $n+1$ extremals for each $y \in [c,d]$, then the conclusion of the Corollary holds.

Theorem 3.7 and its Corollary suggest that for y sufficiently close to y_1, a set of $n+1$ alternating extremals for $F_{y_1} - P_{\alpha(y_1)}$ might be a good initial guess in the Remez exchange algorithm to find $P_{\alpha(y)}$.

Algorithm 2. (1) Choose some finite point set

$$Y = \{y_1, y_2, \ldots, y_N\} \subset [c,d].$$

(2) Use the Remez exchange algorithm to find $P_{\alpha(y_1)}$, the best uniform approximation in $\langle \phi_1, \phi_2, \ldots, \phi_n \rangle$ to F_{y_1} on $[c,d]$.

(3) Choose a set of $n+1$ alternating extremals for $F_{y_{i-1}} - P_{\alpha(y_{i-1})}$ as an initial guess for a corresponding point set in the Remez exchange algorithm to find $P_{\alpha(y_i)}$, $i = 2, 3, \ldots, N$.

(4) Use the Remez algorithm to find $\hat{Q}_{\alpha_j}$, the best uniform approximation in $\langle \psi_1, \psi_2, \ldots, \psi_m \rangle$ to $a_j(\cdot)$ on Y.

A Fortran V program for Algorithm 2, using double precision arithmetic, was written for the Univac 1108 computer. Several examples were run, using this program and the corresponding program for Algorithm 1. Some of the comparative times are tabulated below. In each case, the resulting approximation was identical for both programs.

As before, in each example $D = [-1,1] \times [-1,1]$,

$$Y = \{-1.00, -.98, \ldots, 1.00\}, \qquad \hat{D} = [-1,1] \times Y,$$

$$\phi_1(x) \equiv 1, \phi_2(x) \equiv x, \ldots, \phi_n(x) \equiv x^{n-1},$$

and

$$\psi_1(y) \equiv 1, \psi_2(y) \equiv y, \ldots, \psi_m(y) \equiv y^{m-1}.$$

The tabulated times are in seconds.

TABLE 2

Function	n	m	Time of Algorithm 1	Time of Algorithm 2
$\frac{1}{x+y+3}$	4	4	46.245	22.712
	5	5	48.321	22.989
	6	6	51.586	25.207
$\sqrt{x+y+3}$	4	4	50.408	23.409
	5	5	52.256	20.062
	6	6	52.619	29.639
e^{xy}	6	6	59.323	46.450
e^{2xy}	6	6	84.194	57.003

4. THE DEGREE OF APPROXIMATION

In this section it is shown that each $F \in C(D)$ can be approximated arbitrarily close by a P.C. approximation of sufficiently high degrees n and m.

DEFINITION 4.1. Let $\mathscr{F}$, $\|\cdot\|$ be a normed linear space. A sequence $\{\phi_j\}$ in $\mathscr{F}$ is called fundamental if and only if the span of $\{\phi_j\}$ is dense in $\mathscr{F}$.

EXAMPLE 3. $1, x, x^2, \ldots$ is fundamental in $C[0,1]$.

EXAMPLE 4. $1, \sin x, \cos x, \sin 2x, \cos 2x, \ldots$ is fundamental in $C_{2\pi}$, the space of continuous functions of one real variable, with period 2π.

DEFINITION 4.2. $\{\phi_i\}$ is called a Markoff system on $[a,b]$ if every initial segment $\{\phi_1, \phi_2, \ldots, \phi_k\}$ is a Chebyshev system on $[a,b]$.

THEOREM 4.3. *Let $\{\phi_j\}$ be a Markoff system on $[a,b]$, fundamental in $C[a,b]$, and let $\{\psi_i\}$ be a Markoff system on $[c,d]$, fundamental in $C[c,d]$. Given an $\epsilon > 0$ and an $F \in C(D)$ (where $D = [a,b] \times [c,d]$) there exists an $N = N(\epsilon)$ and for each $n > N(\epsilon)$ there exists a corresponding $M = M(\epsilon, n)$, such that $n > N$ and $m > M \Rightarrow \|F - T_A\| < \epsilon$, where T_A is the* P.C. *approximation to F on D, and $\|\cdot\|$ is the uniform norm on D.*

Proof.

$$\max_D |F(x,y) - P_{\alpha(y)}(x)| = \max_{y \in [c,d]} \{ \max_{x \in [a,b]} |F_y(x) - P_{\alpha(y)}(x)| \} = \max_{y \in [c,d]} \rho(y).$$

By the uniform continuity of $\rho(\cdot)$ on $[c,d]$, there exists a $\delta > 0$, such that

$$y_1, y_2 \in [c,d] \quad \text{and} \quad |y_2 - y_1| < \delta \Rightarrow \rho(y_1) < \rho(y_2) + \epsilon/4.$$

Let k be the largest integer j such that $2j\delta \leqq d - c$. Let $Y = \{y_0, y_1, \ldots, y_k\}$, where $y_j = C + 2j\delta, j = 0, 1, \ldots, k$.

Since $\{\phi_j\}$ is fundamental in $C[a,b]$, given any $\epsilon > 0$, for each $i = 1, 2, \ldots, k$, there exists an N_i such that

$$n > N_i \Rightarrow \rho(y_i) = \max_{x \in [a,b]} |F_{y_i}(x) - P_{\alpha(y_i)}(x)| < \epsilon/4,$$

where $P_{\alpha(y_i)}(\cdot)$ is the b.a. in $\langle \phi_1, \phi_2, \ldots, \phi_n \rangle$ to $F_{y_i}(\cdot)$ on $[a,b]$.

Set $N = \max(N_1, N_2, \ldots, N_k)$. Then

$$n > N \Rightarrow \rho(y_i) < \epsilon/4, \qquad \text{for } i = 1, 2, \ldots, k.$$

Given any $y \in [c,d]$, there exists a corresponding $y_i \in Y$ such that $|y - y_i| < \delta$, which implies

$$\rho(y) < \rho(y_i) + \epsilon/4 < \epsilon/2.$$

Therefore,

$$n > N \Rightarrow \max_{y \in [c,d]} \rho(y) < \epsilon/2.$$

Since $\phi_1, \phi_2, \ldots, \phi_n$ is a Chebyshev system on $[a,b]$, $\|\phi_i\| > 0$ for $j = 1, 2, \ldots, n$.

Since $\{\psi_i\}$ is fundamental in $C[c,d]$, given any $\epsilon > 0$, for each $j = 1, 2, \ldots, n$ there exists an M_j such that

$$m > M_j \Rightarrow \max_{y \in [c,d]} |a_j(y) - Q_{\alpha_j}(y)| < \epsilon/[2n\|\phi_j\|],$$

where $Q_{\alpha_j}(\cdot)$ is the best uniform approximation in $\langle \psi_1, \psi_2, \ldots, \psi_m \rangle$ to $a_j(\cdot)$ on $[c,d]$.

Set $M = \max(M_1, M_2, \ldots, M_n)$. Then

$$m > M \Rightarrow \max_D |P_{\alpha(y)}(x) - T_A(x,y)| = \max_D \left| \sum_{j=1}^{n} (a_j(y) - Q_{\alpha_j}(y))\, \phi_j(x) \right|$$

$$\leqq \sum_{j=1}^{n} \max_{y \in [c,d]} |a_j(y) - Q_{\alpha_j}(y)| \max_{x \in [a,b]} |\phi_j(x)|$$

$$\leqq \sum_{j=1}^{n} (\epsilon/[2n\|\phi_j\|])\, \|\phi_j\| = \frac{\epsilon}{2}.$$

Therefore,

$$n > N \quad \text{and} \quad m > M \Rightarrow \|F - T_A\| = \max_D |F(x,y) - T_A(x,y)|$$
$$\leqq \max_D |F(x,y) - P_{\alpha(y)}(x)| + \max_D |P_{\alpha(y)}(x) - T_A(x,y)|$$
$$< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

COROLLARY. *Under the hypotheses of Theorem* 4.3, *there exists an* $N = N(\epsilon)$ *and for each* $n > N(\epsilon)$ *there exists a corresponding* $M = M(\epsilon, n)$ *such that* $n > N$ *and* $m > M \Rightarrow \|T_A - P_{A*}\| < \epsilon$, *where*

$$P_{A*} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \phi_j \psi_i$$

is a b.a. *to* F *on* D.

Proof. Let N and M be, respectively, N and M of Theorem 4.3, but with ϵ replaced by $\epsilon/2$. Then

$$n > N \quad \text{and} \quad m > M \Rightarrow \|T_A - P_{A*}\| \leqq \|F - T_A\| + \|F - P_{A*}\|$$
$$\leqq 2\|F - T_A\|$$
$$< 2\frac{\epsilon}{2} = \epsilon.$$

At the present time there is no known effective scheme for computing a best uniform approximation to a function of two variables. Present research is being directed towards schemes somewhat like the Remez exchange algorithm. A good initial guess is needed if such an iterative procedure is to converge, and if the computation time is to be reasonably short. Theorem 4.3 and its Corollary suggest that the product Chebyshev approximation might be a good initial guess for a best uniform approximation.

5. THE PRODUCT CHEBYSHEV APPROXIMATION TO A CONTINUOUS FUNCTION OF THREE OR MORE VARIABLES

For simplicity of notation, attention is restricted to the case of three variables.

Let $D = [a,b] \times [c,d] \times [e,f]$, where $a < b$, $c < d$, and $e < f$, and let $F \in C(D)$.

For each $(y,z) \in [c,d] \times [e,f]$, define the continuous function $F_{y,z}$ on $[a,b]$ by

$$F_{y,z}(\cdot) = F(\cdot, y, z).$$

Let $\phi_1, \phi_2, \ldots, \phi_n$ be a Chebyshev system of continuous real-valued functions on $[a,b]$. Let $\|\cdot\|_p$ be any one of the L_p norms, $1 \leqq p \leqq \infty$.

For each $(y,z) \in [c,d] \times [e,f]$, there exists a unique "polynomial"

$$P_{\alpha(y,z)} = \sum_{j=1}^{n} a_j(y,z)\,\phi_j, \qquad \alpha(y,z) = (a_1(y,z), a_2(y,z), \ldots, a_n(y,z)) \in E_n$$

which is the best $\|\cdot\|_p$ approximation to $F_{y,z}$ on $[a,b]$.

For each $(y,z) \in [c,d] \times [e,f]$, let

$$\rho(y,z) = \min_{\alpha \in E_n} \|F_{y,z} - P_\alpha\|_p.$$

The following extensions of Theorems 2.1 and 2.3 are straightforward.

THEOREM 5.1. *$\rho(y,z)$ is continuous on $[c,d] \times [e,f]$.*

THEOREM 5.2. *$\alpha(y,z)$ is continuous on $[c,d] \times [e,f]$.*

Let $\{\psi_i\}$ and $\{\theta_k\}$ be Chebyshev systems of continuous real-valued functions on $[c,d]$ and $[e,f]$, respectively.

DEFINITION 5.3. Let T_{α_j} be the product L_p approximation to $a_j(\cdot)$ on $[c,d] \times [e,f]$ relative to z, for $j = 1,2,\ldots,n$.

The function $T_A = \sum_{j=1}^{n} T_{\alpha_j}\phi_j$ is called the product L_p approximation to F on D, relative to y, z.

The product L_p approximation depends on the order in which the variables are specified, as in Example 1 of Section 3.

The product L_p approximation is extended to continuous functions F defined on a more general set D, in [*18*].

ACKNOWLEDGMENTS

The author thanks Professor David G. Moursund, his thesis advisor, for helpful guidance during the preparation of this work, which is an essential portion of the author's doctoral dissertation at Michigan State University.

Appreciation is also expressed to the staff of the University of Utah Computer Center for the computing facilities used in calculating the examples in this paper.

REFERENCES

1. N. I. ACHIESER, "Theory of Approximation" (English translation). Frederick Ungar, New York, 1956.

2. E. W. CHENEY, "Introduction to Approximation Theory." McGraw-Hill, New York, 1966.

3. P. J. DAVIS, "Interpolation and Approximation." Blaisdell, New York, 1963.

4. W. FRASER AND J. F. HART, On the computation of rational approximations to continuous functions. *Comm. ACM.* **5** (1962), 401–403.
5. A. HAAR, Die Minkowskische Geometrie und die Annäheruug an stetige Funktionen. *Math. Ann.* **18** (1918), 294–311.
6. G. G. LORENTZ, "Approximation of Functions." Holt, Rinehart and Winston, New York, 1966.
7. JOHN C. MAIRHUBER, On Haar's theorem concerning Chebychev approximation problems having unique solutions. *Proc. Amer. Math. Soc.* **7** (1956), 609–615.
8. E. P. NOVODVORSKII AND I. S. PINSKER, The process of equating maxima. *Uspehi Mat. Nauk.* **6** (1951), 174–181.
9. E. YA. REMEZ, "General Computational Methods of Chebyshev Approximation" (English translation). United States Atomic Energy Commission, Washington, D.C., 1962.
10. E. YA. REMEZ, Sur le calcul effectif des polynomes d'approximation de Tchebichef. *Compt. Rend. Acad. Sci., Paris,* **199** (1934), 337–340.
11. J. R. RICE, "The Approximation of Functions," Vol. 1. Addison-Wesley, Reading, Mass., 1964.
12. J. R. RICE, Tchebycheff approximation in several variables. *Trans. Amer. Math. Soc.* **109** (1963), 444–466.
13. T. J. RIVLIN AND E. W. CHENEY, A comparison of uniform approximations on an interval and a finite subset thereof. *SIAM J. Numer. Anal.* **3** (1966), 311–320.
14. T. J. RIVLIN AND H. S. SHAPIRO, Some uniqueness problems in approximation theory, *Commun. Pure Appl. Math.* **13** (1960), 35–47.
15. T. J. RIVLIN AND H. S. SHAPIRO, A unified approach to certain problems of approximation and minimization. *J. Soc. Ind. Appl. Math.* **9** (1961), 670–699.
16. I. J. SCHOENBERG, On the question of unicity in the theory of best approximation. *Ann. N.Y. Acad. Sci.* **86** (1960), 682–692.
17. L. VEIDINGER, On the numerical determination of the best approximations in the Chebyshev sense. *Numer. Math.* **2** (1960), 99–105.
18. STANLEY E. WEINSTEIN, Product approximations of functions of several variables. (In preparation).