APPROXIMATIONS OF FUNCTIONS OF SEVERAL VARIABLES: PRODUCT TCHEBYCHEFF APPROXIMATIONS

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

STANLEY EDWIN WEINSTEIN

1967

This is to certify that the

thesis entitled

APPROXIMATIONS OF FUNCTIONS OF SEVERAL VARIABLES: PRODUCT TCHEBYCHEFF APPROXIMATIONS

presented by

Stanley Edwin Weinstein

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Mathematics

David Moursund
Major professor

Date 8/1/67

ABSTRACT

APPROXIMATIONS OF FUNCTIONS OF SEVERAL VARIABLES: PRODUCT TCHEBYCHEFF APPROXIMATIONS

by Stanley Edwin Weinstein

In this thesis we define a Tchebycheff-like approximation to a continuous function F of two or more variables, which we call the Product Tchebycheff approximation, and which possesses the desired property of uniqueness.

Let D be a compact set in E_2 and let F be continuous on D. Let D_x and D_y denote the projections of D onto the x-axis and onto the y-axis respectively and let $D_x(y_1) = [x \varepsilon D_x (x,y_1) \varepsilon D]$. Let $\Phi_1, \Phi_2, \ldots, \Phi_n$ be a Tchebycheff system of continuous real-valued functions on D_x. For each $y \varepsilon D_y$ we define the continuous function F_{y_1} by

$$F_{y_1}(x) = F(x,y_1)$$

Then F_{y_1} possesses a unique best uniform approximation in $\Phi_1, \Phi_2, \ldots, \Phi_n$, namely

$$P_{A(y_1)} = \sum_{j=1}^{n} a_j(y_1)\, \Phi_j$$

We define a restriction on the set D, which we call

property K. This class of sets includes all compact convex sets as well as all sets which are cross products of compact sets.

If D possesses property K then the deviation function ρ defined on D_y by

$$\rho(y) = \max_{x \varepsilon D_x(y)} | F_y(x) - P_{A(y)}(x)|$$

is continuous. Moreover if we further restrict the point $y_1 \varepsilon D_y$ such that the set $D_x(y_1)$ contains n or more points then the coeficients $a_1(y), a_2(y), \ldots, a_n(y)$ are all continuous at y_1.

We define the Product Tchebycheff approximation on sets which possess property K and such that for each $y \varepsilon D_y$ $D_x(y)$ contains n or more points. Let $\psi_1, \psi_2, \ldots, \psi_m$ be a Tchebycheff system on D_y. Then each Function a_j possesses a unique best uniform approximation

$$Q_j = \sum_{i=1}^{m} a_{ij} \psi_i \; , \quad j = 1, 2, \ldots, n.$$

The polynomial

$$PT = \sum_{j=1}^{n} Q_j \Phi_j = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij} \psi_i \Phi_j$$

is called the Product Tchebycheff approximation to F on D, relative to the variable y.

Two algorithms for the computation of this approximation are presented along with several examples.

It is also shown that by a suitable choice of base functions, the resulting Product Tchebycheff approximation will approximate F arbitrarily close.

The Product Tchebycheff approximation is extended to continuous functions of three or more variables.

APPROXIMATIONS OF FUNCTIONS OF SEVERAL VARIABLES:
PRODUCT TCHEBYCHEFF APPROXIMATIONS

By

Stanley Edwin Weinstein

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

1967

346738
12-8-67

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to Dr. David G. Moursund for his patience and guidance in the preparation of this thesis. The numerous discussions we had during the research was of an immeasurable value to me. I would also like to thank Dr. G. D. Taylor and Dr. C. S. Duris, both of whom have contributed several helpful suggestions.

I acknowledge the support of Michigan State University who made available to me the computing facilities used in the calculation of the examples in this thesis.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION: THE PROBLEM OF THE LACK OF UNIQUENESS OF A BEST APPROXIMATION

Section 1: The Best Approximation Problem

The central theme of this thesis is the problem of finding a best approximation to an arbitrary element F in a normed linear space $\mathcal{F},||\cdot||$. This can be formulated as follows:

Definition 1.1.1: Let $\mathcal{F},||\cdot||$ be a normed linear space, and let $\phi_1,\phi_2,\ldots,\phi_n$ be a set of n linearly independent elements of $\mathcal{F}$. Given an arbitrary element $F\varepsilon\mathcal{F}$, the problem of finding an $A^* = (a_1^*,a_2^*,\ldots,a_n^*)\varepsilon E_n$, such that

$$||F - \sum_{i=1}^{n} a_i^*\phi_i|| \leq ||F - \sum_{i=1}^{n} a_i\phi_i||$$

$$\text{for all } A = (a_1,a_2,\ldots,a_n)\varepsilon E_n,$$

is called the best approximation (b.a.) problem. Both the point $A^*\varepsilon E_n$, and the corresponding polynomial

$$P_{A^*} = \sum_{i=1}^{n} a_i^* \phi_i$$

are referred to as best $||\cdot||$ approximations to F.

<u>Definition 1.1.2</u>: If $\phi_1, \phi_2, \ldots, \phi_n$ are n linearly independent elements of the linear space $\mathcal{F}$, then the set of all linear combinations

$$P_A = \sum_{i=1}^{n} a_i \phi_i \quad ; \quad A = (a_1, a_2, \ldots, a_n) \varepsilon E_n$$

is a linear subspace of $\mathcal{F}$, of dimension n, called the space spanned by $\phi_1, \phi_2, \ldots, \phi_n$, or the span of $\phi_1, \phi_2, \ldots, \phi_n$, which is denoted by

$$< \phi_1, \phi_2, \ldots, \phi_n > .$$

We now restate the b.a. problem with the use of definition 1.1.2.

<u>B. A. Problem (Restated) 1.1.3</u>: Find the point(s) of the subspace $< \phi_1, \phi_2, \ldots, \phi_n >$ which are closest to $F \varepsilon \mathcal{F}$, where the distance between any two elements $F_1, F_2 \varepsilon \mathcal{F}$ is defined by

$$d(F_1, F_2) = ||F_1 - F_2|| \quad .$$

In this thesis, we are primarily concerned with the following case of the b. a. problem:

Let D be a compact set in euclidean k-dimensional space E_k, and let $\mathcal{F}$ = C(D), the linear space of all continuous, real-valued functions on D, with $||\cdot||$ defined on C(D) by

$$||F|| = \sup_{x \varepsilon D} |F(x)| \qquad \text{for each } F \varepsilon C(D).$$

This norm is called the uniform, Tchebycheff or minimax norm. A corresponding solution to the b.a. problem is called a best uniform, best Tchebycheff or best minimax approximation to $F \varepsilon C(D)$ on D.

The existance of a solution to this b.a. problem is a direct implication of the existance of a solution to 1.1.3; see Davis (5).

We now consider the question of the uniqueness of the solution to this b.a. problem.

Section 2: Tchebycheff Systems, Haar's Theorem

Definition 1.2.1: The family of real-valued functions $\phi_1, \phi_2, \ldots, \phi_n$, is called a Tchebycheff system on D if and only if for $A = (a_1, a_2, \ldots, a_n) \varepsilon E_n$ the polynomial

$$P_A = \sum_{i=1}^{n} a_i \phi_i$$

vanishes in at most n-1 distinct points of D unless

$$a_1 = a_2 = \dots = a_n = 0.$$

Definition 1.2.1 is equivalent to saying that $\phi_1, \phi_2, \dots, \phi_n$ is a Tchebycheff system on D, if and only if the determinant

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

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

<u>Remark 1.2.2</u>: Let $D_1 \subset D_2$; then if $\phi_1, \phi_2, \dots, \phi_n$ is a Tchebycheff system on D_2 it is also a Tchebycheff system on D_1.

<u>Remark 1.2.3</u>: The fact that $\phi_1, \phi_2, \dots, \phi_n$ is a Tchebycheff system on D, implies that $\phi_1, \phi_2, \dots, \phi_n$ are linearly independent on D.

<u>Example 1.2.4</u>: Let $D = [a,b]$, $a \leq b$ and let

$$\begin{aligned} \phi_1(x) &= 1 \\ \phi_2(x) &= x \\ \phi_3(x) &= x^2 \\ &\vdots \\ \phi_n(x) &= x^{n-1}. \end{aligned}$$

Then, for each $A = (a_1, a_2, \ldots, a_n) \varepsilon E_n$ the polynomial P_A defined on $[a,b]$ by

$$P_A(x) = \sum_{i=1}^{n} a_i \phi_i(x) = a_1 + \sum_{i=2}^{n} a_i x^{i-1}$$

has at most n-1 distinct zeros on $[a,b]$ unless

$$P_A(x) \equiv 0$$

or equivalently, unless

$$a_1 = a_2 = \ldots = a_n = 0.$$

Therefore, $\phi_1, \phi_2, \ldots, \phi_n$ is a Tchebycheff system on D.

Example 1.2.5: Let D = the x-axis modulo 2π, or equivalently, let $D = [0, 2\pi]$ with the points 0 and 2π identified as one point, and let

$$\begin{aligned}
\phi_1(x) &= 1 \\
\phi_2(x) &= \cos x \\
\phi_3(x) &= \sin x \\
&\vdots \\
\phi_{2n}(x) &= \cos nx \\
\phi_{2n+1}(x) &= \sin nx \quad .
\end{aligned}$$

Then, for each $A = (a_1, a_2, \ldots, a_{2n+1}) \varepsilon E_{2n+1}$ the polynomial P_A defined on D by

$$P_A(x) = \sum_{i=1}^{2n+1} a_i \phi_i(x) = a_1 + a_2 \cos x + a_3 \sin x + \ldots$$

$$+ a_{2n} \cos nx + a_{2n+1} \sin nx$$

has at most 2n distinct zeros on D, unless

$$P_A(x) \equiv 0$$

or equivalently unless

$$a_1 = a_2 = \ldots = a_{2n+1} = 0.$$

Therefore, $\phi_1, \phi_2, \ldots, \phi_{2n+1}$ is a Tchebycheff system on D.

The next example illustrates that whether or not a set of functions $\phi_1, \phi_2, \ldots, \phi_n$ is a Tchebycheff system on a set D, may depend upon the set D.

Example 1.2.6: Let $D_1 = [0,1]$ and let

$$\phi_1(x) = 1$$
$$\phi_2(x) = x^2.$$

Then, for $A = (a_1, a_2) \varepsilon E_2$

$$P_A(x) = a_1\phi_1(x) + a_2\phi_2(x) = a_1 + a_2x^2$$

has at most one distinct zero in D_1, unless $a_1 = a_2 = 0$. Thus, ϕ_1,ϕ_2 is a Tchebycheff system on D_1.

However, if $D_2 = [-1,1]$, then $P_A(x) = \frac{1}{4} - x^2$ vanishes at both $x_1 = \frac{1}{2}$ and $x_2 = -\frac{1}{2}$. Thus, ϕ_1,ϕ_2 is not a Tchebycheff system on D_2.

Now, let $\phi_1,\phi_2,\ldots,\phi_n$ be n real-valued continuous functions on a compact set D, (which contains at least n distinct points) in E_k.

Haar (8) proved the following theorem which helped to answer the question of the uniqueness of a solution to the b.a. problem in the above setting.

<u>Theorem 1.2.7 (Haar)</u>: A necessary and sufficient condition that a unique solution exist to the b.a. problem for every given $F \varepsilon C(D)$ is that the base functions $\phi_1,\phi_2,\ldots,\phi_n$ form a Tchebycheff system on D.

Thus if $\phi_1,\phi_2,\ldots,\phi_n$ is not a Tchebycheff system on D, then for some $F \varepsilon C(D)$, the corresponding b.a. problem has a non-unique solution. However, as is seen in the following example, some $F \varepsilon C(D)$ may have a unique best approximation on D.

<u>Example 1.2.8</u>: Let $D = [-1,1]$ and let

$$\phi_1(x) = 1$$

$$\phi_2(x) = x^2.$$

Then, as in example 1.2.6, ϕ_1, ϕ_2 is not a Tchebycheff system on D. However the function $F \varepsilon C(D)$ defined by

$$F(x) = 1 + x^2$$

has a unique best uniform approximation in ϕ_1, ϕ_2, namely P_A* which is defined by

$$P_{A*}(x) = 1 + x^2 \quad .$$

We now state two corollaries to Haar's theorem.

Corollary 1.2.9: Let $D = [a,b]$, $a < b$ and let

$$\phi_1(x) = 1$$
$$\phi_2(x) = x$$
$$\vdots$$
$$\phi_n(x) = x^{n-1} \quad .$$

Then as in example 1.2.4, $\phi_1, \phi_2, \ldots, \phi_n$ is a Tchebycheff system on D and thus by Haar's theorem, for each $F \varepsilon C(D)$ there exists a unique algebraic polynomial

$$P_A* = a_1*\phi_1 + a_2*\phi_2 + \ldots + a_n*\phi_n$$

which minimizes the quantity

$$\max_{x \varepsilon [0,1]} |F(x) - P_A(x)| \quad .$$

Corollary 1.2.10: Let D = the x-axis modulo 2π and let

$$\begin{aligned}
\phi_1(x) &= 1 \\
\phi_2(x) &= \cos x \\
\phi_3(x) &= \sin x \\
&\vdots \\
\phi_{2n}(x) &= \cos nx \\
\phi_{2n+1}(x) &= \sin nx \quad .
\end{aligned}$$

Then, if F is an arbitrary continuous function with period 2π, (that is, $F(x + 2\pi) = F(x)$, for all real x), there exists a unique trigonometric polynomial

$$P_A^* = a_1^*\phi_1 + a_2^*\phi_2 + \ldots + a^*_{2n+1}\phi_{2n+1}$$

which minimizes the quantity

$$\max_{x\varepsilon D} |F(x) - P_A(x)| \quad .$$

The following example (see Buck (2)) presents a function which possesses a non-unique, best uniform approximation.

Example 1.2.11: Consider the Banach space of continuous real-valued function on the compact set $D = [0,1] \times [0,1] \subset E_2$, with $||\cdot||$, the uniform norm.

Let $\phi_1(x,y) = 1$

$\phi_2(x,y) = x$

$\phi_3(x,y) = y$

$\phi_4(x,y) = x^2$

$\phi_5(x,y) = y^2$,

and let F be defined by

$$F(x,y) = xy.$$

Using results from functional analysis, Buck first shows that

$$\min_{A\varepsilon E_5} ||F - P_A|| \geq \frac{1}{4}$$

where as usual for each $A = (a_1,a_2,a_3,a_4,a_5)\varepsilon E_5$

$$P_A = a_1\phi_1 + a_2\phi_2 + a_3\phi_3 + a_4\phi_4 + a_5\phi_5 \quad .$$

He then observes that

$$||xy - (\tfrac{1}{2}(x^2 + y^2) - \tfrac{1}{4})|| = \tfrac{1}{4}$$

and

$$||xy - (x + y - \tfrac{1}{2}(x^2 + y^2) - \tfrac{1}{4})|| = \tfrac{1}{4} \quad .$$

Therefore, $P_{A_1}{}^*$ and $P_{A_2}{}^*$ defined by

$$P_{A_1}^*(x) = \frac{1}{2}(x^2 + y^2) - \frac{1}{4}$$

and

$$P_{A_2}^*(x) = x + y - \frac{1}{2}(x^2 + y^2) - \frac{1}{4}$$

are both solutions to the b.a. problem. This in turn implies that the b.a. problem has an infinite number of solutions, since if $0 \leq \alpha \leq 1$, then

$$||xy - (\alpha P_{A_1}^* + (1-\alpha)P_{A_2}^*)||$$

$$\leq ||\alpha(xy - P_{A_1}^*)|| + ||(1-\alpha)(xy - P_{A_2}^*)||$$

$$= \alpha||xy - P_{A_1}^*|| + (1-\alpha)||xy - P_{A_2}^*||$$

$$= \alpha \frac{1}{4} + (1-\alpha) \frac{1}{4} = \frac{1}{4} .$$

Hence, $\alpha P_{A_1}^* + (1-\alpha)P_{A_2}^*$ is a solution to the b.a. problem for all α such that $0 \leq \alpha \leq 1$.

Haar has shown that if the domain D contains a triod set T (see Figure 1.2.1), then there does not exist a non-trivial ($n \geq 2$) Tchebycheff system on D_1 (see Lorentz (9), p. 25).

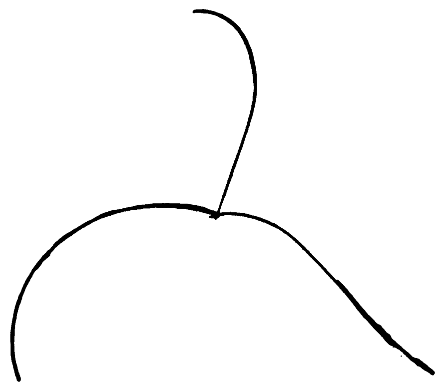


Figure 1.2.1

Thus for any non-trivial base system $\phi_1, \phi_2, \ldots, \phi_n$, the b.a. problem does not have a unique solution for some $F \epsilon C(D)$.

Mairhuber (1 0) established necessary and sufficient conditions for a set D to serve as the domain of definition of a Tchebycheff system.

<u>Theorem 1.2.12 (Mairhuber)</u>: A compact subset D of E_k containing at least n points, $n \geq 2$, may serve as the domain of definition of a non-trivial Tchebycheff system of real-valued continuous functions, if and only if D is homeomorphic to a closed subset of the circumference of a circle.

Therefore, except for trivial Tchebycheff systems ($n = 1$) or D homeomorphic to a closed subset of the circumference of a circle, the b.a. problem has a non-unique solution for some $F \epsilon C(D)$.

It is desirable to have a unique best approximation for all $F \epsilon C(D)$. The primary virtue is one of communication.

We can speak of "the" best uniform approximation or "the" least squares approximation of degree n to $F \epsilon C(D)$. Secondly, it is often true that the property of uniqueness will simplify the algorithm(s) used to solve the b.a. problem.

The primary goal of this thesis is to define and investigate a Tchebycheff-like approximation which possesses the desired property of uniqueness for all $F \epsilon C(D)$. We shall call this approximation the Product Tchebycheff Approximation to F on D.

CHAPTER II

THE PRODUCT TCHEBYCHEFF APPROXIMATION TO A CONTINUOUS FUNCTION

Section 1: Introduction

Let D be a compact set in E_2. We define each of the following compact sets in E_1:

D_x, the projection of D onto the x-axis,

D_y, the projection of D onto the y-axis,

$D_x(y_1)$, the projection of the intersection of the set D and the line $y = y_1$, onto the x-axis.

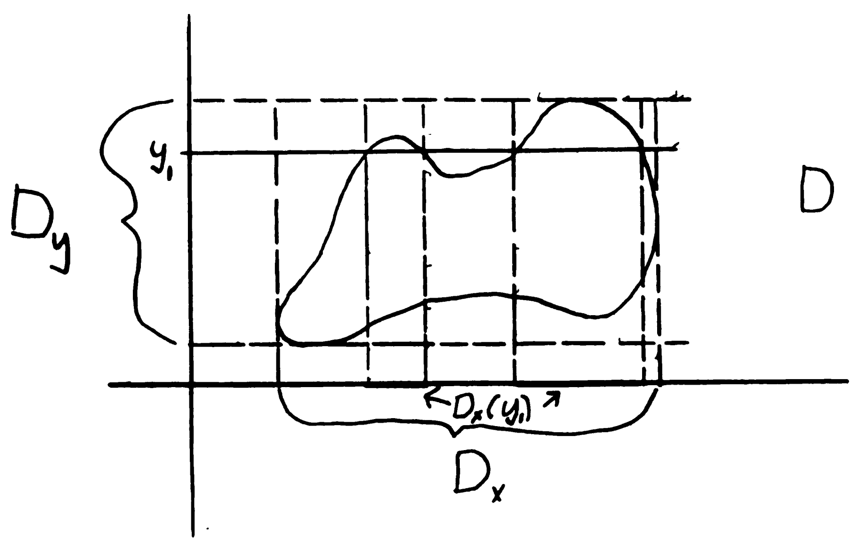


Figure 2.2.1

Let $F \varepsilon C(D)$, and we define the following associated family of functions:

<u>Definition 2.1.1</u>: For each $y_1 \varepsilon D_y$ define the function F_{y_1} on $D_x(y_1)$, by

$$F_{y_1}(x) = F(x,y_1).$$

We also define a family of norms.

<u>Definition 2.1.2</u>: For each $y_1 \varepsilon D_y$, and for each real-valued function H defined on $D_x(y_1)$ we define

$$||H||_{y_1} = \sup_{x \varepsilon D_x(y_1)} |H(x)| \quad .$$

Now, let $\phi_1, \phi_2, \ldots, \phi_n$ be a Tchebycheff system of continuous functions on D_x.

<u>Definition 2.1.3</u>: For each $y_1 \varepsilon D_y$, such that $D_x(y_1)$ contains at least n points, Haar's Theorem shows the existence of a unique polynomial

$$P_{A(y_1)} = \sum_{i=1}^{n} a_i(y_1)\phi_i$$

(where $A(y_1) = (a_1(y_1), a_2(y_1), \ldots, a_n(y_1)) \varepsilon E_n$)

which is the best $||\cdot||_{y_1}$ approximation to the continuous function F_{y_1} on $D_x(y_1)$. As usual both $P_{A(y_1)}$ and the associated parameter $A(y_1)$ will be referred to as the best $||\cdot||_{y_1}$ approximation to F_{y_1}.

The basic idea of this thesis, is to obtain polynomial approximations to the coefficient functions $a_1(y), a_2(y), \ldots, a_n(y)$, for $y \varepsilon D_y$.

We will consider the continuity of $a_1(y), a_2(y), \ldots, a_n(y)$ on D_y under an additional restriction on the set D, which we shall call property K. If these functions are continuous on D_y, they themselves will have best uniform approximations by polynomials in $\psi_1, \psi_2, \ldots, \psi_m$, a family of m real-valued continuous functions on the compact set D_y.

Section 2: Property K

Definition 2.2.1: A compact set D is said to possess property K relative to the variable y, if given $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon) > 0$, such that $y_1, y_2 \varepsilon D_y$, and $|y_1 - y_2| < \delta$ implies that if $(x_1, y_1) \varepsilon D$ $(x_1 \varepsilon D_x(y_1))$ then there exists an $x_2 \varepsilon D_x$ where $|x_2 - x_1| < \varepsilon$ and $(x_2, y_2) \varepsilon D$ $(x_2 \varepsilon D_x(y_2))$.

Remark 2.2.2: The following definition is equivalent to Definition 2.2.1.

(2.2.2) A compact set D is said to possess property K relative to the variable y if given $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon) > 0$, such that $y_1, y_2 \varepsilon D_y$, and $|y_1 - y_2| < \delta$ implies that if $(x_1, y_1) \varepsilon D$ $(x_1 \varepsilon D_x(y_1))$ then there exists an $x_2 \varepsilon D_x$ where $\sigma[(x_1, y_1);(x_2, y_2)] < \varepsilon$ and $(x_2, y_2) \varepsilon D$ $(x_2 \varepsilon D_x(y_2))$.

Note that here and throughout the remainder of this thesis the letter "σ" will be used to denote the usual Euclidean metric on the space under consideration at that time.

σ_n, will denote the usual Euclidean metric on Euclidean n space E_n, for any positive integer n.

<u>Proof</u>: If as in Definition 2.2.2

$$\sigma[(x_1,y_1);(x_2,y_2)] < \varepsilon$$

then $|x_2 - x_1| \leq \sigma[(x_1,y_1);(x_2,y_2)] < \varepsilon$ which satisfies 2.2.1.

On the other hand if $y_1,y_2 \varepsilon D_y$ with $|y_1 - y_2| < \delta_1(\varepsilon/2)$ as in Definition 2.2.1, we can choose

$$\delta = \min(\delta_1(\varepsilon/2), \varepsilon/2) \quad .$$

Hence $y_1,y_2 \varepsilon D_y$ and $|y_1 - y_2| < \delta$ implies that there exists an $x_2 \varepsilon D_x(y_2)$ where

$$\sigma[(x_1,y_1);(x_2,y_2)] \leq |x_2 - x_1| + |y_2 - y_1|$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

which satisfies 2.2.2.

We may similarly define property K relative to the variable x. We shall see in Example 2.2.5 that property K relative to x and property K relative to y are not equivalent.

<u>Example 2.2.3</u>: Any rectangle D = $\{(x,y): a \leq x \leq b, c \leq y \leq d\}$ has property K relative to the variable y as well as property K relative to the variable x.

<u>Example 2.2.4</u>: Let D = the intersection of the horizontal band

$$H_{a,b} = \{(x,y): a \leq y \leq b\}$$

with a family of a finite number of non-horizontal lines.

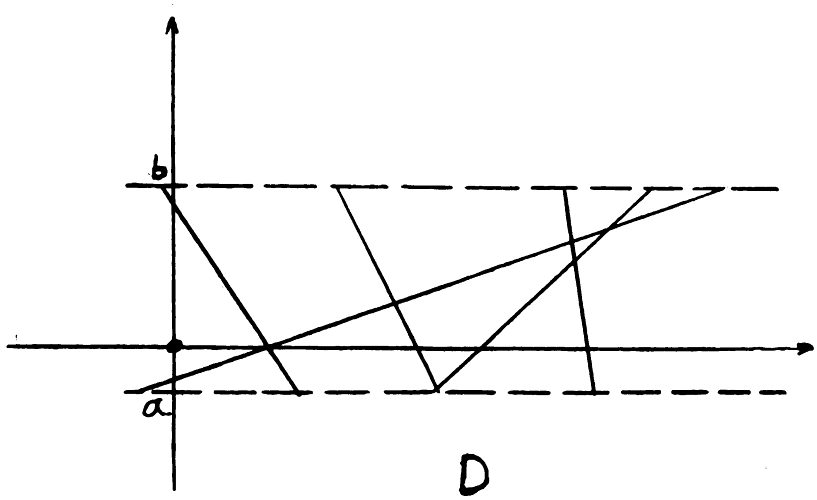


Figure 2.2.1

Then D has property K relative to the variable y.

Note that as a special case of this any single line segment (we need not exclude a horizontal line segment) has property K relative to both x and y.

<u>Example 2.2.5</u>: Let $D = \{(x,y): x + y = 1 \text{ and } 0 \leq x \leq 1$ or $x = 0 \text{ and } 0 \leq y \leq 1\}$.

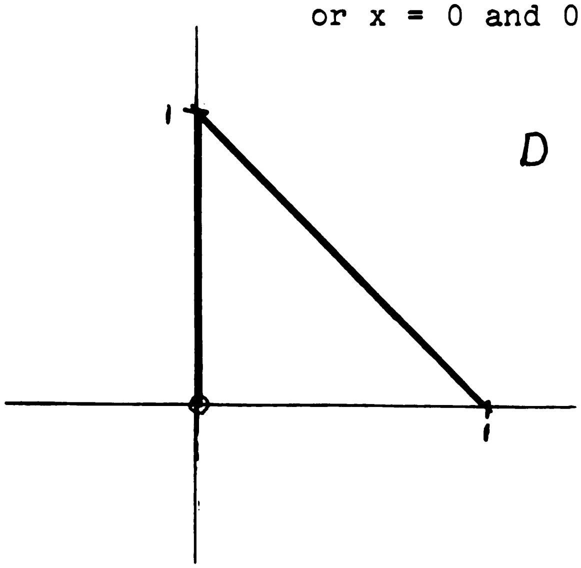


Figure 2.2.2

Then given $\varepsilon > 0$, if $(x_1,y_1)\varepsilon D$ then

(1) $(x_1,y_1) = (0,y_1)$ for some $0 \leq y_1 \leq 1$

or (2) $(x,y_1) = (1 - y_1,y_1)$ for some $0 \leq y_1 \leq 1$

In case (1) Let $\delta = \delta(\varepsilon) = 1$ then for any $y_2 \varepsilon D_y = [0,1]$, $|y_2 - y_1| < \delta = 1$ and $(0,y_2)\varepsilon D$ implies $|0 - 0| = 0$ which satisfies 2.2.1.

In case (2) let $\delta = \delta(\varepsilon) = \varepsilon$ then for $y_2 \varepsilon D_y = [0,1]$ if $|y_2 - y_1| < \delta = \varepsilon$ then for $x_2 = 1 - y_2$

$$|x_2 - x_1| = |1 - y_2 - (1 - y_1)| = |y_1 - y_2| < \varepsilon$$

and definition 2.2.1 is satisfied.

Thus if we choose $\delta = \min(1,\varepsilon)$, Definition 2.2.1 is satisfied relative to the variable y.

However, if we choose $\varepsilon = \frac{1}{2}$ and $(x_1,y_1) = (0,0)$, then $x_2 \varepsilon D_x$ and $|x_2 - x_1| = x_2 < \frac{1}{2}$ implies that if $x_2 \neq 0$, then there is one and only one corresponding y_2, such that $(x_2,y_2)\varepsilon D$, namely $y_2 = 1 - x_2$, and therefore, $|y_2 - y_1| = y_2 = 1 - x_2 > \frac{1}{2}$. Hence, D does not possess property K, relative to the variable x.

Theorem 2.2.6: Let K_1 and K_2 be compact sets in E_1. Then, $D = K_1 \times K_2 \subset E_2$ has property K relative to both x and y.

Proof: Clearly D_1 which is the cross-product of two compact sets, is itself compact in E_2.

Now given $\varepsilon > 0$ and an arbitrary point $(x_1,y_1)\varepsilon D$, then $x_1 \varepsilon K_1$ and $y_1 \varepsilon K_2$. For any $y_2 \varepsilon D_y = K_2$ we have the point $(x_1,y_2)\varepsilon D$. Thus $|x_1 - x_1| = 0 < \varepsilon$ and δ may be chosen arbitrarily large. A similar argument shows that D has property K with respect to the variable x.

The rectangle of Example 2.2.3 is a special case of Theorem 2.2.6, where $K_1 = [a,b]$ and $K_2 = [c,d]$.

Theorem 2.2.7: Any compact convex set $D \subset E_2$ has property K relative to both x and y.

Proof: We first show that property K relative to y holds at a given point in D.

Since D_y is the projection of the compact convex set D, it is a closed bounded interval, say $D_y = [y_{min}, y_{max}]$.

Now let $\varepsilon > 0$ be given and let (x_1, y_1) be an arbitrary point of D. If $y_1 \neq y_{min}$ then by the convexity of D the point

$$(x_2, y_2) = (\alpha x_1 + (1-\alpha)\bar{x},\ \alpha y_1 + (1-\alpha)y_{min})$$

also belongs to D for each fixed $\alpha \varepsilon [0,1]$, (where $\bar{x}$ is some point in $D_x(y_{min})$).

Now, we can choose

$$\delta_2 = \frac{\varepsilon}{2}\left|\frac{y_1 - y_{min}}{x_1 - \bar{x}}\right| > 0 \quad \text{if } \bar{x} \neq x_1$$

or

$$\delta_2 = |y_1 - y_{min}| > 0 \text{ if } \bar{x} = x_1 .$$

Then since

$$|y_2 - y_1| = (1-\alpha)|y_1 - y_{min}|$$

and

$$|x_2 - x_1| = (1-\alpha)|x_1 - \bar{x}|$$

we have

(1) for $y_2 \varepsilon [y_{min}, y_1]$

$$|y_2 - y_1| < \delta_2 \Rightarrow |x_2 - x_1| < \frac{\varepsilon}{2} .$$

On the other hand, if $y_1 = y_{min}$ then the interval $[y_{min}, y_1]$ reduces to the point y_1 and then (1) holds trivially for any choice of $\delta_2 > 0$.

Now, we can similarly find a $\delta_3 = \delta_3(\varepsilon/2)$ such that

(2) For $y_2 \varepsilon [y_1, y_{max}]$

$$|y_2 - y_1| < \delta_3 = |x_2 - x_1| < \frac{\varepsilon}{2} \quad .$$

We can choose $\delta_1 = \min(\delta_2, \delta_3)$. Then

(3) For $y_2 \varepsilon D_y = [y_{min}, y_{max}]$

$$|y_2 - y_1| < \delta_1 = |x_2 - x_1| < \frac{\varepsilon}{2} \quad .$$

Hence property K holds relative to the variable y at the point (x_1, y_1). We will now extend this to S_1, an open neighborhood of (x_1, y_1).

Let $S_1 = \{(x,y) : |x - x_1)| < \frac{\varepsilon}{2}$ and $|y - y_1| < \frac{\delta_1}{2}\}$ and let $(\bar{x}, \bar{y}) \varepsilon S_1 \cap D$.

Then $\qquad |\bar{x} - x_1| < \frac{\varepsilon}{2}$

and $\qquad |\bar{y} - y_1| < \frac{\delta_1}{2}$

For $y_2 \varepsilon D_y$ and $\bar{y} \varepsilon S_1 \cap D_y$ where $|y_2 - \bar{y}| < \frac{\delta_1}{2}$ we have

$$|y_2 - y_1| \leq |y_2 - \bar{y}| + |\bar{y} - y_1| < \frac{\delta_1}{2} + \frac{\delta_1}{2} = \delta_1 \quad .$$

Thus by property K at (x_1, y_1) there exists a point $(x_2, y_2) \varepsilon D$ where $|x_2 - x_1| < \frac{\varepsilon}{2}$.

Hence $$|x_2 - \bar{x}| \leq |x_2 - x_1| + |x_1 - \bar{x}|$$
$$< \varepsilon/2 + \varepsilon/2 = \varepsilon$$

We have shown that for arbitrary $(\bar{x},\bar{y})\varepsilon S_1 \cap D$, and arbitrary $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon) = \frac{\delta_1}{2}$ such that for $y_2 \varepsilon D_y$ where $|y_2 - \bar{y}| < \delta$ there exists a point $(x_2,y_2)\varepsilon D$ where $|x_2 - \bar{x}| < \varepsilon$.

Thus property K relative to the variable y holds in all of S_1.

This argument can be repeated for any point $(x_u,y_u)\varepsilon D$ and its corresponding open neighborhood S_u. By the compactness of D a finite number of such neighborhoods $S_1,S_2,\ldots,S_n$ is sufficient to cover D.

We choose

$$\delta = \min[\frac{\delta_1}{2},\frac{\delta_2}{2},\ldots,\frac{\delta_k}{2}]$$

(where δ_i corresponds to $(x_i,y_i)\varepsilon D$ and its associated neighborhood S_i for $i = 1,2,\ldots,k$) and Definition 2.2.1 is satisfied.

Hence, D has property K, relative to the variable y.

By a similar argument we can show that D also possesses property K, relative to the variable x.

Note that convexity is not a necessary condition for a set to possess property K as is evidenced by examples 2.2.4 and 2.2.5 and Theorem 2.2.6.

<u>Theorem 2.2.8</u>: Let D be a compact set in E_2, which possesses property K, relative to the variable y. Let y_1 be a fixed point of D_y, and let $x_{11}, x_{12}, \ldots, x_{1n}$ be n distinct points of $D_x(y_1)$, where n is an arbitrary positive integer. Then, there exists a $\delta = \delta(y_1) > 0$, such that $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta$$

implies that there are n disjoint closed intervals $S_1, S_2, \ldots, S_n$ such that

(1) $x_{1i} \varepsilon S_i$ $i = 1,2,\ldots,n$

and (2) There is a point $x_{2i} \varepsilon D_x(y_2)$ where $x_{2i} \varepsilon S_i$, $i = 1,2,\ldots,n$.

<u>Proof</u>: Let $\varepsilon = \min\limits_{\substack{i > j \\ i = 2,\ldots,n \\ j = 1,2,\ldots,n-1}} |x_{1i} - x_{1j}| > 0.$

Note that ε depends on y_1.

Then by property K, there exists a $\delta = \delta(\varepsilon/3) > 0$, (depending on y_1) such that if $x_{1i} \varepsilon D_x(y_1)$ then $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta$$

implies there exists an $x_{2i} \varepsilon D_x(y_2)$ satisfying $|x_{2i} - x_{1i}| < \varepsilon/3$, $i = 1,2,\ldots,n$.

Hence $S_i = \{x \varepsilon D_x : |x - x_{1i}| \leq \varepsilon/3\}$ satisfies the desired conditions (1) and (2).

Corollary 2.2.9: Let D be a compact set in E_2, which possesses property K, relative to the variable y. Let y_1 be a fixed point of D_y such that $D_x(y_1)$ contains $\geq$ n points for some arbitrary positive integer n. Then, there exists a $\delta = \delta(y_1) > 0$ such that $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta$$

implies that $D_x(y_2)$ contains $\geq$ n points.

The following example, illustrates a compact connected set which does not possess property K relative to either x or y. This example will be frequently referred to in later sections.

Example 2.2.10: Let $D = \{(x,y): x = 0 \text{ and } 0 \leq y \leq 1 \text{ or } y = 1 \text{ and } 0 \leq x \leq 1\}$.

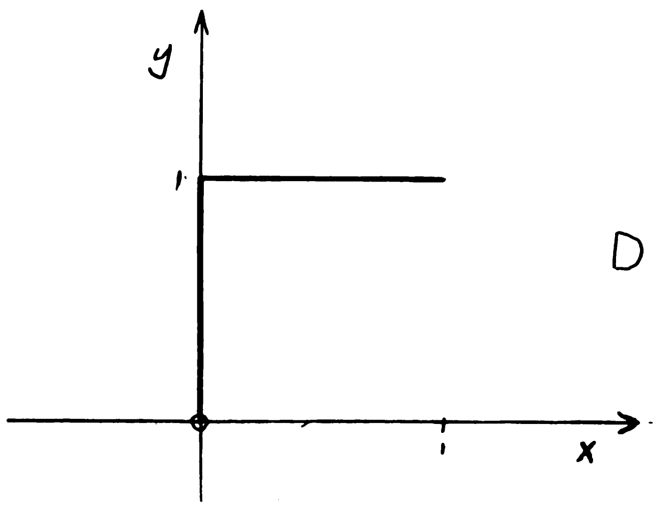


Figure 2.2.3

Then, by a similar argument to that found in the later part of Example 2.2.5 we can show that D does not possess property K relative to x or y.

An alternative proof is to use Corollary 2.2.9.

If we choose $y_1 = 1 \epsilon D_y = [0,1]$ then $D_x(y_1)$ has ≥ 2 points. However, for all $y_2 \epsilon [0,1)$, $D_x(y_2)$ has only 1 point. Therefore, D does not possess property K relative to the variable y. A similar argument shows that D does not possess property K relative to the variable x.

Section 3: Continuity of the Coefficients $a_1(y), a_2(y), \ldots, a_n(y)$

In the remainder of this chapter, for each $y_1 \epsilon D_y$, let F_{y_1} and $||\cdot||_{y_1}$ be as defined in 2.1.1 and 2.1.2 respectively, and for each $A = (a_1, a_2, \ldots, a_n) \epsilon E_n$ let

$$P_A = a_1\phi_1 + a_2\phi_2 + \ldots + a_n\phi_n.$$

<u>Definition 2.3.1</u>: For each $y_1 \epsilon D_y$ let

$$\rho(y_1) = \inf_{A \epsilon E_n} ||F_{y_1} - P_A||_{y_1} .$$

We know that if $F \epsilon C(D)$, then $F_{y_1} \epsilon C(D_x(y_1))$. Also if $\phi_1, \phi_2, \ldots, \phi_n$ is a system of n real-valued, continuous functions on D_x, then they are likewise real-valued, continuous functions on $D_x(y_1)$, and therefore we know that there is some polynomial P_{A_1} which is a best $||\cdot||_{y_1}$ approximation to F_{y_1}. Consequently, we have the existence of a polynomial P_{A_1}, such that

$$\rho(y_1) = ||F_{y_1} - P_{A_1}||_{y_1} \quad .$$

Theorem 2.3.2: Let D be a compact set in E_2, which possesses property K, relative to the variable y, and let $\phi_1, \phi_2, \ldots, \phi_n$ be a Tchebycheff system of real-valued, continuous functions on D_x. Let $F \varepsilon C(D)$. Then $\rho(y)$ is a continuous function for $y \varepsilon D_y$.

Proof: Let y_1 be a fixed point in D_y and let $\varepsilon > 0$ be given. Then for all $y_2 \varepsilon D_y$

$$||F_{y_2} - P_{A_2}||_{y_2} \leq ||F_{y_2} - P_{A_1}||_{y_2} \quad ,$$

where P_{A_i} is a best $||\cdot||_{y_i}$ approximation to F_{y_i}, $i = 1,2$. (Note that because $D_x(y_i)$ may contain less than n points, P_{A_i} might not be a unique best approximation.)

Now, let $\delta_1 = \delta_1(\varepsilon, y_1)$ correspond to ε in the definition of the uniform continuity of the function G defined on D, by

$$G(x,y) \equiv F_y(x) - P_{A_1}(x) \equiv F(x,y) - P_{A_1}(x).$$

Also, by property K relative to the variable y for the set D, given $\delta_1(\varepsilon, y_1) > 0$, there exists a $\delta_2 = \delta_2(\delta_1(\varepsilon, y_1)) > 0$ such that for $y_2 \varepsilon D_y$

$$|y_2 - y_1| < \delta_2$$

implies that for each $x_2 \varepsilon D_x(y_2)$, there is a corresponding $x_1 \varepsilon D_x(y_1)$ satisfying,

$$\sigma_2[(x_1,y_1);(x_2,y_2)] < \delta_1(\varepsilon) .$$

This in turn implies that

$$|F_{y_2}(x_2) - P_{A_1}(x_2)| < |F_{y_1}(x_1) - P_{A_1}(x_1)| + \varepsilon .$$

Combining these results we have,

(1) given $\varepsilon > 0$, and y_1 fixed in D_y there exists a $\delta_2 = \delta_2(\delta_1(\varepsilon, y_1)) > 0$ such that $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta_2$$

imply that for some $x_2 \varepsilon D_x(y_2)$ and a corresponding $x_1 \varepsilon D_x(y_1)$

$$\begin{aligned}
\rho(y_2) - \rho(y_1) &= ||F_{y_2} - P_{A_2}||_{y_2} - ||F_{y_1} - P_{A_1}||_{y_1} \\
&\leq ||F_{y_2} - P_{A_1}||_{y_2} - ||F_{y_1} - P_{A_1}||_{y_1} \\
&= |F_{y_2}(x_2) - P_{A_1}(x_2)| - ||F_{y_1} - P_{A_1}||_{y_1} \\
&< |F_{y_1}(x_1) - P_{A_1}(x_1)| + \varepsilon - ||F_{y_1} - P_{A_1}||_{y_1} \\
&\leq ||F_{y_1} - P_{A_1}||_{y_1} + \varepsilon - ||F_{y_1} - P_{A_1}||_{y_1} \\
&= \varepsilon .
\end{aligned}$$

Now, we consider two cases:

Case 1: $D_x(y_1)$ contains $\leq$ n points.

Then there exists a polynomial P_{A_1} such that

$$||F_{y_1} - P_{A_1}||_{y_1} = 0 .$$

Therefore, $\rho(y_1) = 0$.

Now by (1)

$$|y_2 - y_1| < \delta_2$$
$$\Rightarrow \rho(y_2) - \rho(y_1) = \rho(y_2) < \varepsilon .$$

Hence,
$$|y_2 - y_1| < \delta_2$$
$$\Rightarrow \quad |\rho(y_2) - \rho(y_1)| = |\rho(y_2)| = \rho(y_2) < \varepsilon ,$$

since $\rho(y)$ is a non-negative function.

<u>Case 2</u>: $D_x(y_1)$ contains $\geq$ n+1 points

Then by Theorem 2.2.8 and Corollary 2.2.9, there exists a $\delta_3 = \delta_3(y_1) > 0$, such that $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta$$

implies that (1) $D_x(y_2)$ contains $\geq$ n+1 points and these points $x_{2,1}, x_{2,2}, \ldots, x_{2,n+1}$ each belongs to a corresponding closed internal $S_1, S_2, \ldots, S_{n+1}$ where $S_i \cap S_j = \phi$, $i \neq j$.

Now, by Haar's Theorem, F_{y_1} and F_{y_2} each has a unique best uniform approximation on their respective domains. Denote these approximations by P_{A_1} and P_{A_2} respectively.

As before we have

$$||F_{y_1} - P_{A_1}||_{y_1} \leq ||F_{y_1} - P_{A_2}||_{y_1} \quad .$$

Let $$M = \max_{\substack{y_2 \varepsilon D_y \\ |y_2 - y_1| < \delta_3}} ||F_{y_2}||_{y_2}$$

Then, since the zero polynomial is a possible approximation to F_{y_2}, we have

$$||F_{y_2} - P_{A_2}||_{y_2} \leq ||F_{y_2}||_{y_2} \quad .$$

Therefore,

$$\begin{aligned} ||P_{A_2}||_{y_2} &\leq ||F_{y_2} - P_{A_2}||_{y_2} + ||F_{y_2}||_{y_2} \\ &\leq 2||F_{y_2}||_{y_2} \\ &\leq 2M \quad . \end{aligned}$$

Now, consider the set $\mathcal{A} = \{A_2 : P_{A_2}$ is the best $||\cdot||_{y_2}$ approximation to F_{y_2} for some $y_2 \varepsilon D_y$ where $|y_2 - y_1| < \delta_3\}$.

We first assert that $\mathcal{A}$ is a bounded set. We begin by finding expressions for the coefficients $a_{21}, a_{22}, \ldots, a_{2n}$, where $A_2 = (a_{21}, a_{22}, \ldots, a_{2n}) \varepsilon \mathcal{A}$.

The function G defined on the compact set $S = S_1 \times S_2 \times \ldots \times S_n$, by

$$G(x_1, x_2, \ldots, x_n) = \begin{vmatrix} \phi_1(x_1) & \phi_2(x_1) & \ldots & \phi_n(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \ldots & \phi_n(x_2) \\ \vdots & \vdots & & \vdots \\ \phi_1(x_n) & \phi_2(x_n) & \ldots & \phi_n(x_n) \end{vmatrix} \quad .$$

is continuous there and therefore assumes its minimum absolute value m at some point $(\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n) \varepsilon S$.

Since $\phi_1, \ldots, \phi_n$ is a Tchebycheff system on D_x, if $m = 0$ then for some $i \neq j$, $\overline{x}_i = \overline{x}_j$. However, $\overline{x}_i \varepsilon S_i$ and

$\bar{x}_j \varepsilon S_j$ and $S_i \cap S_j = \Phi$. Therefore, $m > 0$.

Now, for $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta_3$$

and the corresponding point $A_2 \varepsilon \mathcal{A}$, we consider the determinants

$$D_1 = \begin{vmatrix} P_{A_2}(x_{21}) & \phi_2(x_{21}) & \cdots & \phi_n(x_{21}) \\ \vdots & \vdots & & \vdots \\ P_{A_2}(x_{2n}) & \phi_2(x_{2n}) & \cdots & \phi_n(x_{2n}) \end{vmatrix}$$

$$D_2 = \begin{vmatrix} \phi_1(x_{21}) & P_{A_2}(x_{21}) & \phi_3(x_{21}) & \cdots & \phi_n(x_{21}) \\ \vdots & \vdots & \vdots & & \vdots \\ \phi_1(x_{2n}) & P_{A_2}(x_{2n}) & \phi_3(x_{2n}) & \cdots & \phi_n(x_{2n}) \end{vmatrix}$$

$$\vdots$$

$$\text{and } D_n = \begin{vmatrix} \phi_1(x_{21}) & \cdots & \phi_{n-1}(x_{21}) & P_{A_2}(x_{21}) \\ \vdots & & \vdots & \vdots \\ \phi_1(x_{2n}) & \cdots & \phi_{n-1}(x_{2n}) & P_{A_2}(x_{2n}) \end{vmatrix},$$

where $x_{21}, x_{22}, \ldots, x_{2i}$ are points of $D_x(y_2)$, such that

$$x_{2i} \varepsilon S_i,\ i = 1,2,\ldots,n.$$

Since, $$\max_{x = x_{21}, x_{22}, \ldots, x_{2n}} |P_{A_2}(x)| \leq ||P_{A_2}||y_2 \leq 2M$$

we note that the determinants $D_1, D_2, \ldots,$ and D_n are bounded in absolute value. Let their bounds be $B_1, B_2, \ldots, B_n$ respectively.

$$\begin{pmatrix} \phi_1(x_{21}) & \phi_2(x_{21}) & \cdots & \phi_n(x_{21}) \\ \phi_1(x_{22}) & \phi_2(x_{22}) & \cdots & \phi_n(x_{22}) \\ \vdots & \vdots & & \vdots \\ \phi_1(x_{2n}) & \phi_2(x_{2n}) & \cdots & \phi_n(x_{2n}) \end{pmatrix} \begin{pmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2n} \end{pmatrix} = \begin{pmatrix} P_{A_2}(x_{21}) \\ P_{A_2}(x_{22}) \\ \vdots \\ P_{A_2}(x_{2n}) \end{pmatrix} .$$

We have by Cramer's rule

$$|a_{21}| = \frac{|D_1|}{|G(x_{21}, x_{22}, \ldots, x_{2n})|} \leq \frac{B_1}{m}$$

$$|a_{22}| = \frac{|D_2|}{|G(x_{21}, x_{22}, \ldots, x_{2n})|} \leq \frac{B_2}{m}$$

$$|a_{2n}| = \frac{|D_n|}{|G(x_{21}, x_{22}, \ldots, x_{2n})|} \leq \frac{B_n}{m} .$$

Hence, $\mathcal{A}$ is a bounded set and its closure $\bar{\mathcal{A}}$ is compact.

The function H defined on $D \times \bar{\mathcal{A}}$ by

$$H(x,y,A) = F_y(x) - P_A(x)$$

is uniformly continuous there.

Thus, given $\varepsilon > 0$, there exists a $\delta_4 = \delta_4(\varepsilon) > 0$ such that if (x_1,y_1,A_2) and $(x_2,y_2,A_2)\varepsilon D_x\mathcal{A}$ then

$$\sigma_{n+2}[(x_1,y_1,A_2);(x_2,y_2,A_2)] < \delta_4$$

implies that,

$$|F_{y_1}(x_1) - P_{A_2}(x_1)| < |F_{y_2}(x_2) - P_{A_2}(x_2)| + \varepsilon.$$

Also by property K, relative to the variable y, for the set D, given $\delta_4(\varepsilon) > 0$, there exists a $\delta_5 = \delta_5(\delta_4(\varepsilon)) > 0$ such that for $y_2\varepsilon D_y$

$$|y_2 - y_1| < \delta_5$$

implies that for each $x_1\varepsilon D_x(y_1)$ there is a corresponding $x_2\varepsilon D_x(y_2)$ satisfying,

$$\sigma_2[x_1,y_1);(x_2,y_2)] < \delta_4(\varepsilon)$$

or equivalently,

$$\sigma_{n+2}[(x_1,y_1,A_2);(x_2,y_2,A_2)] < \delta_4(\varepsilon) \quad ,$$

which in turn implies that

$$|F_{y_1}(x_1) - P_{A_2}(x_1)| < |F_{y_2}(x_2) - P_{A_2}(x_2)| + \varepsilon.$$

We can choose $\delta_6 = \min(\delta_3,\delta_5) > 0$. Then combining these results we have,

(2) given $\varepsilon > 0$ and y_1 fixed in D_y such that $D_x(y_1)$ contains $\geq$ n+1 points, there exists a $\delta_6 = \delta_6(\varepsilon,y_1) > 0$ such that for $y_2\varepsilon D_y$.

$$|y_2 - y_1| < \delta_6$$

implies that, for some $x_1 = x_1(y_2) \varepsilon D_x(y_1)$ and a corresponding $x_2 = x_2(y_2) \varepsilon D_x(y_2)$,

$$\begin{aligned}
\rho(y_1) - \rho(y_2) &= ||F_{y_1} - P_{A_1}||_{y_1} - ||F_{y_2} - P_{A_2}||_{y_1} \\
&\leq ||F_{y_1} - P_{A_2}||_{y_1} - ||F_{y_2} - P_{A_2}||_{y_2} \\
&= |F_{y_1}(x_1) - P_{A_2}(x_1)| - ||F_{y_2} - P_{A_2}||_{y_2} \\
&< |F_{y_2}(x_2) - P_{A_2}(x_2)| + \varepsilon - ||F_{y_2} - P_{A_2}||_{y_2} \\
&\leq ||F_{y_2} - P_{A_2}||_{y_2} + \varepsilon - ||F_{y_2} - P_{A_2}||_{y_2} \\
&= \varepsilon \quad .
\end{aligned}$$

Now we choose $\delta = \min(\delta_2, \delta_6)$ and we have for case 2, by (1) and (2)

$$|y_2 - y_1| < \delta$$

implies that

$$|\rho(y_2) - \rho(y_1)| < \varepsilon \quad .$$

We will use the continuity of $\rho(y)$ to prove the continuity of $A(y)$ under suitable conditions. With this goal in mind we state the following theorem due to Remez (16).

<u>Theorem 2.3.3 (Remez)</u>: Let D be a compact set in E_1, which contains $\geq$ n+1 points. Let $F \epsilon C(D)$ and let $\phi_1(x) \equiv 1$, $\phi_2(x) \equiv x, \ldots, \phi_n(x) \equiv x^{n-1}$. Let $||\cdot||$ be the uniform norm on $C(D)$, and let $P_A{}^*$ be the best $||\cdot||$ approximation to F on D. Then, given $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon) > 0$ such that

$$||F - P_A|| \leq ||F - P_A*|| + \delta$$

implies that

$$\sigma(A;A*) < \varepsilon .$$

Essentially what this says is that if the deviation of a polynomial P_A is close to the deviation of the best approximation P_A* then the corresponding parameters A and A* must be close.

The proof of the following more general theorem closely parallels the Remez proof. This theorem is one of a sequence of results from which the continuity of A(y) will follow.

Theorem 2.3.4: Let D be a compact set in E_1, which contains n or more points. Let $F \varepsilon C(D)$ and let $\phi_1, \phi_2, \ldots, \phi_n$ be a linearly independent system of continuous real-valued functions on D. Let $||\cdot||$ be the uniform norm on C(D) and let

$$\rho = \inf_{A \varepsilon E_n} ||F - P_A|| \quad .$$

Then given $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon) > 0$, such that

$$||F - P_A|| < \rho + \delta$$

implies that there is a best $||\cdot||$ approximation P_A*, to F on D, such that

$$\sigma(A;A^*) < \varepsilon \ .$$

Essentially what this theorem says is that if the deviation of a polynomial P_A is close to the deviation of a best $||\cdot||$ approximation to F, then the parameter A is close to the parameter of some best $||\cdot||$ approximation to F.

<u>Proof</u>: Let $M = ||F||$, and let

$$T = \{A \varepsilon E_n : ||P_A|| \leq 2M\} \ .$$

Since $\phi_1, \phi_2, \ldots, \phi_n$ is a linearly independent system on D, there exists a set of distinct points $x_1, x_2, \ldots, x_n$ in D, such that

$$\begin{vmatrix} \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_n(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \cdots & \phi_n(x_2) \\ \vdots & \vdots & & \vdots \\ \phi_1(x_n) & \phi_2(x_n) & \cdots & \phi_n(x_n) \end{vmatrix} = m \neq 0 \ .$$

For arbitrary $A \varepsilon T$ the determinants

$$D_1 = \begin{vmatrix} P_A(x_1) & \phi_2(x_1) & \cdots & \phi_n(x_1) \\ \vdots & \vdots & & \vdots \\ P_A(x_n) & \phi_2(x_n) & \cdots & \phi_n(x_n) \end{vmatrix}$$

$$D_2 = \begin{vmatrix} \phi_1(x) & P_A(x_1) & \phi_3(x_1) & \cdots & \phi_n(x_1) \\ \vdots & \vdots & \vdots & & \vdots \\ \phi_1(x_n) & P_A(x_n) & \phi_3(x_n) & \cdots & \phi_n(x_n) \end{vmatrix}$$

$\vdots$

and

$$D_n = \begin{vmatrix} \phi_1(x_1) & \cdots & \phi_{n-1}(x_1) & P_A(x_1) \\ \vdots & & \vdots & \vdots \\ \phi_1(x_n) & \cdots & \phi_{n-1}(x_n) & P_A(x_n) \end{vmatrix}$$

are bounded respectively by $B_1, B_2, \ldots, B_n$. Therefore, by Cramer's rule

$$|a_1| \leq \left|\frac{B_1}{m}\right|, \ |a_2| \leq \left|\frac{B_2}{m}\right|, \ldots, \ |a_n| \leq \left|\frac{B_n}{m}\right| ,$$

for each $A = (a_1, a_2, \ldots, a_n) \varepsilon T$. Hence, T is bounded.

Now, if A' is a fixed point of T', the complement of T, then

$$||P_{A'}|| > 2M \quad .$$

By the continuity of $||P_A||$ as a function of $A \varepsilon E_n$, given $\varepsilon = ||P_{A'}|| - 2M$ we can find a $\delta = \delta(\varepsilon, A')$ such that $A \varepsilon E_n$ and

$$\sigma(A;A') < \delta$$

imply that

$$\begin{aligned} ||P_A|| &> ||P_{A'}|| - \varepsilon \\ &= ||P_{A'}|| - (||P_{A'}|| - 2M) \\ &= 2M. \end{aligned}$$

This implies that

$$\{A\varepsilon E_n : \sigma(A;A') < \delta\} \subset T'.$$

Therefore, T' is open and T is compact.

For some $\hat{M} > M$, we define the set $\hat{T} = \{A\varepsilon E_n : ||P_A|| \leq 2\hat{M}\}$. We can similarly show that $\hat{T}$ is compact.

Let $\mathcal{A}^* = \{A^*\varepsilon E_n : P_A{}^*$ is a best $||\cdot||$ approximation to F on D$\}$.

For each $A^*\varepsilon\,\mathcal{A}^*$

$$||P_A{}^*|| \leq ||F|| + ||F - P_A{}^*|| \leq 2||F|| = 2M.$$

Therefore,

$$\hat{T} \supset T \supset \mathcal{A}^* \neq \phi.$$

The linear independence of $\phi_1, \phi_2, \ldots, \phi_n$ on D implies that $||\phi_1|| \neq 0$. If $P = \frac{2\hat{M}}{||\phi_1||}\,\phi_1$, then $||P|| = 2\hat{M}$. Therefore, $\hat{T} - T$ is a non-empty subset of $\hat{T} - \mathcal{A}^*$.

Now, given $\varepsilon > 0$, we define the set $T_\varepsilon = \{A\varepsilon\hat{T} : \sigma(A;A^*) \geq \varepsilon$, for all $A^*\varepsilon\,\mathcal{A}^*\}$, and note that for ε sufficiently small T_ε is a non-empty compact subset of $\hat{T} - a^*$.

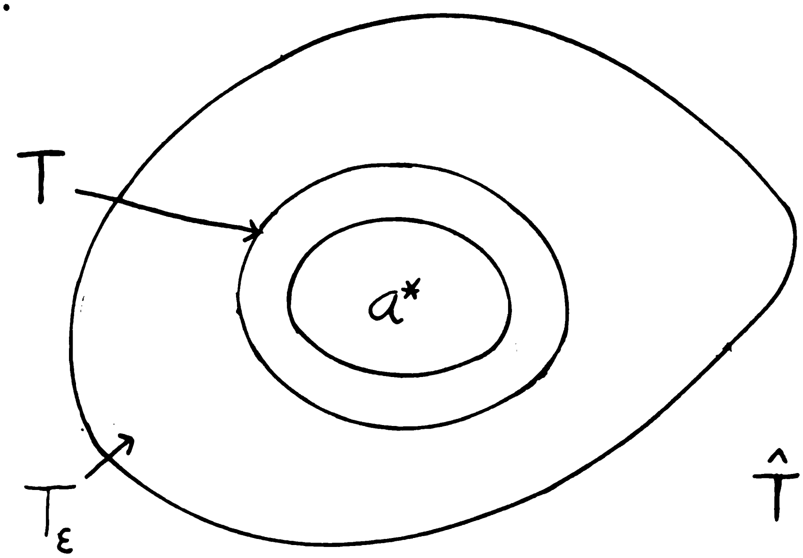


Figure 2.3.1

Since $||F - P_A||$ is a continuous function of $A \varepsilon E_n$, it assumes its minimum $\rho_\varepsilon > \rho$ on the compact set T_ε.

Choose $\delta = \min(\rho_\varepsilon - \rho, 2\hat{M} - M - \rho)$. Then $A \varepsilon E_n$ and

$$||F - P_A|| < \rho + \delta$$

imply (1) $||F - P_A|| < \rho_\varepsilon$

and (2) $||F - P_A|| < 2\hat{M} - M$.

Therefore $A \notin T_\varepsilon$ by the definition of ρ_ε.

Also $2\hat{M} - M > ||F - P_A|| \geq ||P_A|| - M$ implies that $2\hat{M} > ||P_A||$.

This implies that $A \varepsilon \hat{T}$.

Hence $A\varepsilon\hat{T} - T_\varepsilon$, and therefore

$$\sigma(A;A^*) < \varepsilon, \text{ for some } A^*\varepsilon\mathcal{A}^* \quad .$$

If y_1 is a fixed point of D_y such that $D_x(y_1)$ contains less than n points, then F_{y_1} does not possess a unique best $||\cdot||_{y_1}$ approximation in $\phi_1,\phi_2,\ldots,\phi_n$. Therefore, $A(y)$ is not well defined at y_1. This problem is considered at the end of this section. We presently restrict our attention to the case where $D_x(y_1)$ contains n or more points.

In the remainder of this chapter the phrase "let D, F, and $\phi_1,\phi_2,\ldots,\phi_n$ be as usual" will be used to mean that D is a compact set in E_2, which possesses property K relative to the variable y, $F\varepsilon C(D)$ and $\phi_1,\phi_2,\ldots,\phi_n$ is a Tchebycheff system on continuous real-valued functions on D_x.

<u>Lemma 2.3.5</u>: Let D, F and $\phi_1,\phi_2,\ldots,\phi_n$ be as usual, and let y_1 be a fixed point of D_y, such that $D_x(y_1)$ contains n or more points. Then given $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon,y_1) > 0$, such that $y_2\varepsilon D_y$ and

$$|y_2 - y_1| < \delta$$

imply that

$$||F_{y_1} - P_{A(y_2)}||_{y_1} < \rho(y_1) + \varepsilon.$$

<u>Proof</u>: By Corollary 2.2.9 there exists a $\delta_1 = \delta_1(y_1) > 0$, such that $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta_1$$

imply that $D_x(y_2)$ contains n or more points, and therefore $A(y)$ is well-defined at y_2.

As in Theorem 2.3.2 we can show that the set $\mathcal{A} = \{A(y_2) \varepsilon E_n : y_2 \varepsilon D_y \text{ and } |y_2 - y_1| < \delta_1\}$ is bounded. Then as before let $\delta_2 = \delta_2(\varepsilon/2) > 0$ correspond to the definition of the uniform continuity of the function G, defined on $D \times \mathcal{A}$ by

$$G(x,y,A) = F_y(x) - P_A(x).$$

Also, by property K, given $\delta_2 = \delta_2(\varepsilon/2) > 0$ there exists a $\delta_3 = \delta_3(\delta_2) > 0$ such that $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta_3$$

imply that for each $x_1 \varepsilon D_x(y_1)$ there is a corresponding $x_2 \varepsilon D_x(y_2)$ satisfying

$$\sigma_2[(x_2,y_2);(x_1,y_1)] < \delta_2,$$

or equivalently

$$\sigma_{n+2}[(x_2,y_2,A(y_2));(x_1,y_1,A(y_2))] < \delta_2.$$

Choose $\delta_4 = \min(\delta_1,\delta_3)$. Then we have shown that for each $y_2 \varepsilon D_y$

$$|y_2 - y_1| < \delta_4$$

implies that there exists an $x_1 \varepsilon D_x(y_1)$ and a corresponding $x_2 \varepsilon D_x(y_2)$ such that

$$||F_{y_1} - P_{A(y_2)}||_{y_1} = |F_{y_1}(x_1) - P_{A(y_2)}(x_1)|$$

$$< |F_{y_2}(x_2) - P_{A(y_2)}(x_2)| + \varepsilon/2$$

$$\leq \rho(y_2) + \varepsilon/2.$$

Let $\delta_5 = \delta_5(\varepsilon/2) > 0$ correspond to the definition of the uniform continuity of $\rho(y)$ on D_y.

Choose $\delta = \min(\delta_4, \delta_5) > 0$. Then $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta$$

imply that

$$||F_{y_1} - P_{A(y_2)}||_{y_1} < \rho(y_2) + \varepsilon/2$$

$$< \rho(y_1) + \varepsilon.$$

We can now prove that under a suitable hypotheses $A(y)$ is continuous at $y_1 \varepsilon D_y$.

<u>Theorem 2.3.6</u>: Let D, F and $\phi_1, \phi_2, \ldots, \phi_n$ be as usual and let y_1 be a fixed point of D_y, such that $D_x(y_1)$ contains n or more points. Then the function $A(y)$ as defined in 2.1.3 is continuous at y_1.

Proof: $D_x(y_1)$ is a compact set in E_1, and $F_{y_1} \varepsilon C(D_x(y_1))$. $\phi_1, \phi_2, \ldots, \phi_n$ a Tchebycheff system on D_x implies that it is a Tchebycheff system on $D_x(y_1)$ which in turn implies that it is a linearly independent system on $D_x(y_1)$. Therefore, by Theorem 2.3.4 given $\varepsilon > 0$, there exists a $\delta_1 = \delta_1(\varepsilon) > 0$ such that

$$||F_{y_1} - P_A||_{y_1} < \rho(y_1) + \delta_1$$

implies that

$$\sigma[A;A(y_1)] < \varepsilon \ .$$

Now by Lemma 2.3.5 there exists a $\delta = \delta(\delta_1(\varepsilon), y_1) > 0$ such that $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta$$

imply that

$$||F_{y_1} - P_{A(y_2)}||_{y_1} < \rho(y_1) + \delta_1,$$

which in turn implies that

$$\sigma[A(y_2);A(y_1)] < \varepsilon$$

Corollary 2.3.7: Under the hypothesis of Theorem 2.3.6, $a_1(y), a_2(y), \ldots, a_n(y)$ as defined in 2.1.3, are all continuous at y_1.

We have previously mentioned that A(y) is not well-defined for $y_1 \varepsilon D_y$, such that $D_x(y_1)$ contains less than n points. Example 2.3.8 below illustrates such a situation.

Example 2.3.8: Let $D = \{(x,y): 0 \leq x \leq 1$ and $y \geq 0$ and $y \leq 1 - x\}$.

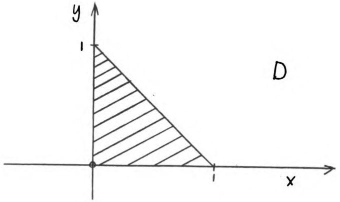


Figure 2.3.2

Let $F(x,y) = \sqrt{x}$ and let $\phi_1(x) = 1$, $\phi_2(x) = x$.
Then for $0 \leq y_1 < 1$, $D_x(y_1) = [0, 1 - y_1]$ and

$$P_{A(y_1)} = \frac{\sqrt{1 - y_1}}{8} \phi_1 + \frac{1}{\sqrt{1 - y_1}} \phi_2 \quad .$$

For $y_1 = 1$, $D_x(y_1) = \{0\}$ and

$$P_{(0,\alpha)} = \alpha\phi_2$$

is a best $||\cdot||_{y_1}$ approximation to F_{y_1}, for each real α.

Hence $$A(y) = (a_1(y), a_2(y)) = \left(\frac{\sqrt{1-y}}{8}, \frac{1}{\sqrt{1-y}}\right)$$

$$\text{for } 0 \leq y < 1,$$

and is not well-defined at $y = 1$. Moreover, there is no value we could assign to $A(1)$ to make $A(y)$ continuous at $y = 1$.

Section 4: The Product Tchebycheff Approximation

Corollary 2.3.7 motivates our seeking a polynomial approximation to each of the coefficient functions, $a_1(y), a_2(y), \ldots, a_n(y)$.

Definition 2.4.1: We will say that D, F, $\phi_1, \phi_2, \ldots, \phi_n$, and $\psi_1, \psi_2, \ldots, \psi_m$ satisfy condition P.T., if (1) D, F and $\phi_1, \phi_2, \ldots, \phi_n$ are as usual, (2) For each $y_1 \varepsilon D_y$, the set $D_x(y_1)$ contains $\geq$ n points and (3) $\psi_1, \psi_2, \ldots, \psi_m$ is a Tchebycheff system of continuous real-valued functions on D_y which contains m or more points.

These are the necessary conditions to define the Product Tchebycheff approximation.

Definition 2.4.2: Let D, F, $\phi_1, \ldots, \phi_n$, and $\psi_1, \ldots, \psi_m$ satisfy condition P. T. Then by Haar's theorem for each $k = 1, 2, \ldots, n$, the continuous function $a_k(y)$ has a unique best uniform approximation on D_y. We will denote this corresponding best approximation by

$$Q_{A_k} = \sum_{i=1}^{m} a_{ik}\psi_i; \; A_k = (a_{1k}, a_{2k}, \ldots, a_{mk}) \varepsilon E_m.$$

The polynomial

$$PT_A = \sum_{j=1}^{n} Q_{A_j} \phi_j$$

is called the Product Tchebycheff (P. T.) approximation to F on D, relative to the variable y.

We can similarly define the P. T. approximation to F on D, relative to the variable x. The following example shows that these two approximations (in cases where both are defined) need not be the same.

<u>Example 2.4.3</u>: Let $D = [-1,1]\times[-1,1]$.

Let
$$F(x,y) = \begin{cases} -2yx^2 + y & \text{for } -1 \le y \le 0 \\ 2yx & \text{for } \;\; 0 < y \le 1 \end{cases}$$

and let
$$\phi_1(x) = 1$$
$$\psi_1(y) = 1.$$

Then $F \varepsilon C(D)$.

For $-1 \le y_1 \le 0$

$$\max_{x\varepsilon[-1,1]} F_{y_1}(x) = -y_1 \quad , \quad \min_{x\varepsilon[-1,1]} F_{y_1}(x) = y_1$$

$$\Rightarrow \qquad P_{A(y_1)}(x) = 0.$$

Therefore, $\max\limits_{y\varepsilon[-1,1]} P_{A(y_1)}(x) = \min\limits_{y\varepsilon[-1,1]} P_{A(y_1)}(x) = 0,$

$$\Rightarrow \qquad PT_A(x,y) = 0.$$

Now, we define F_{x_1} and $P_{B(x_1)}$ analagously to F_{y_1} and $P_{A(y_1)}$, and we let PT_B denote the product Tchebycheff approximation, relative to the variable x.

For $-1 \leq x_1 \leq 0$

$$\max_{y \in [-1,1]} F_{x_1}(y) = 0 \quad , \quad \min_{y \in [-1,1]} F_{x_1}(y) = 2x_1 - 1$$

$$\Rightarrow \quad P_{B(x_1)}(y) = x_1 - \tfrac{1}{2}.$$

For $0 < x_1 \leq \frac{1}{2}$

$$\max_{y \in [-1,1]} F_{x_1}(y) = 2x_1 \quad , \quad \min_{y \in [-1,1]} F_{x_1}(y) = 2x_1 - 1$$

$$\Rightarrow \quad P_{B(x_1)}(y) = 2x_1 - \tfrac{1}{2}.$$

For $\frac{1}{2} < x_1 \leq 1$

$$\max_{y \in [-1,1]} F_{x_1}(y) = 2x_1 \quad , \quad \min_{y \in [-1,1]} F_{x_1}(y) = 0$$

$$\Rightarrow \quad P_{B(x_1)}(y) = x_1.$$

Therefore

$$\max_{x_1 \in [-1,1]} P_{B(x_1)}(y) = 1 \quad , \quad \min_{x_1 \in [-1,1]} P_{B(x_1)}(y) = -\tfrac{3}{2}$$

$$\Rightarrow \quad PT_B(x,y) = -\tfrac{1}{4} \; .$$

Throughout the remainder of this chapter the P. T. or product Tchebycheff approximation will refer to the product Tchebycheff approximation, relative to the variable y.

In Chapter II, Section 9, the P. T. approximation is extended to functions of three or more variables.

In Chapter II, Section 8, we define a similar approximation on sets which do not possess property K.

We now consider a special case of product Tchebycheff approximation:

<u>Theorem 2.4.4</u>: Let the compact set $D \subset E_2$ have property K, relative to the variable y and let $F \varepsilon C(D)$. Let $\phi_1(x) = 1$, $\psi_1(y) = 1$.

Then the product Tchebycheff approximation to F on D is defined by

$$PT_A(x,y) = \frac{1}{4}\left[\max_{y_1 \varepsilon D_y}\left\{\max_{x_1 \varepsilon D_x(y_1)} F(x_1,y_1) + \min_{x_1 \varepsilon D_x(y_1)} F(x_1,y_1)\right\}\right.$$

$$\left. + \min_{y_1 \varepsilon D_y}\left\{\max_{x_1 \varepsilon D_x(y_1)} F(x_1,y_1) + \min_{x_1 \varepsilon D_x(y_1)} F(x_1,y_1)\right\}\right].$$

<u>Proof</u>: For each $y_1 \varepsilon D_y$

$$(1) \qquad P_{A(y_1)}(x) = a_1(y_1)\phi_1(x) = a_1(y_1)$$

$$= \frac{1}{2}\left\{\max_{x_1 \varepsilon D_x(y_1)} F_{y_1}(x_1) + \min_{x_1 \varepsilon D_x(y_1)} F_{y_1}(x_1)\right\}$$

$$= \frac{1}{2}\left\{\max_{x_1 \varepsilon D_x(y_1)} F(x_1,y_1) + \min_{x_1 \varepsilon D_x(y_1)} F(x_1,y_1)\right\}.$$

Also

$$(2) \quad PT_A(x,y) = \frac{1}{2}\left[\max_{y_1 \varepsilon D_y} a_1(y_1) + \min_{y_1 \varepsilon D_y} \{a_1(y_1)\}\right] \;.$$

We combine equations (1) and (2) to obtain the desired result.

If we just restrict D to be an arbitrary compact set in E_2 and F to be an arbitrary bounded function on D, then with $\phi_1(x) \equiv \psi_1(y) \equiv 1$

$$PT_A(x,y) = \frac{1}{4}\left[\sup_{y_1 \varepsilon D_y}\left(\sup_{x_1 \varepsilon D_x(y_1)} F(x_1,y_1) + \inf_{x_1 \varepsilon D_x(y_1)} F(x_1,y_1)\right)\right.$$

$$\left. + \inf_{y_1 \varepsilon D_y}\left(\sup_{x_1 \varepsilon D_x(y_1)} F(x_1,y_1) + \inf_{x_1 \varepsilon D_x(y_1)} F(x_1,y_1)\right)\right] .$$

The following example illustrates some of the ideas presented up to this point:

<u>Example 2.4.5</u>: Let $D = [0,1]\times[0,1]$ and let $F(x,y) = x + y$, $\phi_1(x) = 1$, $\psi_1(y) = 1$.

For $y_1 \varepsilon D_y = [0,1]$

$$\max_{x\varepsilon[0,1]} F_{y_1}(x) = 1 + y_1, \quad \min_{x\varepsilon[0,1]} F_{y_1}(x) = y_1$$

$$\Rightarrow \quad P_{A(y_1)}(x) = a_1(y_1)\phi_1(x) = a_1(y_1) \;.$$

$$= \frac{1}{2} + y_1 \;.$$

Therefore,

$$\max_{y\varepsilon[0,1]} a_1(y_0) = \frac{3}{2} \qquad , \qquad \min_{y\varepsilon[0,1]} a_1(y) = \frac{1}{2}$$

$$\Rightarrow PT_A(x,y) = 1.$$

Note that in this example the Product Tchebycheff approximation to F is also the unique best uniform approximation to F by a constant. We will see that in general the P. T. approximation is not a best uniform approximation. However, in Chapter II, Section 7, we investigate the P. T. approximation using suitable sequences of base functions $\{\phi_j\}$, $\{\psi_i\}$. We shall show that by choosing n and m properly the P. T. approximation can be made arbitrarily close to a best uniform approximation.

Now

$$\rho(y_1) = ||F_{y_1} - P_{A(y_1)}||_{y_1}$$

$$= \max_{x\varepsilon[0,1]} |(x+y) - (\tfrac{1}{2} + y)|$$

$$= \frac{1}{2} .$$

Hence, both $\rho(y)$ and $A(y) = (a_1(y))$ are continuous on D_y, as was asserted in Theorems 2.3.2 and 2.3.6.

We now consider an example in which the set D does not possess property K, relative to the variable y and in which both $\rho(y)$ and $A(y)$ are discontinuous on D_y.

Example 2.4.6: Let

$D = \{(x,y): x = 0 \text{ and } 0 \leq y \leq 1, \text{ or } 0 \leq x \leq 1 \text{ and } y = 1\}$.

It was shown in Example 2.2.10 that D does not possess property K relative to either x or y.

Now, let $F(x,y) = x + y$, $\phi_1(x) = 1$, $\psi_1(y) = 1$. Then for $0 \leq y_1 < 1$

$$F_{y_1}(x) = 0 + y_1 = y_1 \quad \text{on } D_x(y_1) = \{0\} ,$$

and therefore,

$$\begin{aligned} P_{A(y_1)}(x) &= a_1(y_1)\phi_1(x) \\ &= a_1(y_1) \\ &= y_1 \quad . \end{aligned}$$

For $y_1 = 1$, $D_x(1) = [0,1]$

$$F_1(x) = x + 1$$

and therefore,

$$\begin{aligned} P_{A(1)}(x) &= a_1(1)\phi_1(x) \\ &= a_1(1) \\ &= \frac{1}{2}\left[\max_{x\varepsilon[0,1]}(x + 1) + \min_{x\varepsilon[0,1]}(x + 1)\right] \\ &= \frac{3}{2} \quad . \end{aligned}$$

Therefore $a_1(y)$ is discontinuous on $D_y = [0,1]$.

Now, for $0 \leq y_1 < 1$

$$\rho(y_1) = ||F_{y_1} - P_{A(y_1)}||_{y_1}$$

$$= ||y_1 - y_1||_{y_1}$$

$$= 0.$$

For $y_1 = 1$

$$\rho(1) = ||F_1 - P_{A(1)}||_1$$

$$= \max_{x\varepsilon[0,1]} |(x + 1) - \tfrac{3}{2}|$$

$$= \frac{1}{2} .$$

Therefore, $\rho(y)$ is discontinuous on $D_y = [0,1]$.

<u>Theorem 2.4.7</u>: Let D, F and $\phi_1, \phi_2, \ldots, \phi_n, \psi_1, \psi_2, \ldots, \psi_m$, satisfy condition P.T., and let PT_A be the Product Tchebycheff approximation to F on D. Then for any real constant λ, the polynomial λPT_A is the Product Tchebycheff approximation to the function λF on D.

<u>Proof</u>: For each $y\varepsilon D_y$ let

$$P_{A(y)} = \sum_{j=1}^{n} a_j(y)\phi_j$$

be the best $||\cdot||_y$ approximation to F_y. Then the best $||\cdot||_y$ approximation to λF_y is

$$\lambda P_{A(y)} = \sum_{j=1}^{n} \lambda a_j(y)\phi_j \quad .$$

Thus the coefficients of $\phi_1, \phi_2, \ldots, \phi_n$ in $\lambda P_{A(y)}$ are respectively $\lambda a_1(y), \lambda a_2(y), \ldots, \lambda a_n(y)$.

Let Q_{A_j} be the best uniform approximation to the continuous function $a_j(y)$ on D_y, for $j = 1,2,\ldots,n$. Then λQ_{A_j} is the best uniform approximation to $\lambda a_j(y)$ on D_y, for $j = 1,2,\ldots,n$.

Hence the Product Tchebycheff approximation to λF is

$$\sum_{j=1}^{n} \lambda Q_{A_j} \phi_j = \lambda \sum_{j=1}^{n} Q_{A_j} \phi_j$$

$$= \lambda PT_A.$$

The following example illustrates that the Product Tchebycheff approximation and the unique best uniform approximation may be distinct from one another.

<u>Example 2.4.8</u>: Let $D = [-1,1]\times[-\frac{1}{2},1]$ and let

$$F(x,y) = x^2 y$$

$$\phi_1(x) = 1, \psi_1(y) = 1 \quad .$$

Then, $D_x = [-1,1]$ and $D_y = [-\frac{1}{2},1]$.

$$P_{A(y)}(x) = a_1(y)\phi_1(x)$$

$$= a_1(y)$$

$$= \frac{1}{2}\left[\max_{x\,[-1,1]} x^2 y + \min_{x\,[-1,1]} x^2 y\right]$$

$$= \begin{cases} \frac{1}{2}[0 + y], & \text{for } -\frac{1}{2} \le y \le 0 \\ \frac{1}{2}[y + 0], & \text{for } \ 0 < y \le 1 \end{cases}$$

$$= \frac{y}{2} \text{ for } y \varepsilon D_y = [-\tfrac{1}{2},1].$$

Therefore,

$$\begin{aligned}
PT_A(x,y) &= Q_{A_1}(y)\phi_1(x) \\
&= Q_{A_1}(y) \\
&= \frac{1}{2}\left[\max_{y\varepsilon[-\frac{1}{2},1]} \frac{y}{2} + \min_{y\varepsilon[-\frac{1}{2},1]} \frac{y}{2}\right] \\
&= \frac{1}{2}\left[\quad \frac{1}{2} \quad + \quad (-\frac{1}{4}) \quad\right] \\
&= \frac{1}{8} \quad .
\end{aligned}$$

The unique best uniform approximation to F on D is;

$$\begin{aligned}
P_{A*}(x,y) &= \frac{1}{2}\left[\max_{(x,y)\varepsilon D} x^2y + \min_{(x,y)\varepsilon D} x^2y\right] \\
&= \frac{1}{2}\left[\quad 1 \quad + \quad (-\frac{1}{2}) \quad\right] \\
&= \frac{1}{4} \quad .
\end{aligned}$$

We also note that the best uniform approximation to x^2 on $D_x = [-1,1]$, by a constant is $\frac{1}{2}$, and the best uniform approximation to y on $D_y = [-\frac{1}{2},1]$, by a constant is $\frac{1}{4}$. The product of these two constants is $\frac{1}{8}$, which is the Product Tchebycheff approximation to x^2y on $D = D_x \times D_y$ by a constant. This result is generalized in the following theorem.

<u>Theorem 2.4.9</u>: Let D,F and $\phi_1,\phi_2,\ldots,\phi_n$, $\psi_1,\psi_2,\ldots,\psi_m$ satisfy condition P.T., and let

$$F(x,y) = H(x)G(y), \text{ on } D = D_x \times D_y.$$

Let $P_{A*} = \sum_{j=1}^{n} a_j{}^*\phi_j$ and Q_{B*} be the best uniform approximations respectively to H on D_x in $\phi_1,\phi_2,\ldots,\phi_n$ and to G on D_y in $\psi_1,\psi_2,\ldots,\psi_m$.

Then, the Product Tchebycheff approximation to F on D is

$$PT_A = P_{A*}Q_{B*}.$$

Proof: For each fixed $y_1 \varepsilon D_y$, the best $||\cdot||_{y_1}$ approximation to $F_{y_1} = G(y_1)H$ is

$$\begin{aligned} P_{A(y_1)} &= G(y_1)P_{A*} \\ &= \sum_{j=1}^{n} a_j{}^*G(y_1)\phi_j \quad . \end{aligned}$$

For each $j = 1,2,\ldots,n$, the best uniform approximation to the coefficient function $a_j{}^*G$ on D_y is

$$a_j{}^*Q_{B*}.$$

Hence

$$\begin{aligned} PT_A &= \sum_{j=1}^{n} a_j{}^*Q_{B*}\phi_j \\ &= \left[\sum_{j=1}^{n} a_j{}^*\phi_j\right]Q_{B*} \\ &= P_{A*}Q_{B*} \quad . \end{aligned}$$

We now consider a special case for which there is a finite algorithm to find the Product Tchebycheff approximation.

Corollary 2.4.10: Let $D = [-1,1]\times[-1,1]$ and let $F(x,y) = x^n y^m$

$$\phi_1(x) = 1,\ \phi_2(x) = x; \ldots,\ \phi_n(x) = x^{n-1}$$
$$\psi_1(y) = 1,\ \psi_2(y) = y; \ldots,\ \psi_m(y) = y^{m-1}.$$

Then, the Product Tchebycheff approximation to F on D is

$$PT_A(x,y) = (2^{1-n}T_n(x) - x^n)(2^{1-m}T_m(y) - y^m)$$

where T_n and T_m are the well-known Tchebycheff polynomials defined by

$$T_k(x) = \cos(k \cos^{-1} x),\ k = 1,2,\ldots \quad .$$

Proof: The best uniform approximation to x^n on $D_x = [-1,1]$ is $2^{1-n}T_n(x) - x^n$, and the best uniform approximation to y^m on $D_y = [-1,1]$ is $2^{1-m}T_m(y) - y^m$.

Hence, we apply Theorem 2.4.9 to achieve the desired result.

Theorem 2.4.11: Let D, F and $\phi_1,\phi_2,\ldots,\phi_n$, $\psi_1,\psi_2,\ldots,\psi_m$ satisfy condition P.T. and let

$$F(x,y) = H(x) + G(y) \text{ on } D = D_x\times D_y.$$

Let $\phi_1(x) = 1$ and $\psi_1(y) = 1$ (or equivalently any non-zero constants), and let

$$P_{A*} = \sum_{j=1}^{n} a_j{}^*\phi_j = a_1{}^* + \sum_{j=2}^{n} a_j{}^*\phi_j \quad ,$$

and Q_{B*} be the best uniform approximations respectively to H on D_x in $\phi_1 = 1, \phi_2, \ldots, \phi_n$ and to G on D_y in $\psi_1 = 1, \psi_2, \ldots, \psi_m$.

Then, the Product Tchebycheff approximation to F on D is

$$PT_A = P_{A*} + Q_{B*}.$$

<u>Proof</u>: For each fixed $y_1 \varepsilon D_y$, the best $||\cdot||_{y_1}$ approximation to the function F_{y_1} defined by

$$F_{y_1}(x) = H(x) + G(y_1)$$

is

$$P_{A(y_1)} = P_{A*} + G(y_1)$$

$$= a_1{}^* + G(y_1) + \sum_{j=2}^{n} a_j{}^*\phi_j.$$

For each $j = 2,3,\ldots,n$ the best uniform approximation to the constant function $a_j{}^*$ on D_y is $a_j{}^*$ itself. Also, the best uniform approximation to the coefficient function $a_1{}^* + G(y)$ on D_y is

$$a_1{}^* + Q_{B*}.$$

Hence,

$$PT_A = a_1{}^* + Q_{B*} + \sum_{j=2}^{n} a_j{}^*\phi_j$$

$$= P_{A*} + Q_{B*} \quad .$$

The following counter-example illustrates that it is necessary to include both a non-zero constant as one of the base functions $\phi_1, \phi_2, \ldots, \phi_n$ and another non-zero constant as one of the base functions $\psi_1, \psi_2, \ldots, \psi_m$, in the hypothesis of Theorem 2.4.11.

<u>Example 2.4.12</u>: Let $D = [1,2]\times[1,2]$ and let

$$F(x,y) = x + y$$

$$\phi_1(x) = x$$

$$\psi_1(y) = 1, \psi_2(y) = y.$$

Note that $\phi_1(x) = x$ is a trivial Tchebycheff system on $D_x = [1,2]$ since for each real $\alpha \neq 0$, $\alpha x = 0$ has no solutions in $[1,2]$.

Assertion: The best $||\cdot||_{y_1}$ approximation to F_{y_1} defined by

$$F_{y_1}(x) = x + y_1$$

is

$$P_{A(y_1)} = (1 + \frac{2}{3} y_1)\phi_1 ,$$

for each fixed $y_1 \varepsilon D_y = [1,2]$.

<u>Proof</u>: For each $y_1 \varepsilon [1,2]$

$$F_{y_1}(x) - (1 + \frac{2}{3} y_1)x = x + y_1 - (1 + \frac{2}{3} y_1)x$$

$$= y_1(1 - \frac{2}{3} x),$$

which is a strictly decreasing function of x.

$$y_1(1 - \frac{2}{3}x) = \begin{cases} \frac{y_1}{3} > 0 & \text{for } x = 1 \\ -\frac{y_1}{3} < 0 & \text{for } x = 2 \end{cases} .$$

Therefore,

$$||F_{y_1} - (1 + \tfrac{2}{3}y_1)\phi_1||_{y_1} = \max_{x\varepsilon[1,2]} |F_{y_1}(x) - (1 + \tfrac{2}{3}y_1)x|$$

$$= \frac{y_1}{3} .$$

To decrease the norm of the error we must find a polynomial αx, which is negative at $x = 1$ and positive at $x = 2$. However, $\alpha \cdot 1 < 0 \Rightarrow \alpha < 0$

$$\Rightarrow \alpha \cdot 2 < 0. \quad \text{q.e.d. (assertion)}$$

The best uniform approximation to the coefficient function $1 + \frac{2}{3}y$, on D_y by a polynomial of the form $a + by$ is $1 + \frac{2}{3}y$ itself. Therefore,

$$PT_A(x,y) = (1 + \tfrac{2}{3}y)x.$$

Now the best uniform approximation to x on D_x by a polynomial of the form ax is

$$P_{A*}(x) = x,$$

and the best uniform approximation to y on D_y by a polynomial of the form $a + by$ is

$$Q_{B*}(y) = y.$$

Therefore,

$$P_{A*}(x) + Q_{B*}(y) = x + y$$
$$\neq (1 + \tfrac{2}{3}y)x = PT_A(x,y).$$

Similarly if we let F and D be as above and let

$$\phi_1(x) = 1, \ \phi_2(x) = x$$
$$\psi_1(y) = y \quad ,$$

then we can show that the Product Tchebycheff approximation to F on D is PT_A as defined by

$$PT_A(x,y) = (1 + \tfrac{2}{3}x)y$$
$$\neq x + y \quad .$$

Section 5: The First Product Tchebycheff Algorithm

Most best uniform approximation problems require an infinite algorithm for their solutions. One of the more frequently used procedures is the Remez exchange algorithm, sometimes referred to as Remez' second algorithm.

We precede the description of this procedure with a definition and theorem which form an integral part of the underlying theory.

Definition 2.5.1: Let D be a compact set in E_1 and let $F \varepsilon C(D)$. Let $\phi_1, \phi_2, \ldots, \phi_n$ be a Tchebycheff system of continuous functions on D, and let $||\cdot||$ be the uniform norm on D. Then $x_1 \varepsilon D$ is called a positive (negative) extremal or an E + (E-) point for $F - P_A = F - \sum_{i=1}^{n} a_i \phi_i$

if $F(x_1) - P_A(x_1) = ||F - P_A||$ $(F(x_1) - P_A(x_1) = -\ ||F - P_A||)$.

Theorem 2.5.2: P_A is the best uniform approximation to F on D if and only if there are n+1 points

$$x_1 < x_2 < \ldots < x_{n+1} \quad \text{in D}$$

which are alternately E+ and E- points for $F - P_A$.

Note that this characteristic point set need not be unique.

The Remez exchange algorithm is an iterative procedure which seeks to find a set of n+1 alternating extremals.

Remez Exchange Algorithm 2.5.3

(1) Choose $x_1^{(o)} < x_2^{(o)} < \ldots x_{n+1}^{(o)}$ in D as an initial guess for a set of n+1 alternating extremals.

(2) For $x_1^{(k)} < x_2^{(k)} < \ldots < x_{n+1}^{(k)}$ in D solve the linear system

$$P_{A^{(k)}}(x_j^{(k)}) + (-1)^j \rho^{(k)} = F(x_j^{(k)}),\ j = 1,2,\ldots,n+1$$

for $A^{(k)} = (a_1^{(k)}, a_2^{(k)}, \ldots, a_n^{(k)})$ and $\rho^{(k)}$.

If $||F - P_{A^{(k)}}|| = \rho^{(k)}$, then $\{x_j^{(k)}\}$ is an extremal point set for $F - P_{A^{(k)}}$, and therefore by Theorem 2.5.2 $P_{A^{(k)}}$ is the best uniform approximation to F on D.

If $||F - P_{A^{(k)}}|| > \rho$, then we choose

$$x_1^{(k+1)} < x_2^{(k+1)} < \ldots < x_{n+1}^{(k+1)} \text{ in D}$$

such that

(1) each $x_j^{(k+1)}$, $j = 1,2,\ldots,n+1$ is a relative maxima or minima of $F - P_{A^{(k)}}$

(2) for some x_{max} $\{x_j^{(k+1)}\}$

$$|F(x_{max}) - P_{A^{(k)}}(x_{max})| = ||F - P_{A^{(k)}}||$$

and (3) $\operatorname{sgn}[f(x_j^{(k+1)}) - P_{A^{(k)}}(x_j^{(k+1)})]$

$$= (-1)^{j+1}\operatorname{sgn}[F(x_1^{(k+1)}) - P_{A^{(k)}}(x_1^{(k+1)})] \quad ,$$

$$j = 1,2,\ldots,n+1 \quad .$$

Repeat step (2).

The convergence of this procedure is outlined in Remez (17), and proved in Novodvorskii and Pinsker (13). Verdinger (26) shows that if F is differentiable then the rate of convergence is quadratic.

In practice some $\varepsilon > 0$ is prescribed and the iterations are terminated when

$$||F - P_{A^{(k)}}|| - \rho^{(k)} < \varepsilon \quad .$$

Fraser and Hart (6) suggest that it is often advantageous to choose the extrema of T_n, (the Tchebycheff polynomial of degree n) for the set $\{x_j^{(o)}\}$, when $D = [-1,1]$, or equivalently let

$$\{x_j\} = \{\frac{a+b}{2} + \frac{b-a}{2} \cos \frac{j-1}{n} \pi\}$$

when D = [a,b] .

The de la Vallée Poussin algorithm for finding the best uniform approximation is described in Rice (18) as follows:

Let $X = \{x_i : i=1,2,\ldots\}$ be a dense subset of the compact set D, and let

$$X_m = \{x_i : i = 1,2,\ldots,m\} \subset X.$$

We define the density of X_m in D by

$$\delta_m = \max_{x \varepsilon D} \sigma_1[x;X_m] .$$

Let P_{A_m} be the best Tchebycheff approximation to F on X_m, and let P_{A*} be the best Tchebycheff approximation to F on D.

One would hope that

$$\lim_{m \to \infty} P_{A_m} = P_{A*} .$$

For a restricted class of functions, Rice establishes the following stronger result.

<u>Theorem 2.5.4</u>: Let D be a compact set in E, and let $F \varepsilon C(D)$. Let $\phi_1,\phi_2,\ldots,\phi_n$ be a Tchebycheff system of continuous base functions on D, and let F and $\phi_1,\phi_2,\ldots,\phi_n$ satisfy a Hölder Condition with exponent α. Then,

$$|| P_{A^*} - P_{A_m} || \leq K\delta_m^{\alpha} ,$$

Where K is a constant which depends only on F and $\{\phi_i\}$, and $||.||$ is the uniform norm on D.

Now if Y is any subset of D, we define the density of Y in D to be the number

$$\delta_Y = \max_{x\varepsilon D} \inf_{y\varepsilon Y} \sigma(x;y) \quad .$$

Rivlin and Cheney (20) prove the following more general theorem:

<u>Theorem 2.5.5</u>: 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_{A^*} in M. For any $Y \subset D$ let P_{A_Y} denote a best approximation to F from M on the set Y. Then as $\delta_Y \to 0$, $P_{A_Y} \to P_{A^*}$ uniformly.

We can now define the first Product Tchebycheff algorithm.

(1) Choose some finite point set $Y \subset D_y$.

(2) For each $y \varepsilon Y$, use the Remez exchange algorithm to find $P_{A(y)}$ the best $||.||_y$ approximation to F_y on $D_x(y)$.

(3) For each $j = 1,2,...,n$ use the Remez exchange algorithm to find $\hat{Q}_{A_j}$ the best uniform approximation to $a_j(y)$ on the set Y.

Then $\hat{PT}_A = \sum_{j=1}^{n} \hat{Q}_{A_j} \phi_j$ is the Product Tchebycheff approximation to F on $\hat{D} = \{(x,y) \varepsilon D: y \varepsilon Y\}$.

Now given $\varepsilon > 0$, we see by Theorem 2.5.5 that we can choose Y in D_y such that the density of Y in D_y is sufficiently small to have

$$\max_{y \varepsilon D_y} | \hat{Q}_{A_j}(y) - A_{A_j}(y) | < \varepsilon / \max_{x \varepsilon D_x} | \phi_j(x) |$$

For $j = 1,2,\ldots,n.$

Hence,

$$\max_{(x,y)\varepsilon D} | \hat{PT}_A(x,y) - PT_A(x,y) |$$

$$= \max_{(x,y)\varepsilon D} \left| \sum_{j=1}^{n} [\hat{Q}_{A_j}(y) - Q_{A_j}(y)] \phi_j(x) \right|$$

$$\leq \varepsilon \ .$$

Therefore, if Y is chosen so that the density of Y in D_y is sufficiently small then $\hat{PT}_A$ provides a good estimate for PT_A.

A program utilizing this algorithm was written for the Control Data Corporation 3600 digital computer, with the following restrictions:

(1) $Y = \{y_i: i = 0,1,\ldots,100\}$.

(2) For each $y_i \varepsilon Y$, the corresponding set $D_x(y_i)$ was a closed bounded interval.

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

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

The norm of $F - \hat{PT}_A$ was estimated by N^* where,

$$D_x = [a,b] \ , \ h_1 = \frac{b-a}{100}$$

$$D_y = [c,d] \ , \ h_2 = \frac{d-c}{100}$$

and $N^* = \max\limits_{\substack{x=a,a+h_1,\ldots,b \\ y=c,c+h_2,\ldots,d}} |F(x,y) - \hat{PT}_A(x,y)|$.

In all cases single-precision arithmetic was used.

Some of the results obtained are described below.

I. $F(x,y) = \dfrac{1}{x+y+1}$, $D = [0,1] \times [0,1]$

$Y = \{0.00, 0.01, 0.02, \ldots, 1.00\}$

$\hat{D} = [0,1] \times Y$

$n = 3$, $m = 3$

$$\begin{aligned}\hat{PT}_A(x,y) = {} & .0315641979\ x^2y^2 - .3075877279\ x^2y \\ & + .3102885292\ x^2 \\ & + .3689560988\ xy^2 + .2346341978\ xy \\ & - .7160087106\ x \\ & + .3348278992\ y^2 - .8141539737\ y \\ & + .9855581665\ ,\end{aligned}$$

$N^* = .09333756560.$

II. $F(x,y) = \dfrac{1}{x+y+10}$ $D = [0,1] \times [0,1]$

$Y = 0.00, 0.01, 0.02, \ldots, 1.00$

$\hat{D} = [0,1] \times Y.$

(a) $n = 2$, $m = 2$

$$\hat{PT}_A(x,y) = .009090909094\ xy - .01283994207\ x$$
$$- .009065006705\ y + .09978401812\ ,$$
$$N^* = .003965014850.$$

(b) $n = 3,\ m = 3$

$$\hat{PT}_A(x,y) = .002546904012\ x^2y^2 + .002313865341\ x^2y$$
$$+ .0005496653156\ x^2$$
$$+ .03203501519\ xy^2 - .02999156816\ xy$$
$$- .005997745788\ x$$
$$+ .0008660139467\ y^2 - .009951005503\ y$$
$$+ .09999485663\ ,$$
$$N^* = .003938859616.$$

III. It was noted in example 2.3.8 that if

$$F(x,y) = \sqrt{x}$$
$$\phi_1(x) = 1,\ \phi_2(x) = x$$

and $D = \{(x,y):\ 0 \le y \le 1 \text{ and } 0 \le x \le 1 - y\}$,

Then $P_{A(y)}$ is not well-defined at $y = 1$.

We will approximate F on D as follows:

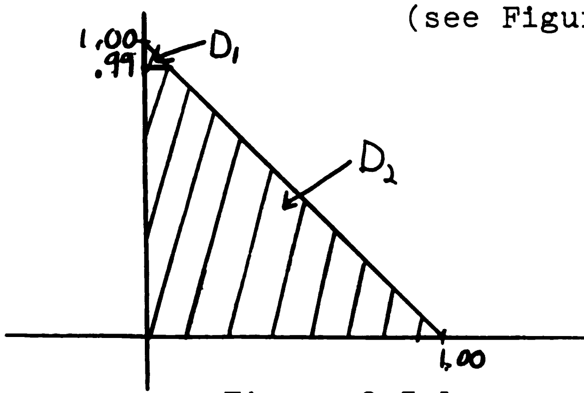
Let $D = D_1 \cup D_2$ where

$$D_1 = \{(x,y):\ .99 \le y \le 1 \text{ and } 0 \le x \le 1 - y\}$$

and

$$D_2 = \{(x,y):\ 0 \le y \le .99 \text{ and } 0 \le x \le 1 - y\}$$

(see Figure 2.5.1).



We approximate F on D_1 by the constant

$$\frac{1}{2}\left[\max_{x\varepsilon[0,.01]} \sqrt{x} + \min_{x\varepsilon[0,.01]} \sqrt{x}\right]$$

$$= \frac{1}{2}[.1 + 0] = .05 ,$$

and we approximate F on D_2 by a product Tchebycheff approximation as follows:

$$Y = \{0.00, 0.0099, 0.0198, 0.0297, \ldots, 0.99\}$$

$$\hat{D} = \{(x,y) : y \varepsilon Y,\ 0 \le x \le 1 - y\}$$

$$n = 2,\ m = 2$$

$$P\hat{T}_A(x,y) = -1.010101010\ xy + 4.539406039\ x - .1136363637\ y + .1365047380 ,$$

$$N^* = .1365047380 .$$

Section 6: The Revised Product Tchebycheff Algorithm

The Remez exchange algorithm described in 2.5.3 seeks to find a set of n + 1 alternating extremals which characterize the unique best uniform approximation. If we desire to find the Product Tchebycheff approximate we must find $P_{A(y)}$, the unique best uniform approximation to F_y on $D_x(y)$, for a number of distinct values of $y \varepsilon D_y$.

In this section we shall show that for $y_1, y_2 \varepsilon D_y$, a set of characteristic extremals for $F_{y_1} - P_{A(y_1)}$ will often be a good initial guess at a corresponding set for $F_{y_2} - P_{A(y_2)}$, whenever y_2 is sufficiently close to y_1.

We shall assume throughout the remainder of this section that D,F and $\phi_1, \phi_2, \ldots, \phi_n, \psi_1, \psi_2, \ldots, \psi_m$ satisfy condition P.T. .

Definition 2.6.1: Let R be the continuous function defined on D by

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

$$\text{Then } \rho(y) = \max_{x\varepsilon D_x(y)} |R(x,y)| \; .$$

Definition 2.6.2: For each $y \varepsilon D_y$ we define

$$E(y) = \{x \varepsilon D_x(y): \quad |R(x,y)| = \rho(y)\} \; ,$$

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

E(y) is a non-empty compact subset of $D_x(y)$.

Definition 2.6.3: For each $y \varepsilon D_y$ and each $\varepsilon > 0$ we define

$$E(y)_\varepsilon = \{x \varepsilon D_x: \quad \sigma[x;E(y)] < \varepsilon\}$$

(where $\sigma[x;Y] = \inf_{y\varepsilon Y} \sigma[x;y]$) .

Theorem 2.6.4: Given $\varepsilon > 0$ and $y_1 \varepsilon D_y$, there exists a $\delta = \delta(\varepsilon, y_1) > 0$ such that $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta$$

imply that

$$E(y_2) \subset E(y_1)_\varepsilon \; .$$

Proof:

Case 1: $E(y_1)_{\frac{\varepsilon}{2}} \supset D_x(y_1)$.

Then for each $x_1 \varepsilon D_x(y_1)$

$$\sigma[x_1;E(y_1)] > \frac{\varepsilon}{2} \; .$$

By property K, there exists a $\delta = \delta(\varepsilon/2) > 0$, such that $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta$$

imply that for each $x_2 \varepsilon D_x(y_2)$ there exists a corresponding $x_1 \varepsilon D_x(y_1)$ satisfying

$$|x_1 - x_2| < \frac{\varepsilon}{2} .$$

This implies that

$$\sigma[x_2;E(y_1)] \leq |x_2 - x_1| + \sigma[x_1;E(y_1)]$$

$$< \quad \frac{\varepsilon}{2} \quad + \quad \frac{\varepsilon}{2} = \varepsilon .$$

Therefore, $D_x(y_2) \subset E(y_1)_\varepsilon$.

Then since $E(y_2) \subset D_x(y_2)$ we have $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta$$

imply that

$$E(y_2) \subset E(y_1)_\varepsilon .$$

Case 2: $E(y_1)_{\frac{\varepsilon}{2}} \not\supset D_x(y_1)$

or equivalently $D_x(y_1) - E(y_1)_{\frac{\varepsilon}{2}}$ is a non-empty compact set.

Let $M = M(\varepsilon,y_1) = \max\limits_{x\varepsilon D_x(y_1)-E(y_1)_{\frac{\varepsilon}{2}}} |R(x,y_1)| < \rho(y_1).$

Then by the uniform continuity of $\rho(y)$ on D_y, since $\frac{\rho(y_1)-M}{2} > 0$ there exists a $\delta_1 = \delta_1 \left(\frac{\rho(y_1)-M}{2}\right) > 0$ such that

$y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta_1$$

imply that

$$\rho(y_1) > \rho(y_1) - \left(\frac{\rho(y_1)-M}{2}\right) = \frac{\rho(y_1)+M}{2}$$

.

Now by the uniform continuity of R on the compact set D, there exists a $\delta_2 = \delta_2 \left(\frac{\rho(y_1)-M}{2}\right) > 0$ such that $(x_1,y_1),(x_2,y_2) \varepsilon D$ and $\sigma[(x_1,y_1);\ (x_2,y_2)] < \delta_2$

imply that

$$|R(x_2,y_2)| < |R(x_1,y_1)| + \frac{\rho(y_1)-M}{2} .$$

Also, if $x_2 \varepsilon D_x(y_2) - E(y_1)_\varepsilon$

then $\sigma_1[x_2,E(y_1)] \geq \varepsilon$. Then $x_1 \varepsilon D_x(y_1)$

and

$$\sigma_2[(x_1,y_1);\ (x_2,y_2)] < \frac{\varepsilon}{2}$$

imply that

$$|x_1 - x_2| < \frac{\varepsilon}{2} .$$

Therefore,

$$\sigma[x_1;E(y_1)] \geq \sigma[x_2;E(y_1)] - |x_1 - x_2|$$

$$> \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2} ,$$

which implies that $x_1 \varepsilon D_x(y_1) - E(y_1)_{\frac{\varepsilon}{2}}$.

Now if we choose $\delta_3 = \min(\delta_2, \frac{\varepsilon}{2}) > 0$

then by property K, there exists a

$\delta_4 = \delta_4(\delta_3) > 0$, such that $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta_4$$

imply that for each $x_2 \varepsilon D_x(y_2) - E(y_1)_\varepsilon$

there is a corresponding $x_1 \varepsilon D_x(y_1)$ such that

$$\sigma[(x_1,y_1); (x_2,y_2)] < \delta_3 .$$

Therefore, $x_1 \varepsilon D_x(y_1) - E(y_1)_{\frac{\varepsilon}{2}}$ which implies that

$|R(x_1,y_1)| \leq M$. Thus,

$$\begin{aligned} |R(x_2,y_2)| &\leq |R(x_1,y_1)| + \frac{\rho(y_1)-M}{2} \\ &\leq M + \frac{\rho(y_1)-M}{2} \\ &= \frac{\rho(y_1)+M}{2} . \end{aligned}$$

If $D_x(y_2) - E(y_1)_\varepsilon = \Phi$, then

$E(y_1)_\varepsilon \supset D_x(y_2) \supset E(y_2)$, and the proof is complete.

If $D_x(y_2) - E(y_1)_\varepsilon \neq \Phi$, then we

choose $\delta = \min\ (\delta_1,\delta_4) > 0$. $y_2 \varepsilon D_y$ and

$$|y_2 - y_1| < \delta$$

imply

(1) $\rho(y_2) > \dfrac{\rho(y_1)+M}{2}$

and (2) For each $x_2 \varepsilon D_x(y_2) - E(y_1)_\varepsilon$

$$|R(x_2,y_2)| < \frac{\rho(y_1)+M}{2} < \rho(y_2) ,$$

or equivalently

$$E(y_2) \cap [D_x(y_2) - E(y_1)_\varepsilon] = \Phi .$$

Then since $E(y_2) \subset D_x(y_2)$ we have

$$E(y_2) \subset E(y_1)_\varepsilon \ .$$

We now consider the implications of theorem 2.6.4:

Example 2.6.5: Let $D = [0,4] \times [0,2]$

and let $F(x,y) = y\sin\pi x$, for $0 \leq y \leq 1$

$y\sin\pi x$, for $0 \leq x \leq 1$ and $1 < y \leq 2$

$\sin\pi x$, for $1 < x \leq 2$ and $1 < y \leq 2$.

Let $\phi_1(x) = 1$.

Then for $0 \leq y \leq 1$

$||F_y||_y = y$, and $F_y\left(\frac{1}{2}\right) = - F_y\left(\frac{3}{2}\right) = F_y\left(\frac{5}{2}\right) = - F_y\left(\frac{7}{2}\right) = y$.

Therefore the best uniform approximation to

F_y by a constant is 0 and

$$E(y) = \{\tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \tfrac{7}{2}\} \ .$$

Also for $1 < y \leq 2$

$||F_y||_y = y$, and $F_y\left(\frac{1}{2}\right) = - F_y\left(\frac{3}{2}\right) = y$.

Therefore the best uniform approximation for

F_y by a constant is 0 and

$$E(y) = \{\tfrac{1}{2}, \tfrac{3}{2}\}$$ See Figure 2.6.1.

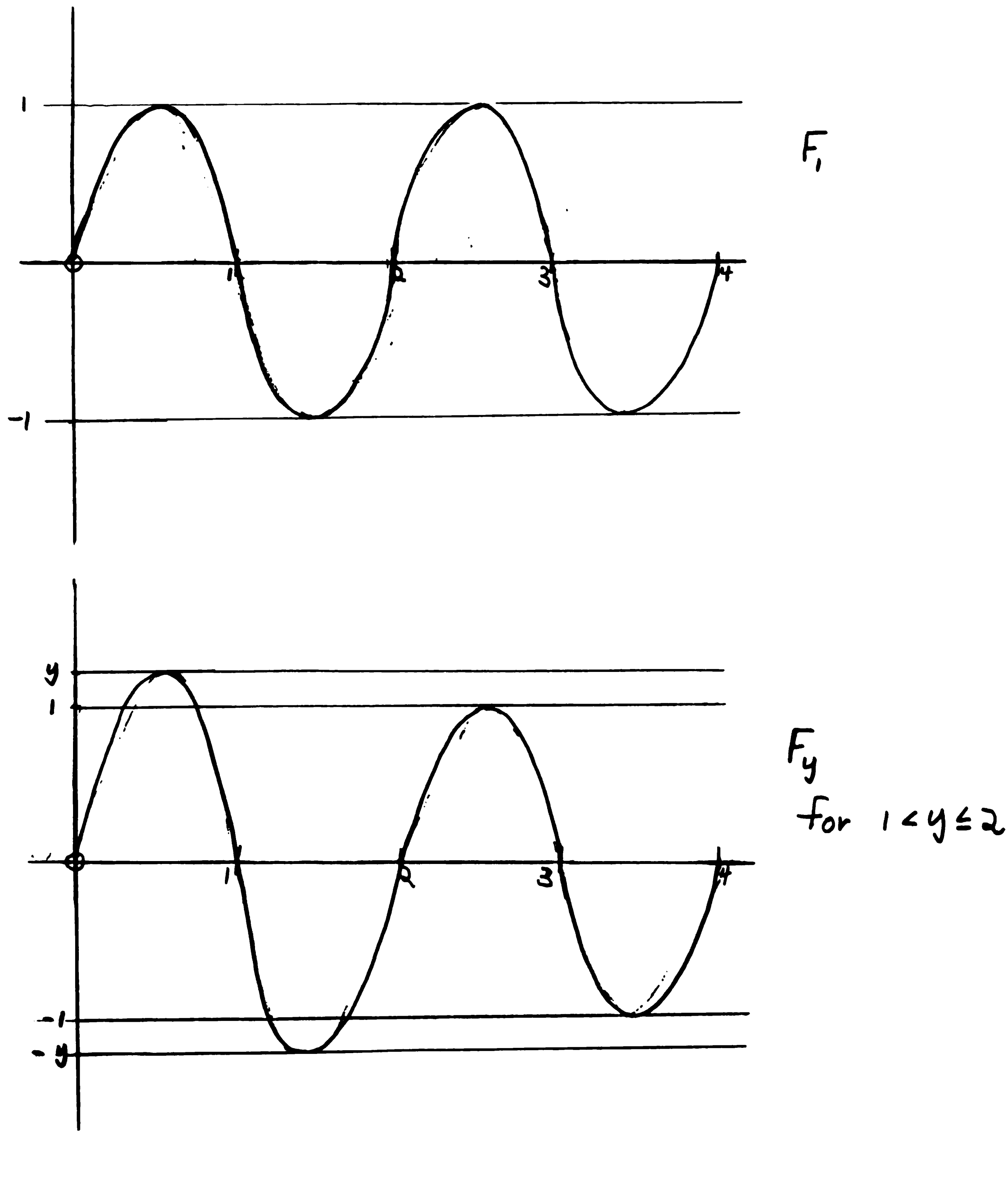


Figure 2.6.1

Note that $F_y\left(\frac{5}{2}\right) = - F_y \left(\frac{7}{2}\right) = 1 < y$,

which implies that $\frac{5}{2}$ and $\frac{7}{2}$ are not extremals for

$F_y - P_{A(y)} = F_y$.

Now if we choose $\varepsilon = \frac{1}{2}$ and $y_1 = 1 \varepsilon D_y = [0,2]$,

then by theorem 2.6.3 there exists a $\delta = \delta\left(\frac{1}{2},1\right) > 0$

such that $y_2 \varepsilon [0,2]$ and

$$|y_2 - 1| < \delta$$

implies that all extremals of $F_{y_2} - P_{A(y_2)} = F_{y_2}$

are within $\frac{1}{2}$ of $E(1) = \{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \frac{7}{2}\}$.

Any choice of $\delta > 0$ will do. In particular we can choose $\delta = 1$. Consider $y_2 = \frac{3}{2}$.

$$E\left(\frac{3}{2}\right) = \{\frac{1}{2}, \frac{3}{2}\} \subset E(1) = \{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \frac{7}{2}\} .$$

Therefore, each point of $E\left(\frac{3}{2}\right)$ lies within $\frac{1}{2}$ of some point of $E\left(\frac{1}{2}\right)$. However there is no point of $E\left(\frac{3}{2}\right)$ which lies within $\frac{1}{2}$ of $\frac{5}{2}$ or $\frac{7}{2}$ which belong to $E\left(\frac{1}{2}\right)$.

This can be explained by the fact that while

$$|y_2 - y_1| = |\tfrac{3}{2} - 1| < \delta(\varepsilon,y_1) = \delta\left(\tfrac{1}{2}, 1\right) = 1$$

we can <u>not</u> have

$$|y_1 - y_2| = \tfrac{1}{2} < \delta(\varepsilon,y_2) = \delta\left(\tfrac{1}{2}, \tfrac{3}{2}\right) .$$

However $y_2 \varepsilon [0,2]$ and

$$|y_3 - y_2| = |y_3 - \tfrac{3}{2}| < \tfrac{1}{2}$$

imply that

$$E(y_3) = \{\tfrac{1}{2}, \tfrac{3}{2}\} \subset E\left(\tfrac{3}{2}\right) = \{\tfrac{1}{2}, \tfrac{3}{2}\}.$$

Therefore, we can choose $\delta\left(\tfrac{1}{2}, \tfrac{3}{2}\right) = \tfrac{1}{2}$.

<u>Corollary 2.6.6</u>: Let $\phi_1(x) = 1$, $\phi_2(x) = x, \ldots, \phi_n(x) = x^{n-1}$, and assume that

(1) $\dfrac{d^{n+1}}{dx^{n+1}} F_{y_1}(x) \neq 0$ for all $y_1 \varepsilon D_y$ and for all $x \varepsilon D_x(y_1)$.

Then there are n + 1 continuous functions $x_1(y), x_2(y), \ldots, x_{n+1}(y)$, such that $x_1(y), < x_2(y) < \ldots < x_{n+1}(y)$ is the set of n+1 characteristic extremals for $F_y - P_{A(y)}$, for each $y \varepsilon D_y$.

Proof: (1) implies that $F_y - P_{A(y)}$ has exactly n + 1 extremals, for each $y \varepsilon D_y$. We apply theorem 2.6.4 to complete the proof.

More generally if $F_y - P_{A(y)}$ has exactly n + 1 extremals for each y D_y, then the conclusion of Corollary 2.6.6 holds.

Theorem 2.6.4 and Corollary 2.6.6 suggest that for y_2 and y_1 sufficiently close, a set of n + 1 alternating extremals for $F_{y_1} - P_{A(y_1)}$ might be a good starting point for the Remez exchange algorithm to find $P_{A(y_2)}$.

A program incorporating this modification was written for the Control Data Corporation 3600 digital computer. Some of the results obtained are described below.

I. $F(x,y) = \sin xy, D = [0,1] \times [1,2]$

$Y = \{0.00, 0.01, 0.02, \ldots, 1.00\}$

$\hat{D} = [0,1] \times Y$

$n = 2,\ m = 2$

$$PT_A(x,y) = -\ .841470984\ xy + 2.135441294\ x$$
$$+\ .1655503163\ y - .163795902$$
$$N^* = .5030052628$$

Execution time for revised algorithm

1 minute 59.159 seconds.

This problem was also solved using the first algorithm as described in section 5. The execution time for the first algorithm was

2 minutes 8.249 seconds.

II. The following two examples were run using both the first Product Tchebycheff algorithm and the revised Product Tchebycheff algorithm.

$F(x,y) = \frac{1}{x+y+10}$, $D = [0,1] \quad [0,1]$

$Y = \{0.00, 0.01, 0.02, \ldots, 1.00\}$

$\hat{D} = [0,1] \times Y.$

See problems II(a) and II(b) at the end of section 5, for the corresponding $PT_A(x,y)$ and N^*.

(a) $n = 2$, $m = 2$

Execution time for first algorithm

1 minute 43.629 seconds.

Execution time for revised algorithm

1 minute 12.824 seconds.

(b) $n = 3$, $m = 3$

Execution time for first algorithm

1 minute 12.587 seconds.

Execution time for revised algorithm

1 minute 5.956 seconds.

Section 7: The Degree of Product Tchebycheff Approximation

Definition 2.7.1: Let $\mathfrak{F}$, $\|\cdot\|$ be a normed linear space with base field R. Then a sequence $\{\phi_j\}$ in $\mathfrak{F}$ is called closed in $\mathfrak{F}$ if and only if for each $F \in \mathfrak{F}$ and arbitrary $\epsilon > 0$ there exists an integer $n > 0$ and a corresponding polynomial

$$P_A = \sum_{j=1}^{n} a_j \phi_j$$

such that

$$\|F - P_A\| < \epsilon .$$

Remark 2.7.2: $\{\phi_j\}$ is closed in $\mathfrak{F}$ if and only if the set of all finite linear combinations of the ϕ_j's is a dense subset of $\mathfrak{F}$.

Remark 2.7.3: By Weierstrass' First Approximation Theorem we know that

$$\{1,x,x^2,\ldots\}$$

is a closed sequence in C(D).

Remark 2.7.4: By Weierstrass' Second Approximation Theorem we know that

$$\{1,\sin x,\cos x,\sin 2x,\cos 2x,\ldots\}$$

is a closed sequnece in $C_{2\pi}$, the linear space of all real-valued continuous functions such that

$$F(x+2\pi) = F(x) \text{ for all } x\varepsilon E_1 .$$

Theorem 2.7.5: For each n > 0 and m > 0 let D, F and $\phi_1,\phi_2,\ldots,\phi_n,\psi_1,\psi_2,\ldots,\psi_m$ satisfy condition P.T. and let $\{\phi_j\}_{j=1}^{\infty}$ and $\{\psi_i\}_{i=1}^{\infty}$ be closed sequences in $C(D_x)$ and $C(D_y)$ respectively.

Then given $\varepsilon > 0$ and $F \varepsilon C(D)$ there exists an $N = N(\varepsilon)$ and an $M = M(\varepsilon,n)$ for each $n > N(\varepsilon)$, such that $n > M$ and $m > M$ imply that

$$||F - PT_A|| < \varepsilon$$

where PT_A is the Product Tchebycheff approximation in $\phi_1,\phi_2,\ldots,\phi_n,\psi_1,\psi_2,\ldots,\psi_m$, and $||\cdot||$ is the uniform norm on D.

Proof: $\max\limits_{(x,y)\varepsilon D} |F(x,y) - P_{A(y)}(x)|$

$$= \max_{y \varepsilon D_y} \left[\max_{x \varepsilon D_x(y)} |F_y(x) - P_{A(y)}(x)| \right]$$

$$= \max_{y \varepsilon D_y} \rho(y) .$$

By the uniform continuity of $\rho(y)$ on D_y, given $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon/4) > 0$, such that $y_1, y_2 \varepsilon D_y$ and

$$|y_1 - y_2| < \delta$$

imply that

$$\rho(y_1) < \rho(y_2) + \varepsilon/4 .$$

Since D_y is compact we can choose a finite set

$$Y_k = \{y_1, y_2, \ldots, y_k\} \subset D_y$$

such that for each $y \varepsilon D_y$, there exists a corresponding $y_i \varepsilon Y_k$ satisfying

$$|y - y_i| < \delta(\varepsilon/4) .$$

This implies that

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

Now since $\{\phi_j\}$ is closed in $C(D_x)$, given $\varepsilon > 0$, there exists an $N_i = N_i(\varepsilon/4)$ such that for $n > N_i$

$\rho(y_i) < \varepsilon/4$, for $i = 1,2,\ldots,k$.

Choose $N = \max(N_1, N_2, \ldots, N_k)$. Then $n > N$ implies that

$\rho(y_i) < \varepsilon/4$ for $i = 1,2,\ldots,k$.

Then for each $y \varepsilon D_y$ there exists a corresponding $y_i \varepsilon Y_k$ such that

$$\rho(y) < \rho(y_i) + \varepsilon/4$$

$$< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2} .$$

Therefore $n > N$ implies that

$$\max_{(x,y)} |F(x,y) - P_{A(y)}(x)| = \max_{y \varepsilon D_y} \rho(y) < \frac{\varepsilon}{2} .$$

Now since $\phi_1, \phi_2, \ldots, \phi_n$ is a Tchebycheff set on D_x, they are linearly independent there. Therefore

$$||\phi_j|| = \max_{x \varepsilon D_x} |\phi_j(x)| > 0, \text{ for } j = 1,2,\ldots,n.$$

By the closure of $\{\psi_i\}$ in $C(D_y)$, given $\varepsilon > 0$, there exists an

$M_j = M_j(\varepsilon/2n||\phi_j||)$, such that $m > M_j$ implies that

$$\max_{y\ D_y} |a_j(y) - Q_{A_j}(y)| < \varepsilon/2n||\phi_j|| ,$$

where Q_{A_j} is the best uniform approximation in $\psi_1, \psi_2, \ldots, \psi_m$ to a_j on D_y.

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

Then $m > M$ implies that

$$\max_{(x,y)\varepsilon D} |P_{A(y)}(x) - PT_A(x,y)| = \max_{(x,y)\varepsilon D} \left| \sum_{j=1}^{n} a_j(y)\phi_j(x) - \sum_{j=1}^{n} Q_{A_j}(y)\phi_j(x) \right|$$

$$= \max_{(x,y)\varepsilon D} \left| \sum_{j=1}^{n} \left(a_j(y) - Q_{A_j}(y)\right) \phi_j(x) \right|$$

$$\leq \sum_{j=1}^{n} \max_{(x,y)\varepsilon D} \left|\left(a_j(y) - Q_{A_j}(y)\right) \phi_j(x)\right|$$

$$\leq \sum_{j=1}^{n} \max_{y\varepsilon D_y} |a_j(y) - Q_{A_j}(y)| \cdot \max_{x\varepsilon D_x} |\phi_j(x)|$$

$$< \sum_{j=1}^{n} (\varepsilon/2n||\phi_j||)\ ||\phi_j|| = \frac{\varepsilon}{2}\ .$$

Hence $n > N$ and $m > M$ imply that

$$||F - PT_A|| = \max_{(x,y)\varepsilon D} |F(x,y) - PT_A(x,y)|$$

$$\leq \max_{(x,y)\varepsilon D} |F(x,y) - P_{A(y)}(x)| + \max_{(x,y)\varepsilon D} |P_{A(y)}(x) - PT_A(x,y)|$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\ .$$

Corollary 2.7.6: Under the hypothesis of theorem 2.7.5 given $\varepsilon > 0$ and $F\varepsilon C(D)$ there exists an $N = N(\varepsilon)$ and an $M = M(\varepsilon,n)$ for each $n > N(\varepsilon)$, such that $n > N$ and $m > M$ imply that

$$||PT_A - P^*|| < \varepsilon$$

where PT_A is the Product Tchebycheff approximation and P^* is a best uniform approximation (in $\phi_1, \phi_2, \ldots, \phi_n, \psi_1, \psi_2, \ldots, \psi_m$) to F on D.

Proof: Let $N = N(\varepsilon/2)$ and $M = M(\varepsilon/2,n)$ correspond to the N and M in theorem 2.7.5. Then $n > N$ and $m > M$ imply that

$$||PT_A - P^*|| \leq ||F - PT_A|| + ||F - P^*||$$

$$\leq 2\ ||F - PT_A||$$

$$< 2\ \frac{\varepsilon}{2} = \varepsilon\ .$$

At the present time there is no known effective scheme for computing best uniform approximation to a function of two variables. Present research is being directed towards schemes somewhat like the Remez 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. The previous results (i.e. Theorem 2.7.5 + Corollary 2.7.6) suggest that in certain cases the Product Tchebycheff approximation may be a good initial guess to a best uniform approximation.

Section 8: Approximation on a Domain Which does not Possess Property K

The Product Tchebycheff approximation as defined in Chapter II, Section 4, applies only to sets which possess property K. We shall briefly discuss the problem of approximation of functions on certain sets which do not possess property K.

Let D be a set which does not possess property K. Suppose there is an invertible transformation T which maps E_2 into E_2, such that the set T(D) possesses property K. If $F \varepsilon C(D)$ then $FT^{-1} \varepsilon C(T(D))$. Suppose also that T(D), FT^{-1} and $\phi_1, \phi_2, \ldots, \phi_n, \psi_1, \psi_2, \ldots, \psi_m$ satisfy condition P.T.. Then FT^{-1} has a Product Tchebycheff approximation PT_A on T(D).

We can use PT_A to approximate F on D by defining

$$P(x,y) = PT_A \left[T[(x,y)]\right] .$$

Example 2.8.1: Let D be the shaded area in Figure 2.8.1.

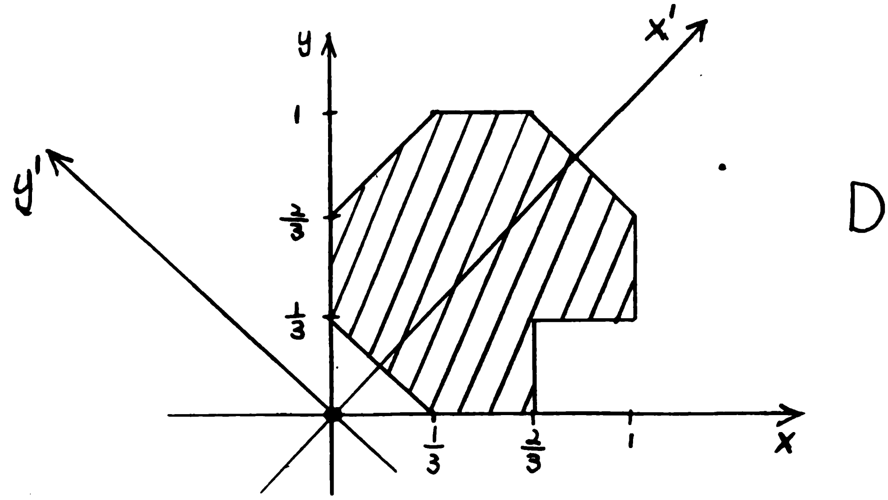


Figure 2.8.1

Then D does not possess property K.

Let T_1: $(x,y) \to (x^1,y^1)$ be the rotation of the x and y axes counterclockwise through an angle of 45°. Then $T_1(D)$ possesses property K.

$$x^1 = \frac{\sqrt{2}}{2}(x+y), \quad y^1 = \frac{\sqrt{2}}{2}(y-x) \ .$$

Therefore T_1^{-1}: $(x^1,y^1) \to (x,y)$ where

$$x = \frac{\sqrt{2}}{2}(x^1-y^1), \quad y = \frac{\sqrt{2}}{2}(x^1+y^1).$$

For this set D we shall consider the approximation of

$F(x,y) = x + y.$

$$\text{Then } FT_1^{-1}(x^1,y^1) = F\left(\frac{\sqrt{2}}{2}(x^1-y^1), \frac{\sqrt{2}}{2}(x^1+y^1)\right)$$

$$= \frac{\sqrt{2}}{2}(x^1-y^1) + \frac{\sqrt{2}}{2}(x^1+y^1)$$

$$= \sqrt{2}\, x^1 \qquad \text{on} \quad T_1(D).$$

Let $\phi_1(x^1) = 1$, $\phi_2(x^1) = x^1$

$\psi_1(y^1) = 1$.

Then we can easily see that the Product Tchebycheff approximation for FT^{-1} on $T_1(D)$ is

$$PT_{A_1}(x^1, y^1) = \sqrt{2}\, x^1 .$$

Therefore we approximate F on D by

$$P_1(x,y) = PT_{A_1}\left(T_1[(x,y)]\right)$$

$$= PT_{A_1}\left(\frac{\sqrt{2}}{2}(x+y), \frac{\sqrt{2}}{2}(y-x)\right)$$

$$= \sqrt{2}\left(\frac{\sqrt{2}}{2}(x+y)\right) = x + y.$$

Now let $T_2\colon (x,y) \to (x'', y'')$ be the rotation of the x and y axis clockwise through an angle of 45°. Then $T_2(D)$ possesses property K (see Figure 2.8.2).

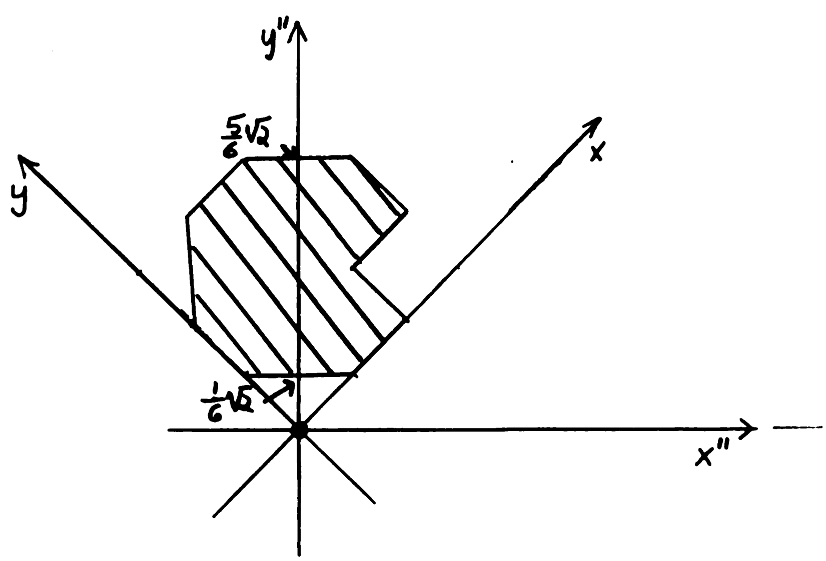


Figure 2.8.2

$$x'' = \frac{\sqrt{2}}{2}(x-y), \quad y'' = \frac{\sqrt{2}}{2}(x+y)$$

Therefore T_2^{-1}: $(x'',y'') \to (x,y)$ where

$$x = \frac{\sqrt{2}}{2}(x''+y''), \quad y = \frac{\sqrt{2}}{2}(y''-x'') \ .$$

Then $FT_2^{-1}(x'',y'') = \left(F\ \frac{\sqrt{2}}{2}(x''+y''),\ \frac{\sqrt{2}}{2}(y''-x'')\right)$

$$= \frac{\sqrt{2}}{2}(x''+y'') + \frac{\sqrt{2}}{2}(y''-x'')$$

$$= \sqrt{2}\, y'' \text{ on } T_2(D).$$

Let $\phi_1(x'') = 1$, $\phi_2(x'') = x''$

$\psi_1(y'') = 1$.

Then the best uniform approximation to

$FT^{-1}_{2y''}(x'') = \sqrt{2}\, y''$ on $T_2(D)_{x''}(y'')$ is

$P_{A(y'')}(x'') = \sqrt{2}\, y''$.

Therefore the Product Tchebycheff approximation to FT_2^{-1} on $T_2(D)$ is

$$PT_{A_2}(x'',y'') = \frac{1}{2}\left[\max_{y''\,\varepsilon\left(\frac{1}{6}\sqrt{2},\frac{5}{6}\sqrt{2}\right)} \sqrt{2}\, y'' + \min_{y''\varepsilon\left(\frac{1}{6}\sqrt{2},\frac{5}{6}\sqrt{2}\right)} \sqrt{2}\, y''\right]$$

$$= \frac{1}{2}\left[\frac{5}{3} + \frac{1}{3}\right] = 1$$

Therefore we can approximate F on D by

$$P_2(x,y) = PT_{A_2}(T_2[(x,y)])$$

$$= PT_{A_2}\left(\frac{\sqrt{2}}{2}(x-y),\ \frac{\sqrt{2}}{2}(x+y)\right)$$

$$= 1 \ .$$

This example illustrates that this technique may possess the undesirable property of lacking a unique solution. However we can specify the transformation used and we can bridge the communication gap.

There are a few other difficulties which we may encounter. It may prove difficult to find a transformation T such that FT^{-1} will have a Product Tchebycheff approximation on $T(D)$. Also, the resulting approximation $PT_A T$ is not in general a polynomial. However if T is a linear transformation, then $PT_A T$ is a polynomial and thus possesses the desirable properties of a polynomial approximation.

Section 9: The Product Tchebycheff Approximation to a Continuous Function of Three or More Variables

We now extend the Product Tchebycheff approximation to continuous functions of three or more variables. For ease of notation we shall restrict our attention to the three variable case. The further extension is straightforward.

Let D be a compact set in E_3. We define the following compact sets.

D_x, D_y, D_z, the projections of D onto the x,y and z axes respectively.

$D_z(x_1, y_1)$, the projection of the intersection of the set D and the line $x = x_1$, $y = y_1$ onto the z - axis.

$D_{x,y}$, the projection of D onto the x,y plane.

Definition 2.9.1: Let $F \varepsilon C(D)$. Then for each $(x_1,y_1) \varepsilon D_{x,y}$ we define the associated function F_{x_1,y_1} on $D_z(x_1,y_1)$ by

$$F_{x_1,y_1}(z) = F(x_1,y_1,z).$$

Let $\phi_1,\phi_2,\ldots,\phi_n$ be a Tchebycheff system of continuous functions on D_z.

Definition 2.9.2: For each $(x_1,y_1) \varepsilon D_{x,y}$ such that $D_z(x_1,y_1)$ contains n or more points, Haar's theorem shows the existence of a unique polynomial

$$P_{A(x_1,y_1)} = \sum_{j=1}^{n} a_j(x_1,y_1)\phi_j \quad ,$$

where $A(x_1,y_1) = \left| a_1(x_1,y_1), a_2(x_1,y_1), \ldots, a_n(x_1,y_1) \right| \varepsilon E_n$, which is the best uniform approximation to the continuous function F_{x_1,y_1} on its domain of definition $D_z(x_1,y_1)$.

We now extend the definition of property K to $D \subset E_3$, and we can show that with the addition of this property A(x,y) is continuous on $D_{x,y}$.

Definition 2.9.3: The compact set D in E_3 is said to possess property K, relative to the variables x,y if and only if given $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon) > 0$ such that (x_1,y_1), $(x_2,y_2) \varepsilon D_{x,y}$ and

$$\sigma_2[(x_1,y_1);\ (x_2,y_2)] < \delta$$

imply that for each $z_1 \varepsilon D_z(x_1,y_1)$ there is a corresponding $z_2 \varepsilon D_z(x_2,y_2)$ satisfying

$$|z_2 - z_1| < \varepsilon .$$

Definition 2.9.4: For each $x_1,y_1)$ in $D_{x,y}$ let

$$\rho(x_1,y_1) = \inf_{A \varepsilon E_n} \sup_{z \varepsilon D_z(x_1,y_1)} |F_{x_1,y_1}(z) - P_A(z)| .$$

The following two theorems are extensions of theorems 2.3.2 and 2.3.6. The proofs are omitted since they do not incorporate any new concepts.

Theorem 2.9.5: Let D be a compact set in E_3, which possesses property K, relative to x,y and let $\phi_1,\phi_2,\ldots,\phi_n$ be a Tchebycheff system of real-valued continuous functions on D_z. Let $F \varepsilon C(D)$. Then ρ is a continuous function on $D_{x,y}$.

Theorem 2.9.6: Let D, F and $\phi_1,\phi_2,\ldots,\phi_n$ satisfy the hypothesis of theorem 2.9.5 and let (x_1,y_1) contains n or more points. Then $A(x,y)$ is continuous at (x_1,y_1).

Corollary 2.9.7: Under the hypothesis of theorem 2.9.6 $a_1(x,y),a_2(x,y),\ldots,a_n(x,y)$ as defined in 2.9.1 are all continuous at $x_1,y_1)$.

Now as in Section 4 we are motivated to seek polynomial approximations to each of the continuous functions $a_1,a_2,\ldots,a_n$. In this case we will approximate each a_j, $j = 1,2,\ldots,n$ by a corresponding Product Tchebycheff approximation of degree two.

D

D

C

<u>Definition 2.9.8</u>: We will say that the compact set $D \subset E_3$ and $\theta_1,\theta_2,\ldots,\theta_k,\psi_1,\psi_2,\ldots,\psi_m,\phi_1,\phi_2,\ldots,\phi_n$ satisfy condition P.T. relative to x,y if and only if

(1) D possesses property K relative to x,y,

(2) $\phi_1,\phi_2,\ldots,\phi_n$ is a Tchebycheff system of continuous real-valued functions on D_z,

(3) For each $(x_1,y_1) \varepsilon D_{x,y}$, the set $D_z(x,y)$ contains n or more points,

(4) The compact set $D_{x,y} \subset E_2$ and $\theta_1,\theta_2,\ldots,\theta_k$, $\psi_1,\psi_2,\ldots,\psi_m$ satisfy condition P.T. relative to x.

<u>Definition 2.9.9</u>: Let the compact set $D \subset E_3$ and $\theta_1,\theta_2,\ldots,\theta_k$, $\psi_1,\psi_2,\ldots,\psi_m,\phi_1,\phi_2,\ldots,\phi_n$ satisfy condition P.T. relative to x,y. For each $j = 1,2,\ldots,n$ let Q_{A_j} be the Product Tchebycheff approximation to the continuous function a_j on $D_{x,y}$ relative to x, (with base functions $\theta_1,\theta_2,\ldots,\theta_k$ and $\psi_1,\psi_2,\ldots,\psi_m$).

Then the polynomial

$$PT_A = \sum_{j=1}^{n} Q_{A_j}\phi_j$$

is called the Product Tchebycheff approximation to F on D relative to x,y.

Note that this approximation depends on the order in which the variables x,y,z are specified. As was illustrated

in example 2.4.3 two distinct orders may produce two distinct Product Tchebycheff approximations.

We conclude this section with a simple example which illustrates the ideas we have presented.

Example 2.9.10: Let $D = [0,1] \times [0,1] \times [0,1]$, and let $F(x,y,z) = x + y + z$

$$\theta_1(x) = 1, \ \psi_1(y) = 1, \ \phi_1(z) = 1.$$

Then for each $(x_1,y_1) \varepsilon [0,1] \times [0,1]$, we approximate

$$F_{x_1,y_1}(z) = x_1 + y_1 + z$$

on $[0,1]$ by its best uniform approximation

$$\begin{aligned} P_{A(x_1,y_1)}(z) &= a_1(x_1,y_1)\phi_1(z) \\ &= a_1(x_1,y_1) \\ &= x_1 + y_1 + \tfrac{1}{2} \ . \end{aligned}$$

Next we approximate $a_1(x,y) = x + y + \frac{1}{2}$ on $[0,1] \times [0,1]$ by its Product Tchebycheff approximation relative to x,

$$\begin{aligned} Q_{A_1}(x,y) &= \tfrac{3}{2}\, \theta_1(x)\psi_1(y) \\ &= \tfrac{3}{2} \ . \end{aligned}$$

Therefore, the Product Tchebycheff approximation to F on D relative to x,y is

$$\begin{aligned} PT_A(x,y,z) &= Q_{A_1}(x,y)\phi_1(x) \\ &= \tfrac{3}{2} \ . \end{aligned}$$

Section 10: Conclusion

Let $w(x,y)$ be a positive continuous weight function on D. Then we define

$$\rho(y) = \inf_{A \varepsilon E_n} \ \operatorname{Sup}_{x \varepsilon D_x(y)} \left| w(x,y) \left(F_y(x) - P_A(x) \right) \right|$$

to be the deviation of a best weighted approximation to F_y on $D_x(y)$. We can show that extensions of theorems 2.3.2 and 2.3.6 hold for this more general problem. Therefore we can extend the Product Tchebycheff approximation to a weighted Product Tchebycheff approximation.

This thesis can also be extended to arbitrary norms and to the approximation of a function and its partial derivatives.

BIBLIOGRAPHY

1. N. I. Achieser, Theory of Approximation (English translation), New York, Frederick Unger, 1956.

2. R. C. Buck, Linear spaces and approximation theory, On Numerical Approximation, ed. Rudolph Langer, Univ. of Wis. Press, 1959, 11-23.

3. R. C. Buck, Survey of recent Russian literature on approximation, On Numerical Approximation, ed. Rudolph Langer, Univ. of Wis. Press, 1959, 341-359.

4. E. W. Cheney, Introduction to Approximation Theory, New York, McGraw-Hill, 1966.

5. P. J. Davis, Interpolation and Approximation, New York, Blaisdell, 1963.

6. W. Fraser and J. F. Hart, On the computation of rational approximations to continuous functions, Comm. ACM. 5 (1962), 401-403.

7. M. Golomb, Approximation by functions of fewer variables, On Numerical Approximation, ed. Rudolph Langer, Univ. of Wis. Press, 1959, 275-327.

8. A. Haar, Die Minkowshische Geometrie and die Annäheruug an stetige Funktionen, Math. Ann. 18 (1918), 294-311.

9. G. G. Lorentz, Approximation of Functions, New York, Holt, Rinehart and Winston, 1966.

10. John C. Mairhuber, On Haar's theorem concerning Chebychev approximation problems having unique solutions, Proc. Amer. Math. Soc. 7 (1956), 609-615.

11. David G. Moursund, Application of numerical maximization and minimization techniques to Chebyshev approximation, SIAM Rev. 9 (1967), 100-106.

12. I. P. Natanson, Konstruktive Funktionen theorie, Deutsche Übersetzung von K. Bogel, Berlin, Akademie Verlag, 1955.

13. E. P. Novodvorskii and I. S. Pinsker, The process of equating maxima, Uspehi Mat. Nauk. 6 (1951), 174-181.

14. R. R. Phelps, Uniqueness of Hahn-Banach extensions and unique best approximation, Trans. Amer. Math. Soc. 95 (1960), 238-255.

15. Anthony Ralston, A First Course in Numerical Analysis, New York, McGraw-Hill, 1965.

16. E. Ya. Remez, General Computational Methods of Chebyshev Approximation, United States Atomic Energy Commission, Washington, D.C., 1962.

17. E. Ya. Remez, Sur le calcul effectif des polynomes d'approximation de Tchebycheff, Compt. Rend. 199 (1934), 337-340.

18. J. R. Rice, The Approximation of Functions, Vol. 1, Reading, Mass., Addison-Wesley, 1964.

19. J. R. Rice, Tchebycheff approximation in several variables, Trans. Amer. Math. Soc. 109 (1963) 444-466.

20. 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.

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

22. T. J. Rivlin and H. S. Shapiro, A unified approach to certain problems of approximation and minimization, J. Soc. Indust. Appl. Math. 9 (1961), 670-699.

23. I. J. Schoenberg, On the question of unicity in the theory of best approximation, Ann. New York Acad. Sci. 86 (1960), 682-692.

24. Edward L. Stiefel, Numerical methods of Tchebycheff approximation, On Numerical Approximation, ed. Rudolph Langer, Univ. of Wis. Press, 1959, 217-232.

25. John Todd, Introduction to the Constructive Theory of Functions, California Institute of Technology, Pasadena, California, 1961.

26. L. Veidinger, On the numerical determination of the best approximations in the Chebyshev sense, Numer. Math. 2 (1960), 99-105.

27. Burton Wendroff, Theoretical Numerical Analysis, New York, Academic Press, 1966.