SOME CONTRIBUTIONS TO THE THEORY AND APPLICATION OF POLYNOMIAL APPROXIMATION

George McArtney Phillips

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

1969

SOME CONTRIBUTIONS TO THE THEORY

AND APPLICATION OF POLYNOMIAL APPROXIMATION

A thesis presented for the degree of Doctor of Philosophy

in the Faculty of Science of the University of St. Andrews

by

G.M. Phillips, M.A., M.Sc.

1969 St. Andrews

In memory of

my

father and mother

I declare that this thesis is of my own composition and that the work of which it is a record has been carried out by myself, except that part of the work was done jointly, as mentioned in the Acknowledgments. It has not been submitted in any previous application for a Higher Degree.

The thesis describes the results of research begun in the Department of Mathematics, University of Southampton, where I was admitted as a research student under the supervision of Professor H.B. Griffiths in October 1966, and continued in the Department of Applied Mathematics, University of St. Andrews under the supervision of Professor S.N. Curle since October 1967, the date of my admission as a research student at the University of St. Andrews.

I hereby certify that G.M. Phillips has fulfilled the conditions of Ordinance No. 12 and Resolution of the University Court No. 1 (St. Andrews) and is qualified to submit the accompanying thesis in application for the degree of Doctor of Philosophy.

Supervisor

ACKNOWLEDGMENTS

I am most indebted to the University of Aberdeen where I was undergraduate, M.Sc. student and Assistant, to the University of Southampton where as Assistant Lecturer and Lecturer I began my research in numerical analysis, and to the University of St. Andrews where as a Lecturer in the Department of Applied Mathematics, I am fortunate enough to be able to continue this work.

I am therefore most grateful to many persons. These include my former M.Sc. supervisor Principal E.M. Wright, Southampton colleagues Professor H.B. Griffiths, Dr. P.A. Samet and Mr. P.J. Taylor, and Professor S.N. Curle of St. Andrews. I also owe much to Professor P.J. Davis, Professor of Applied Mathematics at Brown University, Rhode Island, for the stimulus from his book 'Interpolation and Approximation'.

I would also like to take this rare opportunity of expressing my thanks to my earlier teachers, who include Miss M. Cassie and Mr. R. Gordon at Walker Road School, Aberdeen and Mr. J.A. Flett and Sir J.J. Robertson at Aberdeen Grammar School.

Most of the work for Chapter 4 was done at Southampton in collaboration with Mr. P.J. Taylor.

CONTENTS

Page

INTRODUCTION AND SUMMARY

The fundamental theorem, as far as this work is concerned, is Weierstrass' theorem (1885) on the approximability of continuous functions by polynomials. Since the time of Weierstrass (1815-97) and his equally important contemporary Chebyshev (1821-94), the topic of approximation has grown enormously into a subject of considerable interest to both pure and applied mathematicians.

The subject matter of this thesis, being exclusively concerned with polynomial approximations to a single-valued function of one real variable, is on the 'applied' side of approximation theory. The first chapter lists the definitions and theorems required subsequently. Chapter 2 is devoted to estimates for the maximum error in minimax polynomial approximations. Extensions of this are used to obtain crude error estimates for cubic spline approximations. The following chapter extends the minimax results to deal also with best L_p polynomial approximations, which include best least squares (L_2) and best modulus of integral (L_1) approximations as special cases. Chapter 4 is different in character. It is on the practical problem of approximating to convex or nearly convex data.

Chapter 1

PRELIMINARY DEFINITIONS AND THEOREMS

This chapter contains definitions and theorems which are required in subsequent chapters. The proofs of most of the theorems are readily available in texts and are not repeated here. Where a proof is omitted, a reference is given to a source of a proof.

1.1 Minimax approximations

Theorem 1. (Weierstrass' theorem). Given a function $f(x)$ continuous on $[a,b]$ and any $\epsilon > 0$, there exists a polynomial $q(x)$ such that

$$\max_{a \leq x \leq b} |f(x) - q(x)| < \epsilon$$

(Proof in Davis, 1963).

Definition 1. Given a function $f(x)$ defined on $[0,1]$, the n^{th} Bernstein polynomial for $f(x)$, denoted by $B_n(f;x)$ is defined as

$$B_n(f;x) = \sum_{j=0}^{n} \binom{n}{j} x^j (1-x)^{n-j} f(\tfrac{j}{n}) \tag{1.1}$$

Theorem 2. (Bernstein's theorem). If $f(x)$ is continuous on $[0,1]$ the sequence of polynomials $\left(B_n(f;x)\right)$ converges uniformly to $f(x)$ on $[0,1]$ as n tends to infinity. (Proof in Davis, 1963).

A linear change of variable extends this result to any finite interval $[a,b]$ and provides a constructive proof of Weierstrass' theorem.

Definition 2. An interpolating polynomial for a function $f(x)$ constructed at the distinct points $x = x_0, x_1, \ldots, x_n$ is a polynomial $q(x)$ of lowest degree such that $q(x_j) = f(x_j)$, $j = 0,1,\ldots,n$.

Theorem 3. The interpolating polynomial for a single-valued function $f(x)$ constructed at a distinct set of points $x_0, x_1, \ldots, x_n$ exists and is unique. (Proof in Davis, 1963).

Theorem 4. Let $a \leq x_j \leq b$ for $j = 0,1,\ldots,n$ and let $q(x)$ denote the interpolating polynomial for $f(x)$ constructed at $x_0, \ldots, x_n$. If $f^{(n+1)}(x)$ exists for $a \leq x \leq b$ and is continuous for $a < x < b$, then there exists a point on $[a,b]$, say ξ_x, such that for any x on $[a,b]$

$$f(x) - q(x) = \frac{1}{(n+1)!}(x-x_0)\cdots(x-x_n)\, f^{(n+1)}(\xi_x). \qquad (1.2)$$

(Proof in Davis, 1963). It should be noted that in (1.2) ξ_x

is itself a function of x.

P_n will be used to denote the set of all polynomials, with real coefficients, of degree not greater than n.

Definition 3. Given a function $f(x)$ defined on $[a,b]$, a polynomial $q^*(x) \in P_n$ is said to be a best minimax (or best Chebyshev) approximation to $f(x)$ on $[a,b]$ of degree not greater than n if

$$\inf_{q(x) \in P_n} \max_{a \leq x \leq b} |f(x) - q(x)| \tag{1.3}$$

is attained with $q(x) = q^*(x)$.

Theorem 5. If $f(x)$ is continuous on $[a,b]$ the infimum in (1.3) is attained. That is, the best minimax approximation exists and 'inf' in (1.3) may be replaced by 'min'. (Proof in Davis, 1963).

Theorem 6. The best minimax approximation defined by (1.3) is unique. (Proof in Davis, 1963).

Definition 4. A continuous function $e(x)$ is said to equioscillate at m points on $[a,b]$ if $\max_{a \leq x \leq b} |e(x)|$ is attained at m points $x_1, \ldots, x_m$ belonging to $[a,b]$ and also

$$\text{sign}\left[e(x_{j+1})\right] = -\,\text{sign}\left[e(x_j)\right]$$

for $j = 0,\ldots,m-1$. The x_j are called extreme points or extrema.

<u>Theorem 7.</u> If $f(x)$ is continuous on $[a,b]$ and $q^*(x)$ denotes the best minimax approximation defined by (1.3), then $f(x) - q^*(x)$ equioscillates at $n+2$ points on $[a,b]$. (Proof in Davis, 1963). This theorem is due to Chebyshev, as is also:

<u>Theorem 8.</u> If $f(x)$ is continuous on $[a,b]$ and for some $q(x) \in P_n$, $f(x) - q(x)$ equioscillates at $n+2$ points on $[a,b]$, then $q(x)$ is the best minimax approximation defined by (1.3). (Proof in Davis, 1963).

<u>Definition 5.</u> The modulus of continuity of a function $f(x)$ on $[a,b]$, denoted by

$$\omega(\delta) = \omega(f;\delta) = \omega(f, a, b; \delta)$$

is defined by

$$\omega(\delta) = \sup_{|x_1 - x_2| \leq \delta} |f(x_1) - f(x_2)|. \tag{1.4}$$

The supremum in (1.4) is over all x_1, x_2 belonging to $[a,b]$ and such that $|x_1 - x_2| \leq \delta$. It is clear that, if $f(x)$ is continuous on $[a,b]$, $\omega(\delta) \to 0$ as $\delta \to 0$.

1.2 Orthogonal polynomials

No attempt is made to give a systematic account of orthogonal polynomials here. Only the results required later are quoted.

Definition 6. The Chebyshev polynomial of degree n, denoted by $T_n(x)$, is defined as

$$T_n(x) = \cos n(\cos^{-1} x) \tag{1.5}$$

This is also referred to as the Chebyshev polynomial of the first kind.

Definition 7. Given a function $f(x)$, the Chebyshev series for $f(x)$, when it exists, is defined as

$$\sum_{j=0}^{\infty}{}' \; c_j \; T_j(x), \tag{1.6}$$

where

$$c_j = \frac{2}{\pi} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} f(x) \; T_j(x) \, dx , \tag{1.7}$$

for $j=0,1,\ldots$, and $\sum'$ denotes a sum whose first term is halved.

A Chebyshev series may be integrated to give another Chebyshev series. Suppose $f(x)$ is given by (1.6) and that

$$F(x) = \int_{-1}^{x} f(t)\,dt.$$

It follows that

$$F(x) = \sum_{j=0}^{\infty}{}' C_j T_j(x)$$

where

$$C_j = (c_{j-1} - c_{j+1})/2j, \quad j > 0, \tag{1.8}$$

and C_0 is determined by the lower limit of integration. See Goodwin et al., 1960.

Theorem 9. The infimum

$$\inf_{(c_j)} \max_{-1 \le x \le 1} |x^n + c_{n-1}x^{n-1} + \cdots + c_0| \tag{1.9}$$

is attained when $x^n + c_{n-1}x^{n-1} + \cdots + c_0 = \frac{1}{2^{n-1}} T_n(x)$.

(Proof in Davis, 1963).

<u>Definition 8.</u> The Chebyshev polynomial of the second kind of degree n, denoted by $U_n(x)$, is defined by

$$\mathcal{U}_n(x) = \sin\left((n+1)\cos^{-1}x\right) / \sin(\cos^{-1}x). \qquad (1.10)$$

From (1.5) and (1.10) it is easily checked that

$$T_n'(x) = n\,\mathcal{U}_{n-1}(x). \qquad (1.11)$$

<u>Theorem 10.</u> The infimum

$$\inf_{(c_i)} \int_{-1}^{1} |x^n + c_{n-1}x^{n-1} + \cdots + c_0|\,dx \qquad (1.12)$$

is attained when $x^n + c_{n-1}x^{n-1} + \cdots + c_0 = \frac{1}{2^n}\mathcal{U}_n(x)$

and has the value $1/2^{n-1}$. (Proof in Timan, 1963). This result is due to A.A. Markov.

<u>Definition 9.</u> A sequence of polynomials $q_0(x)$, $q_1(x)$, $q_2(x)$,..., where $q_j(x)$ has degree j, is said to be orthogonal on $[a,b]$ with respect to a function $w(x)$ if

$$\int_a^b w(x)\, q_j(x)\, q_k(x)\, dx = 0 \qquad (1.13)$$

for $j \neq k$ and is non-zero for $j=k$.

The function $w(x)$ is called the weight function.

Theorem 11. The sequence $T_0(x), T_1(x), \ldots$ is orthogonal with respect to $(1-x^2)^{-\frac{1}{2}}$ on $[-1,1]$. (Proof in Davis, 1963).

Theorem 12. The sequence $U_0(x), U_1(x), \ldots$ is orthogonal with respect to $(1-x^2)^{\frac{1}{2}}$ on $[-1,1]$. (Proof in Davis, 1963).

Definition 10. The Legendre polynomials are a sequence of polynomials $Q_0(x), Q_1(x), \ldots$ which are orthogonal with respect to the constant function 1 on $[-1,1]$ and which satisfy

$$Q_j(1) = 1, \qquad j=0,1,\ldots .$$

Theorem 13. The infimum

$$\inf_{(c_j)} \left[\int_{-1}^{1} (x^n + c_{n-1} x^{n-1} + \cdots + c_0)^2 \, dx \right]^{\frac{1}{2}} \tag{1.14}$$

is attained when

$$x^n + c_{n-1} x^{n-1} + \cdots + c_0 = 2^n Q_n(x) \Big/ \binom{2n}{n}$$

and has the value $\left(\frac{2}{2n+1}\right)^{\frac{1}{2}} \cdot \, 2^n \Big/ \binom{2n}{n}$.

(Proof in Davis, 1963).

Theorem 14. The minimum of

$$\int_{-1}^{1} w(x) \left[f(x) - \sum_{j=0}^{n} c_j \, q_j(x) \right]^2 dx ,$$

where $w(x)$ and the $q_j(x)$ satisfy Definition 9, is attained for

$$c_j = \int_{-1}^{1} w(x)\, f(x)\, q_j(x)\, dx \Big/ \int_{-1}^{1} w(x) \left[q_j(x) \right]^2 dx. \tag{1.15}$$

(Proof in Rice, 1964).

Theorem 15. The minimum of

$$\int_{-1}^{1} w(x)\,(x-x_1)^2 \cdots (x-x_n)^2\, dx,$$

over all choices of real numbers $x_1, \ldots, x_n$, is attained when $x_1, \ldots, x_n$ are the zeros of $q_n(x)$, which belongs to the set of polynomials orthogonal with respect to $w(x)$ on $[-1,1]$. (See Definition 9).

1.3 L_p approximations

Definition 11. Given a function $f(x)$ defined on $[a,b]$ and a number $p \geqslant 1$, a polynomial $q^*(x) \in P_n$ is said to be a best L_p polynomial approximation to $f(x)$ on $[a,b]$ of degree not greater than n if

$$\inf_{q(x)\in P_n} \left[\int_a^b |f(x) - q(x)|^p dx \right]^{\frac{1}{p}} \tag{1.16}$$

is attained when $q(x) = q^*(x)$. (If $p \geq 1$, for any function $g(x)$ defined on $[a,b]$, $\left[\int_a^b |g(x)|^p dx \right]^{\frac{1}{p}}$ defines a norm on the linear space of continuous functions defined on $[a,b]$. If $p<1$, one of the norm axioms is violated and so L_p approximations are usually restricted to a choice of $p \geq 1$, especially common choices being $p=1,2$ and ∞. The use of normed linear spaces facilitates the discussion of more general modes of approximation than are required here.)

It may be noted that

$$\lim_{p\to\infty} \left[\int_a^b |f(x) - q(x)|^p dx \right]^{\frac{1}{p}} = \sup_{a\leq x\leq b} |f(x) - q(x)|. \tag{1.17}$$

For this reason the best minimax approximation is sometimes called the best L_∞ approximation.

Theorem 16. If $f(x)$ is continuous on $[a,b]$ the infimum in (1.16) is attained for each $p \geq 1$. That is, the best L_p approximation $(p \geq 1)$ exists and 'inf' in (1.16) may be replaced by 'min'. (Proof in Davis, 1963).

Theorem 17. If $f(x)$ is continuous on $[a,b]$ the best L_p approximation defined by (1.16) is unique. (Proof in Davis, 1963).

Theorem 18. If $f(x)$ is continuous on $[a,b]$, then for any $p \geq 1$ a necessary and sufficient condition for $q(x) \in P_n$ to be the best L_p approximation (defined by (1.16)) is that

$$\int_a^b r(x) \, |f(x) - q(x)|^{p-1} \cdot \operatorname{sign}[f(x) - q(x)] \, dx = 0$$

for all $r(x) \in P_n$. (The proof will be given in Chapter 3).

We shall also require the following two results concerning inequalities, which are proved in Hardy, Littlewood and Pólya, 1934.

Theorem 19. If $0 < p < p'$, then, when the following integrals exist,

$$\left[\frac{1}{b-a}\int_a^b |g(x)|^p dx\right]^{\frac{1}{p}} < \left[\frac{1}{b-a}\int_a^b |g(x)|^{p'} dx\right]^{\frac{1}{p'}}$$

unless $g(x)$ is a constant function.

Theorem 20. (Hölder's inequality for integrals). If $p > 1$ then, if the following integrals exist,

$$\int_a^b |g(x)\,h(x)|\,dx \leq \left[\int_a^b |g(x)|^p dx\right]^{\frac{1}{p}} \cdot \left[\int_a^b |h(x)|^{p'} dx\right]^{\frac{1}{p'}}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$.

1.4 Generalisation of polynomial approximations

In order to deal subsequently with polynomial approximations which also interpolate $f(x)$ at certain points, we shall require a generalisation of some of the foregoing results on minimax approximation.

Definition 12. A finite set of functions $\psi_0(x), \psi_1(x), \cdots, \psi_n(x)$ is said to be a Chebyshev set on $[a,b]$ if the $\psi_j(x)$ are continuous and linearly independent on $[a,b]$ and the function $\sum_{j=0}^{n} c_j \psi_j(x)$ has at most n zeros on $[a,b]$ for any choice of real c_j, $j=0,1,\ldots,n$. (Such a set of functions is described by some authors, e.g. Timan, as satisfying the Haar property).

Theorem 21. If $f(x)$ is continuous on $[a,b]$ and the set of functions $\psi_0(x), \psi_1(x), \cdots, \psi_n(x)$ is a Chebyshev set on $[a,b]$, then

$$\inf_{(c_j)} \max_{a \leq x \leq b} \left| f(x) - \sum_{j=0}^{n} c_j \psi_j(x) \right| \tag{1.18}$$

is attained and the best approximation is unique. (Proof in Rice, 1964).

Theorem 22. If $f(x)$ is continuous on $[a,b]$ and $\psi_0(x), \psi_1(x), \cdots, \psi_n(x)$ is a Chebyshev set on $[a,b]$, then a necessary and sufficient condition for $\sum_{j=0}^{n} c_j \psi_j(x)$ to be the best approximation (as in (1.18)) is that

$$f(x) \quad - \quad \sum_{j=0}^{n} c_j \ \psi_j(x)$$

equioscillates at $n+2$ points on $[a,b]$. (Proof in Rice, 1964).

1.5 Cubic spline approximations

Definition 13. Given a function $f(x)$ defined on $[a,b]$ and a partition Δ : $a=x_0<x_1<\ldots<x_k=b$ of the interval $[a,b]$, a function $S_\Delta(f;x)$ is said to be a cubic spline approximation to $f(x)$ on Δ if

(i) $S_\Delta(f;x)$ is a cubic polynomial (at most) on on each interval $[x_j,x_{j+1}]$, $j=0,\ldots,k-1$,

(ii) $S_\Delta(f;x_j) = f(x_j)$, $j=0,\ldots,k$,

(iii) $S'_\Delta(f;x)$ and $S''_\Delta(f;x)$ are continuous on $[a,b]$.

Two further conditions are required, usually taken to be the values of $S'_\Delta(f;x)$ or $S''_\Delta(f;x)$ at the end-points $x=a$ and $x=b$, in order to specify a particular $S_\Delta(f;x)$ satisfying the three properties above. See, for example, Ahlberg, Nilson and Walsh, 1967.

Chapter 2

ESTIMATES OF THE MINIMAX ERROR

2.1 Minimax approximations over a single interval

Let us use

$$E_n(f) \quad = \quad E_n(f, a, b)$$

to denote

$$\inf_{(c_j)} \max_{a \leq x \leq b} \left| f(x) - \sum_{j=0}^{n} c_j x^j \right| .$$

In 1911, D. Jackson proved:

Theorem 23. If $f(x)$ is continuous on $[a,b]$, there exists a constant C such that

$$E_n(f) = E_n(f, a, b) \leqslant C \cdot \omega\left(f; \frac{b-a}{n}\right). \quad (2.1)$$

(Proof in Timan, 1963). Since by continuity of $f(x)$,

$$\omega(f; \delta) \to 0 \quad \text{as} \quad \delta \to 0,$$

Jackson's inequality (2.1) implies Weierstrass' theorem.

Jackson further proved;

<u>Theorem 24.</u> If $f(x)$ has its k^{th} derivative continuous on $[a,b]$, then for $n > k$

$$E_n(f) \leq M_k \cdot \left(\frac{b-a}{n}\right)^k \cdot \omega\left(f^{(k)}; \frac{b-a}{n}\right), \qquad (2.2)$$

where M_k is a constant depending only on k. (Proof in Timan, 1963).

Further results of this type, involving the modulus of continuity, are quoted in Timan, 1963. A more recent result, given by Meinardus, 1967, is;

<u>Theorem 25.</u> If $f^{(n+1)}(x)$ is continuous on $[a,b]$, there exists a number ξ, $a < \xi < b$, such that

$$E_n(f) = \frac{2}{(n+1)!} \left(\frac{b-a}{4}\right)^{n+1} |f^{(n+1)}(\xi)|. \qquad (2.3)$$

Meinardus' proof is based on a theorem due to Bernstein;

<u>Theorem 26.</u> Let $g(x)$ and $f(x)$ have derivatives of order $n+1$ on $[-1,1]$ and suppose that

$$|f^{(n+1)}(x)| \leq g^{(n+1)}(x), \qquad x \in [-1,1].$$

Then

$$E_n(f) \leq E_n(g).$$

(Proof in Meinardus, 1967).

An alternative proof of Theorem 25 will now be given which depends simply on the theorem concerning the error in the interpolating polynomial and Chebyshev's equioscillation theorem (Theorems 4 and 7). A similar approach enables one to estimate the error in best L_p polynomial approximations, to be dealt with in Chapter 3.

Proof of Theorem 25. If $q^*(x)$, of degree at most n, denotes the best minimax approximation for $f(x)$ on $[a,b]$, then by Theorem 7 there exist $n+2$ points on $[a,b]$ at which $f(x) - q^*(x)$ equioscillates. By continuity there are therefore $n+1$ distinct points, say $x_0^*, x_1^*, \ldots, x_n^*$ on $[a,b]$ where $f(x) - q^*(x) = 0$. That is, $q^*(x)$ may be regarded as the interpolating polynomial for $f(x)$ constructed at $x_0^*, x_1^*, \ldots, x_n^*$. So by Theorem 4 we may write

$$f(x) - q^*(x) = \frac{1}{(n+1)!} (x-x_0^*) \cdots (x-x_n^*) f^{(n+1)}(\xi_x^*), \tag{2.4}$$

where ξ_x^* is some function of x.

Now let $x_0, x_1, \ldots, x_n$ be the zeros of $T_{n+1}\left((2x-b-a)/(b-a)\right)$ and let $q(x)$ denote the interpolating polynomial for $f(x)$ constructed at $x_0, x_1, \ldots, x_n$. Then we also have

$$f(x) - q(x) = \frac{1}{(n+1)!} (x-x_0) \cdots (x-x_n) f^{(n+1)}(\xi), \tag{2.5}$$

where ξ_x is some function of x. Since

$$E_n(f) \leq \max_{a \leq x \leq b} |f(x) - q(x)|$$

it follows from (2.5) that

$$E_n(f) \leq \frac{1}{(n+1)!} \max_{a \leq x \leq b} |(x-x_0) \cdots (x-x_n)| \cdot \max_{a \leq x \leq b} |f^{(n+1)}(x)|. \tag{2.6}$$

Put $y = (2x-b-a)/(b-a)$ and for $j = 0, 1, \ldots, n$ let $y_j = (2x_j - b - a)/(b-a)$. Then

$$|(x-x_0)\cdots(x-x_n)| = \left(\frac{b-a}{2}\right)^{n+1} \cdot |(y-y_0)\cdots(y-y_n)|$$

and since the y_j are the zeros of $T_{n+1}(x)$,

$$\max_{a\le x\le b} |(x-x_0)\cdots(x-x_n)| = \left(\frac{b-a}{2}\right)^{n+1} \cdot \frac{1}{2^n} . \tag{2.7}$$

Using (2.7) in (2.6) we have the upper bound for $E_n(f)$:

$$E_n(f) \le \frac{2}{(n+1)!} \left(\frac{b-a}{4}\right)^{n+1} \cdot \max_{a\le x\le b} |f^{(n+1)}(x)| . \tag{2.8}$$

For a lower bound, we have from (2.4)

$$E_n(f) \ge \frac{1}{(n+1)!} \max_{a\le x\le b} |(x-x_0^*)\cdots(x-x_n^*)| \cdot \min_{a\le x\le b} |f^{(n+1)}(x)| . \tag{2.9}$$

From Theorem 9, concerning the minimax property of the Chebyshev polynomials, it follows that

$$E_n(f) \ge \frac{2}{(n+1)!} \left(\frac{b-a}{4}\right)^{n+1} \cdot \min_{a\le x\le b} |f^{(n+1)}(x)| . \tag{2.10}$$

Theorem 25 follows from the two inequalities (2.8) and (2.10), from the continuity of $f^{(n+1)}(x)$.

A connection with Chebyshev series.

Let us use c_j to denote the coefficient of $T_j(x)$ in the Chebyshev series for a function $f(x)$. Elliott, 1963, proved:

Theorem 27. If $f^{(n+1)}(x)$ is continuous on $[-1,1]$, then

$$c_{n+1} = \frac{1}{2^n} \cdot \frac{1}{(n+1)!} \cdot f^{(n+1)}(\eta), \qquad (2.11)$$

where $-1 \leq \eta \leq 1$.

It is well known that the truncated Chebyshev series is often a very close approximation to the best minimax polynomial of the same degree. For if the coefficients c_j tend to zero rapidly and the Chebyshev series

$$\sum_{j=0}^{\infty}{}' c_j T_j(x)$$

converges uniformly to $f(x)$ on $[-1,1]$,

$$f(x) - \sum_{j=0}^{n}{}' c_j T_j(x)$$

will be approximated closely by $c_{n+1}T_{n+1}(x)$, which equi-oscillates at the $n+2$ extrema of $T_{n+1}(x)$. Hence the similarity of (2.11) and (2.3) with $a=-1$, $b=1$ is not surprising, in view of Theorem 8.

2.2 Best approximations satisfying interpolatory conditions

In this section we investigate

$$\inf_{q(x)\in S_{n+2}} \max_{a\le x\le b} |f(x) - q(x)|, \tag{2.12}$$

where S_{n+2} is used to denote the set of all real polynomials $q(x)$ of degree at most $n+2$ which also satisfy the end-point conditions

$$q(a) = f(a), \qquad q(b) = f(b).$$

Thus in (2.12) we are concerned with finding a best polynomial approximation which interpolates $f(x)$ at the end points of $[a,b]$.

Let us write $L(x)$ for the interpolating polynomial for $f(x)$ constructed at $x = a$ and $x = b$. That is,

$$L(x) = [(x-a)f(b) - (x-b)f(a)]/(b-a). \tag{2.13}$$

Given any $q(x) \in S_{n+2}$, since $q(x) - L(x)$ must vanish at $x=a$ and $x=b$, we have

$$q(x) = L(x) + (x-a)(x-b)\, r(x), \qquad (2.14)$$

say, where $r(x) \in P_n$. We have that

$$(x-a)(x-b)\, r(x) = \sum_{j=0}^{n} c_j\, x^j (x-a)(x-b)$$

and the set of functions

$$x^j (x-a)(x-b), \qquad j = 0, \cdots, n$$

form a Chebyshev set on any interval $[a+\epsilon, b-\epsilon]$, for $0 < \epsilon < b-a$. Regarding $(x-a)(x-b)\,r(x)$ as an approximation for $f(x) - L(x)$, the equioscillation theorem (Theorem 22) applies on any of the above intervals $[a+\epsilon, b-\epsilon]$. In particular, there exist $n+1$ points ξ_j on $[a+\epsilon, b-\epsilon]$ such that

$$f(\xi_j) - L(\xi_j) = (\xi_j - a)(\xi_j - b)\, r(\xi_j). \qquad (2.15)$$

That is, the choice of $r(x)$ corresponding to the best approximation above is an interpolating polynomial for

$$[f(x) - L(x)] \,/\, (x-a)(x-b)$$

constructed at certain points $\xi_0, \xi_1, \dots, \xi_n$. Let us write

$$F(x) = [f(x) - L(x)] \,/\, (x-a)(x-b), \tag{2.16}$$

which is defined on the open interval (a,b). Then if $f(x)$ is $(n+1)$ times differentiable we have from Theorem 4 that

$$F(x) - r(x) = \frac{1}{(n+1)!}(x-\xi_0)\cdots(x-\xi_n)\, F^{(n+1)}(\eta_x). \tag{2.17}$$

Using (2.14) and (2.16) it follows from this that

$$f(x) - q(x) = \frac{1}{(n+1)!}(x-a)(x-b)(x-\xi_0)\cdots(x-\xi_n)\, F^{(n+1)}(\eta_x). \tag{2.18}$$

In (2.18), $\xi_0, \dots, \xi_n$ depend on the choice of ϵ, as do the functions $q(x)$ and η_x. As $\epsilon \to 0$, each ξ_j will have a limit, say ξ_j^*, and we will have

$$f(x) - q^*(x) = \frac{1}{(n+1)!}(x-a)(x-b)(x-\xi_0^*)\cdots(x-\xi_n^*)\, F^{(n+1)}(\eta_x^*), \quad (2.19)$$

where $q^*(x)$ denotes the polynomial for which the infimum in (2.12) is attained. Let us put

$$\mu_n = \inf_{(\xi_j)} \max_{-1\le x\le 1} |(1-x^2)(x-\xi_0)\cdots(x-\xi_n)|. \quad (2.20)$$

By considering (2.20) as the problem of approximating to $x^{n+1}(1-x^2)$ by a linear combination of the functions $x^j(1-x^2)$, $j=0,1,\ldots,n$, we can see that there exists a unique set of points ξ_j at which the infimum is attained. Let $q(x)$ denote the polynomial whose associated $r(x)$ interpolates $F(x)$ at the points on $[a,b]$ corresponding to (by a linear transformation) the minimising ξ_j on $[-1,1]$. Then in a similar way to the last section we have

$$\inf_{q(x)\in S_{n+2}} \max_{a\le x\le b} |f(x) - q(x)| \quad (2.21)$$

$$\geq \frac{\mu_n}{(n+1)!}\left(\frac{b-a}{2}\right)^{n+3} \cdot \min_{a\le x\le b} |F^{(n+1)}(x)|$$

and also

$$\inf_{q(x)\in S_{n+2}} \max_{a\le x\le b} |f(x) - q(x)| \le \frac{\mu_n}{(n+1)!}\left(\frac{b-a}{2}\right)^{n+3} \cdot \max_{a\le x\le b} |F^{(n+1)}(x)|. \tag{2.22}$$

Thus, combining these inequalities, we have the following result :

<u>Theorem 28.</u> Given a function $f(x)$ whose $(n+1)^{\text{th}}$ derivative is continuous on $[a,b]$, there exists a number $\xi\in[a,b]$ such that

$$\inf_{q(x)\in S_{n+2}} \max_{a\le x\le b} |f(x) - q(x)| = \frac{\mu_n}{(n+1)!}\left(\frac{b-a}{2}\right)^{n+3} \cdot |F^{(n+1)}(\xi)|. \tag{2.23}$$

Note that the auxiliary function $F(x)$, which appears on the right side of (2.23), depends on $f(x)$ and on $[a,b]$, as given by (2.16) and (2,13).

Bounds for μ_n.

From (2.20) we have

$$\mu_n \leq \frac{1}{2^n} \max_{-1 \leq x \leq 1} \left|(1-x^2)\, T_{n+1}(x)\right| \qquad (2.24)$$

$$\leq \frac{1}{2^n} .$$

For a lower bound, we have

$$\mu_n > \inf_{(x_j)} \max_{-1 \leq x \leq 1} \left|(x-x_0)(x-x_1)\cdots(x-x_{n+2})\right| \qquad (2.25)$$

$$= \frac{1}{2^{n+2}} .$$

Combining these inequalitims, we obtain

$$\frac{1}{2^{n+2}} < \mu_n \leq \frac{1}{2^n} . \qquad (2.26)$$

Precise values for μ_0 and μ_1.

It follows from the uniqueness and equioscillation property associated with the minimising ξ_j for (2.20) that the ξ_j must be symmetrically placed about the origin. For, on replacing x by $-x$, the polynomial

$$(1-x^2)(x+\xi_0)\cdots(x+\xi_n)$$

will also equioscillate and must therefore be identical (by uniqueness) with the polynomial

$$(1-x^2)(x-\xi_0)\cdots(x-\xi_n).$$

Thus, in particular, we have

$$\mu_0 = \max_{-1\leq x\leq 1} |x(1-x^2)| = \frac{2\sqrt{3}}{9}. \tag{2.27}$$

Also,

$$\mu_1 = \inf_{(\xi)} \max_{-1\leq x\leq 1} |(1-x^2)(x^2-\xi^2)|.$$

Hence we find that

$$\mu_1 = \max_{-1\leq x\leq 1} |(1-x^2)(x^2-\tfrac{1}{5})| = \frac{1}{5}. \tag{2.28}$$

These values for μ_0 and μ_1 are consistent with the inequalities (2.26).

<u>Example.</u> Consider the function $f(x) = (\alpha + x)^{-1}$ on $[-1,1]$ with $\alpha > 1$. In this case (2.13) gives

$$L(x) = \frac{1}{2}\left[\frac{x+1}{\alpha+1} - \frac{x-1}{\alpha-1}\right].$$

From (2.16),

$$F(x) = \left[f(x) - L(x)\right]/(x^2-1)$$

$$= \left[(\alpha^2-1)(\alpha+x)\right]^{-1}.$$

Thus, we have

$$|F^{(n+1)}(x)| = (n+1)!\ (\alpha^2-1)^{-1}(\alpha+x)^{-(n+2)}.$$

Therefore, for some $\xi \in [-1,1]$,

$$\inf_{q(x)\in S_{n+2}} \max_{-1\le x\le 1} |(\alpha+x)^{-1} - q(x)|$$
$$= \mu_n (\alpha^2-1)^{-1}(\alpha+\xi)^{-(n+2)}.$$

2.3 Piecewise approximations

Let us now approximate to $f(x)$ by partitioning $[a,b]$ into k sub-intervals and using a polynomial approximation of degree at most n on each sub-interval. Let us choose the points of sub-division and the k approximating polynomials so as to minimise the maximum error. It is clear that the maximum error, which will be denoted by $E_{n,k}(f)$, will be attained at least once on each sub-interval.

Let us write

$$E_{n,k}(f) \quad = \quad E_{n,k}(f, a, b)$$

to emphasise the dependence of $E_{n,k}(f)$ on the interval $[a,b]$. We already know that a best approximation of this type exists for $k = 1$. We can see by induction on k that the best approximation described above exists for $k = 1,2,3,\ldots$. For, assuming that a best approximation exists when we have $k-1$ sub-intervals $(k \geq 2)$, we can find the best approximation on k sub-intervals by choosing a number δ, $0 < \delta < b-a$, such that

$$E_{n,k-1}(f, a, b-\delta) = E_{n,1}(f, b-\delta, b). \qquad (2.29)$$

Note that the left side of (2.29) is a decreasing function of δ and the right side is an increasing function.

Let $I_1, I_2, \ldots, I_k$ be the sub-intervals of $[a,b]$ corresponding to a best piecewise approximation and let $x_1, x_2, \ldots, x_{k-1}$ be the points of sub-division. Then from (2.3), assuming continuity of $f^{(n+1)}(x)$, we may write, for $j=1,2,\ldots,k$,

$$E_{n,k}(f) = \frac{2}{(n+1)!}\left(\frac{x_j - x_{j-1}}{4}\right)^{n+1} \cdot \left|f^{(n+1)}(\xi_j)\right|, \qquad (2.30)$$

where $\xi_j \in I_j$ and $x_0=a$, $x_k=b$. Thus

$$\left[\tfrac{1}{2}(n+1)!\; E_{n,k}(f)\right]^{\frac{1}{n+1}} = \tfrac{1}{4}(x_j - x_{j-1}) \cdot \left|f^{(n+1)}(\xi_j)\right|^{\frac{1}{n+1}} \qquad (2.31)$$

and, on summing (2.31) for $j=1,2,\ldots,k$, we obtain

$$k\left[\tfrac{1}{2}(n+1)!\; E_{n,k}(f)\right]^{\frac{1}{n+1}} = \tfrac{1}{4}\sum_{j=1}^{k}(x_j - x_{j-1}) \cdot \left|f^{(n+1)}(\xi_j)\right|^{\frac{1}{n+1}}. \qquad (2.32)$$

Assuming that $f(x)$ cannot be represented exactly by a polynomial of degree n on any sub-interval of $[a,b]$, we have that as $k \to \infty$ in (2.32) the length of the largest sub-interval $x_j - x_{j-1}$ will tend to zero. Therefore, as $k \to \infty$, we may replace the right side of (2.32) by the Riemann integral, giving

$$\lim_{k \to \infty} k^{n+1} E_{n,k}(f) = \frac{2}{(n+1)!} \left[\frac{1}{4} \int_a^b \left| f^{(n+1)}(x) \right|^{\frac{1}{n+1}} dx \right]^{n+1}. \tag{2.33}$$

The special case of (2.33), with $n=1$, is given by Ream, 1961. This will be of interest later in this chapter.

Returning to (2.30), at least one sub-interval I_j must have length not greater than $(b-a)/k$, so that

$$E_{n,k}(f) \leqslant \frac{2}{(n+1)!} \left(\frac{b-a}{4k} \right)^{n+1} \max_{a \leq x \leq b} \left| f^{(n+1)}(x) \right|. \tag{2.34}$$

This generalises the inequality (2.8) for $E_n(f)$. Similarly, in order to obtain a lower bound for $E_{n,k}(f)$, we can argue that at least one sub-interval I_j must have length not smaller than $(b-a)/k$. Hence we obtain

$$E_{n,k}(f) \geq \frac{2}{(n+1)!}\left(\frac{b-a}{4k}\right)^{n+1} \min_{a \leq x \leq b} \left|f^{(n+1)}(x)\right|. \qquad (2.35)$$

From these two inequalities we now have:

Theorem 29. If $f^{(n+1)}(x)$ is continuous on $[a,b]$, there exists a number $\xi \in [a,b]$ such that the error in the best piecewise polynomial approximation of degree at most n on each of k sub-intervals is

$$E_{n,k}(f) = \frac{2}{(n+1)!}\left(\frac{b-a}{4k}\right)^{n+1} \left|f^{(n+1)}(\xi)\right|. \qquad (2.36)$$

2.4 Algorithms for deriving piecewise straight line approximations

Stone, 1961, gives an algorithm for finding best least squares approximations to a function $f(x)$ on a finite interval $[a,b]$ by k straight line segments. He justifies the usefulness of his algorithm by showing how it may be applied in the solution of certain non-linear programming problems. Ream, 1961, refers to the relevance of this approximation problem in designing diode function-generators for analogue computers. In the examples given by Ream and Stone, $f''(x)$ is of constant sign. Nearly all functions of practical interest satisfy this condition at least piecewise.

In this section, algorithms will be described for solving the same problem, but finding minimax rather than least squares approximations.

Suppose $f''(x) > 0$ on $[\alpha, \beta]$ and that $cx + d$ is the best minimax straight line approximation for $f(x)$ on $[\alpha, \beta]$. We have from the equioscillation theorem (Theorem 7) that

$$\max_{\alpha \leq x \leq \beta} | f(x) - cx - d |$$

is attained on at least three points. At an interior extreme point, we will have

$$\frac{d}{dx} \left(f(x) - cx - d \right) = 0 .$$

That is,

$$f'(x) - c = 0 . \tag{2.37}$$

Since $f''(x) > 0$, (2.37) can have at most one solution on $[\alpha, \beta]$, whence it follows that two of the extreme points must occur at the end points α and β. The third extreme point will be an interior point, say ξ. If

$$\epsilon = \max_{\alpha \leq x \leq \beta} |f(x) - cx - d|, \tag{2.38}$$

we will have the following equations

$$f(\alpha) - (c\alpha + d) = \epsilon \tag{2.39}$$

$$f(\xi) - (c\xi + d) = -\epsilon \tag{2.40}$$

$$f(\beta) - (c\beta + d) = \epsilon \tag{2.41}$$

$$f'(\xi) - c = 0 \tag{2.42}$$

Given α and β, these four equations may be solved to determine c, d, ξ and ϵ. For we may eliminate d and ϵ from (2.39) and (2.41) to obtain

$$c = [f(\beta) - f(\alpha)]/(\beta-\alpha),$$

the slope of the chord joining the end points. Hence, using some root-finding procedure, most suitably one which 'brackets' the root, such as the rule of false position (regula falsi), ξ may be determined from (2.42). Lastly, d and ϵ are found by solving the two linear equations (2.39) and (2.40).

However, we will be more interested here in using the four equations (2.39) - (2.42) in a different way, as in the following theorem.

Theorem 30. Given α and ϵ the equations (2.39) - (2.42) where $f''(x) > 0$, have at most one solution for c, d, ξ and β.

Proof. From equations (2.39), (2.40) and (2.42) we have, on eliminating c and d,

$$f'(\xi)(\xi-\alpha) + f(\alpha) - f(\xi) - 2\epsilon = 0. \qquad (2.43)$$

Let us write this last equation, in which the only unknown is ξ, as

$$G(\xi) = 0.$$

Then we can see that

$$G'(\xi) = f''(\xi)(\xi - \alpha). \tag{2.44}$$

From this, it is seen that

$$G'(\xi) > 0, \qquad \text{for } \xi > \alpha.$$

Thus the equation (2.43) has at most one solution $\xi > \alpha$. Since from (2.43) $G(\alpha) < 0$, a solution ξ of (2.43) will exist on $[\alpha, b]$ if and only if $G(b) \geqslant 0$. If a solution does exist, we may find c from (2.42) and d from (2.39).

Equation (2.41) is then available to determine β. If we write this equation as

$$H(\beta) = 0,$$

then for $\beta > \xi$, we have

$$H'(\beta) = f'(\beta) - f'(\xi) > 0.$$

So there is at most one solution for β and, since $H(\xi) < 0$, a solution will exist on $[\xi, b]$ if and only $H(b) \geq 0$. This concludes the proof of Theorem 30.

The process of beginning with a pre-assigned minimax error $\in$ and a given value for the left hand end point α and then finding the minimax straight line and the right hand end point β will be used repeatedly in the following algorithm.

Algorithm 1. Given any $\in > 0$, we can construct k sub-intervals

$$[a, x_1], \quad [x_1, x_2], \quad \cdots, \quad [x_{k-1}, b]$$

and k straight lines

$$c_j x + d_j \quad , \qquad j = 1, 2, \cdots, k$$

such that on each sub-interval the largest error in approximating to $f(x)$ by the associated straight line is ϵ. It is assumed that $f''(x) > 0$ on the given interval $[a,b]$.

With the notation used in the proof of Theorem 30, if $G(b) \geqslant 0$ and $H(b) \geqslant 0$, then given the end point a (corresponding to α in the four equations (2.39) - (2.42)) and ϵ we can find the best minimax straight line, say c_1x+d_1, and also ξ_1 and x_1 (these last two numbers corresponding respectively to ξ and β above). The solution of the equations $G(x)=0$ and $H(x)=0$ may be found by the regula falsi method.

Beginning with x_1 (corresponding to α) and ϵ, a second minimax straight line may be constructed up to some point x_2, and so on. At some stage, say with x_{k-1} as the new left hand end point (α), we will find that either $G(b) < 0$ or $H(b) < 0$. The geometrical interpretation of this is that the k^{th} straight line with maximum error ϵ overshoots the right hand end point b.

When this stage is reached, we choose as c_kx+d_k

the straight line which passes through the points

$$(x_{k-1}, f(x_{k-1}) - \epsilon) \quad \text{and} \quad (b, f(b) - \epsilon).$$

Thus, given any $\epsilon > 0$, the algorithm obtains a piecewise straight line approximation to $f(x)$ on $[a,b]$ with maximum error ϵ. It may be noted that the approximating function is continuous over the whole interval $[a,b]$. The last straight line, $c_k x + d_k$ was chosen so as to preserve the continuity of the piecewise polynomial approximation. We also note that the approximation is achieved with the smallest possible number of straight line segments.

Best approximations by k segments.

In the above algorithm, given a pre-assigned maximum error ϵ, we obtained a piecewise straight line approximation for $f(x)$ on $[a,b]$. Now, suppose that we wish to approximate to $f(x)$ piecewise by means of precisely k straight line segments. That is, this time we are given the value of k at the outset. Let us examine an algorithm which finds the appropriate partition of $[a,b]$ and the corresponding minimax error ϵ.

Algorithm 2.

In Algorithm 1 it is evident that the positive integer k is a non-increasing function of the minimax error ϵ, say

$$k = K(\epsilon).$$

We can find lower and upper bounds for ϵ as follows. First, choose $\epsilon_0 > 0$ arbitrarily and use Algorithm 1 to calculate

$$k_0 = K(\epsilon_0).$$

If $k_0 > k$, ϵ_0 will be a lower bound for ϵ. We may then set $\epsilon_1 = 2\epsilon_0$ and calculate

$$k_1 = K(\epsilon_1).$$

If we repeat this calculation for k_1, with ϵ_1 replaced each time by $2\epsilon_1$, at some stage we will obtain a value of $k_1 \leq k$. This will give an upper bound for ϵ, say ϵ_1.

However, if initially we obtain $k_0 \leq k$ we may set $\epsilon_1 = \epsilon_0$ as an upper bound for ϵ and this time repeatedly halve ϵ_0, calculating

$$k_0 = K(\epsilon_0)$$

each time. Finally, we will obtain a value of $k_0 > k$, showing that the current value of ϵ_0 is a lower bound for ϵ.

Once we have obtained lower and upper bounds for ϵ, we may refine them by repeated bisection of the interval $[\epsilon_0, \epsilon_1]$, using Algorithm 1 at each stage to calculate

$$K\left(\tfrac{1}{2}(\epsilon_0 + \epsilon_1)\right).$$

The process is terminated when $\epsilon_1 - \epsilon_0$ is sufficiently small. The operation of Algorithm 1 corresponding to the final value of ϵ_0 gives the values of the sub-dividing points x_j and the minimax straight lines $c_j x + d_j$. Again, the approximating function is continuous on $[a,b]$, being simply a convex polygonal line.

It may be noted that at any stage, the operation of Algorithm 1 corresponding to lower and upper bounds ϵ_0 and ϵ_1 produces respectively lower and upper bounds for the sub-dividing points x_j. This is easily seen geometrically. Rounding error has given no trouble in a very wide range of numerical examples on which Algorithms 1 and 2 have been tried.

Finally, it may be observed that, by considering ϵ as a function of the sub-dividing point x_{k-1}, we could use regula falsi in Algorithm 2 instead of bisection of the interval.

Numerical example.

To illustrate these methods, let us consider the function e^x on the interval $[0,1]$. The table on the following page displays the best minimax approximation to e^x on

$[0,1]$ by four straight line segments, obtained by using Algorithm 2 with $k=4$. The corresponding value of ϵ is 0.006 579, all numbers being given to six decimal places.

From (2.36) we have the a priori bounds

$$0.0039 < \epsilon < 0.0107.$$

j	x_j	c_j	d_j
1	0.300 570	1.166 545	0.993 421
2	0.561 833	1.543 487	0.880 124
3	0.792 888	1.973 057	0.638 777
4	1.000 000	2.455 255	0.256 448

<u>Piecewise approximation to e^x on $[0,1]$</u>

The relation (2.33), interpreted as an asymptotic formula, would predict

$$\epsilon \simeq \frac{1}{64}\left(e^{\frac{1}{2}}-1\right)^2 \simeq 0.006\,576,$$

which is in error only in the sixth decimal place.

<u>An application to quadrature.</u>

Suppose that $f''(x) > 0$ on $[a,b]$ and we wish to approximate to

$$\int_a^b f(x)\,dx$$

with a maximum error of ϵ_0. Then, setting

$$\epsilon = \epsilon_0 / (b-a),$$

we may use Algorithm 1 to obtain a piecewise straight line approximation for $f(x)$ with maximum error ϵ. We will then be able to approximate to the above integral by the area under the convex polygonal line, which gives

$$\sum_{j=1}^{k} \int_{x_{j-1}}^{x_j} (c_j x + d_j)\, dx .$$

We have the inequality

$$\left| \int_a^b f(x)\, dx \;-\; \sum_{j=1}^{k} \int_{x_{j-1}}^{x_j} (c_j x + d_j)\, dx \right| \leq \epsilon_0 .$$

Thus the required integral may be replaced by the approximation

$$\sum_{j=1}^{k} \left[\tfrac{1}{2}\, c_j \,(x_j^2 - x_{j-1}^2) + d_j\,(x_j - x_{j-1}) \right]$$

with an error not greater than ϵ_0. This approach requires a rather large number of evaluations of $f(x)$ and $f'(x)$. On the other hand, it provides a sure bound for the error incurred. The error estimates for the most commonly used quadrature methods involve high order derivatives of the integrand. These estimates are often of little practical use.

It would be worth-while using the quadrature method put forward here only in a situation where the estimate of the error had sufficiently high priority to justify the large number of function evaluations.

2.5 Piecewise approximations satisfying interpolatory conditions

In this section, we combine some of the ideas used in Sections 2.2 and 2.3 .

Let Δ denote a partition of $[a,b]$:

$$a = x_0 < x_1 < \cdots < x_{k-1} < x_k = b.$$

Instead of the function $F(x)$ defined in Section 2.2, let us use $F_\Delta(x)$, where, for $x_{j-1} \leq x \leq x_j$,

$$F_\Delta(x) = [f(x) - L_j(x)] / (x - x_{j-1})(x - x_j),$$

with $L_j(x)$ defined by

$$L_j(x) = [(x - x_{j-1}) f(x_j) - (x - x_j) f(x_{j-1})] / (x_j - x_{j-1}).$$

That is, $F_\Delta(x)$ is defined piecewise on $[a,b]$. Now let

$q_{\Delta}(x)$ denote a function which interpolates $f(x)$ on Δ and is a piecewise polynomial of degree at most $n \geq 2$ on Δ. Let

$$E^{*}_{n,k}(f) = \inf \max_{a \leq x \leq b} |f(x) - q_{\Delta}(x)|, \tag{2.45}$$

where the infimum is over all such polynomials $q_{\Delta}(x)$ and all partitions Δ of $[a,b]$ into k sub-intervals. As in (2.23), we have

$$E^{*}_{n,k}(f) = \frac{\mu_{n-2}}{(n-1)!} \left(\frac{x_j - x_{j-1}}{2}\right)^{n+1} \cdot \left|F^{(n-1)}_{\Delta}(\xi_j)\right|, \tag{2.46}$$

with $x_{j-1} \leq \xi_j \leq x_j$. Arguing now as we did in obtaining (2.34) and (2.35), we obtain the inequality

$$E^{*}_{n,k}(f) \leq \frac{\mu_{n-2}}{(n-1)!} \left(\frac{b-a}{2k}\right)^{n+1} \cdot \sup_{\Delta,\, x} \left|F^{(n-1)}_{\Delta}(x)\right|, \tag{2.47}$$

where the supremum is over all partitions Δ of $[a,b]$ into k sub-intervals, and all $x \in [a,b]$. We also have the lower bound for $E_{n,k}(f)$:

$$E^{*}_{n,k}(f) \geq \frac{\mu_{n-2}}{(n-1)!} \left(\frac{b-a}{2k}\right)^{n+1} \cdot \inf_{\Delta,\, x} \left|F^{(n-1)}_{\Delta}(x)\right|. \tag{2.48}$$

An application to cubic splines.

Let $S_\Delta(x)$ denote a cubic spline approximation for $f(x)$ on a partition Δ :

$$a = x_0 < x_1 < \cdots < x_{k-1} < x_k = b$$

of the interval $[a,b]$. (See Section 1.5).

We have that

$$\max_{a \le x \le b} \left| f(x) - S_\Delta(x) \right| \geq E^*_{3,k}(f).$$

That is,

$$\begin{aligned} &\max_{a \le x \le b} \left| f(x) - S_\Delta(x) \right| \\ &\qquad\qquad \geq \frac{1}{10} \left(\frac{b-a}{2k} \right)^4 . \inf_{\Delta, x} \left| F''_\Delta(x) \right| . \end{aligned} \tag{2.49}$$

We have not made use of the continuity of the first and second derivatives of $S_\Delta(x)$, in obtaining the inequality

(2.49). If, further we did not take into account the interpolatory nature of $S_\Delta(x)$, that is the continuity of $S_\Delta(x)$ itself, we could use the inequality (2.35) and obtain

$$\max_{a\le x\le b} | f(x) - S_\Delta(x) | \geq \frac{1}{12}\left(\frac{b-a}{4k}\right)^4 \cdot \min_{a\le x\le b} | f^{(4)}(x)|. \tag{2.50}$$

For example, let us choose

$$f(x) = (\alpha+x)^{-1}$$

on the interval $[-1,1]$, with $\alpha > 1$. For $x_{j-1} \le x \le x_j$ we then have

$$F_\Delta''(x) = 2(\alpha+x_{j-1})^{-1}(\alpha+x_j)^{-1}(\alpha+x)^{-3}.$$

Thus from (2.49),

$$\max_{-1\le x\le 1} | (\alpha+x)^{-1} - S_\Delta(x)| \geq \frac{1}{5k^4}\cdot\frac{1}{(\alpha+1)^5}.$$

From (2.50) we obtain

$$\max_{-1 \le x \le 1} \left| (\alpha + x)^{-1} - S_\Delta(x) \right| \ge \frac{1}{8k^4} \cdot \frac{1}{(\alpha+1)^5} .$$

As would be expected, this last inequality is weaker than the previous one.

Chapter 3

ESTIMATE OF THE ERROR IN BEST L_p APPROXIMATIONS

3.1 A characterising property

The best minimax polynomial approximations are characterised by the equioscillation property of the error function. The best least squares polynomial approximations are those whose coefficients are the solutions of certain sets of linear equations, called the normal equations.

The characterising property of best L_p polynomial approximations, for any value of $p \geq 1$, is not quite so widely known. For this reason, a proof of Theorem 18 will be given here. It is based on a proof in Timan,1963. For convenience, let us restate :

Theorem 18. If $f(x)$ is continuous on $[a,b]$, then for any value of $p \geq 1$ a necessary and sufficient condition for $q(x) \in P_n$ to be the best L_p approximation for $f(x)$ on $[a,b]$ is that

$$\int_a^b r(x)\,|f(x) - q(x)|^{p-1} \cdot \operatorname{sign}[f(x) - q(x)]\,dx = 0 \quad (3.1)$$

for all $r(x) \in P_n$.

<u>Proof.</u> Suppose that, for some $q(x) \in P_n$, (3.1) holds for all $r(x) \in P_n$. Then we may write

$$\int_a^b |f(x) - q(x)|^p \, dx$$

$$= \int_a^b [f(x) - q(x)] . |f(x) - q(x)|^{p-1} . \operatorname{sign}[f(x) - q(x)] \, dx.$$

Hence, for any $r(x) \in P_n$, we have from (3.1) that

$$\int_a^b |f(x) - q(x)|^p dx$$

$$= \int_a^b [f(x) - r(x)] . |f(x) - q(x)|^{p-1} . \operatorname{sign}[f(x) - q(x)] \, dx$$

$$\leq \int_a^b |f(x) - r(x)| . |f(x) - q(x)|^{p-1} dx \tag{3.2}$$

$$\leq \left[\int_a^b |f(x) - r(x)|^p dx\right]^{\frac{1}{p}} . \left[\int_a^b |f(x) - q(x)|^p dx\right]^{\frac{p-1}{p}}$$

for $p>1$, by Hölder's inequality for integrals (Theorem 20).

Thus for $p>1$

$$\left[\int_a^b |f(x) - q(x)|^p dx \right]^{\frac{1}{p}} \leq \left[\int_a^b |f(x) - r(x)|^p dx \right]^{\frac{1}{p}}. \qquad (3.3)$$

From (3.2) we see that (3.3) holds for $p=1$ also. Since (3.3) holds for all $r(x) \in P_n$, we have proved the sufficiency of the condition (3.1). That is, $q(x)$ is the best approximation.

Conversely, suppose that $q(x)$ is the best approximation and that there exists a non-negative integer $k \leq n$ such that

$$\int_a^b x^k \cdot |f(x) - q(x)|^{p-1} \cdot \operatorname{sign}[f(x) - q(x)]\, dx = \delta \neq 0. \qquad (3.4)$$

Now let

$$r_\epsilon(x) = q(x) + \epsilon x^k. \qquad (3.5)$$

Then, for some $\epsilon \neq 0$ (ϵ not neccessarily positive),

$$\epsilon \int_a^b x^k |f(x) - r_\epsilon(x)|^{p-1} . \operatorname{sign}[f(x) - r_\epsilon(x)]\, dx > 0. \tag{3.6}$$

This follows from (3.4) by taking $|\epsilon|$ sufficiently small and keeping

$$\operatorname{sign}(\epsilon) = \operatorname{sign}(\delta).$$

Hence

$$\begin{aligned} &\int_a^b |f(x) - r_\epsilon(x)|^p\, dx \\ &\quad = \int_a^b [f(x) - r_\epsilon(x)] . |f(x) - r_\epsilon(x)|^{p-1} . \operatorname{sign}[f(x) - r_\epsilon(x)]\, dx. \end{aligned} \tag{3.7}$$

Using (3.6) we have from (3.7) that

$$\begin{aligned} &\int_a^b |f(x) - r_\epsilon(x)|^p\, dx \\ &\quad < \int_a^b [f(x) - q(x)] . |f(x) - r_\epsilon(x)|^{p-1} . \operatorname{sign}[f(x) - r_\epsilon(x)]\, dx \end{aligned}$$

$$\leq \int_a^b |f(x) - q(x)| . |f(x) - r_\epsilon(x)|^{p-1} dx. \tag{3.8}$$

From (3.8), using Hölder's inequality exactly as in the earlier part of the proof, we have

$$\int_a^b |f(x) - r_\epsilon(x)|^p dx < \int_a^b |f(x) - q(x)|^p dx.$$

Since $r_\epsilon(x) \in P_n$ this last inequality provides a contradiction to the assumption that $q(x)$ is the best approximation. This completes the proof.

3.2 The interpolatory property

In Chapter 2, the derivation of the result

$$E_n(f) = \frac{2}{(n+1)!} \left(\frac{b-a}{4}\right)^{n+1} |f^{(n+1)}(\xi)| \tag{3.9}$$

depended on the interesting property that, in the minimax approximation of a continuous function $f(x)$, the best polynomial interpolates $f(x)$ at $n+1$ points on $[a,b]$. This also holds for best L_2 (i.e. least squares) polynomial approximations. See, for example, Davis, 1963.

More generally, this is true for best L_p polynomial approximations, for any value of $p \geq 1$. This result is implicit in Timan, 1963. Here it is stated explicitly:

Theorem 31. For any $p \geq 1$, if $q(x)$ is the best L_p polynomial approximation of degree not greater than n to a continuous function $f(x)$ on $[a,b]$, then there exist $n+1$ points on $[a,b]$ at which $q(x)$ interpolates $f(x)$.

Proof. This follows simply from Theorem 18. Consider the number of changes in sign of $f(x)-q(x)$ on $[a,b]$. Since from Theorem 18

$$\int_a^b |f(x) - q(x)|^{p-1} \cdot \operatorname{sign}[f(x) - q(x)]\, dx = 0$$

it follows that there must be at least one sign change. Let us suppose that sign changes occur only at $x_0, x_1, \dots, x_k$ within $[a,b]$, where $0 \leq k < n$. Then the function

$$(x-x_0)\cdots(x-x_k) \cdot \operatorname{sign}[f(x) - q(x)]$$

has constant sign on $[a,b]$ and therefore

$$\int_a^b (x-x_0)\cdots(x-x_k).\,|f(x)-q(x)|^{p-1}.\,\mathrm{sign}\,[f(x)-q(x)]\,dx$$

is non-zero. Since the polynomial

$$(x-x_0)(x-x_1)\;\cdots\;(x-x_k)$$

belongs to P_n, this contradicts Theorem 18 and completes the proof.

Now let us write, for $p \geq 1$,

$$E_n^{(p)}(f) = \inf_{q(x)\in P_n} \left[\int_a^b |f(x)-q(x)|^p\,dx\right]^{\frac{1}{p}}, \tag{3.10}$$

so that $E_n^{(\infty)}(f)$ coincides with $E_n(f)$ in Chapter 2. By Theorem 31, we may write

$$f(x) - q^*(x) = \frac{1}{(n+1)!}(x-x_0^*)\cdots(x-x_n^*).\,f^{(n+1)}(\xi_x^*), \tag{3.11}$$

where $q^*(x)$ is the polynomial for which the infimum (3.10) is attained. We assume continuity of $f^{(n+1)}(x)$. Thus from (3.11)

$$E_n^{(p)}(f) = \frac{1}{(n+1)!}\left[\int_a^b |(x-x_0^*)\cdots(x-x_n^*)|^p . |f^{(n+1)}(\xi_x^*)|^p dx\right]^{\frac{1}{p}}. \tag{3.12}$$

Let

$$\delta_n^{(p)} = \inf_{(y_j)} \left[\int_{-1}^{1} |(y-y_0)\cdots(y-y_n)|^p dy\right]^{\frac{1}{p}}. \tag{3.13}$$

The infimum is attained (see Nikolskii, 1964) for a set of points $\{y_0, \ldots, y_n\}$ contained in $[-1,1]$. Now let us transfom $a \le x \le b$ into $-1 \le y \le 1$ by putting

$$x = [(b-a)y + (b+a)]/2.$$

Let $x_0, \ldots, x_n$ be the points on $[a,b]$ corresponding to the minimising y_j for (3.13), and let $q(x)$ denote the interpolating polynomial for $f(x)$ constructed at $x = x_0, \ldots, x_n$. We can therefore write

$$f(x) - q(x) = \frac{1}{(n+1)!}(x-x_0)\cdots(x-x_n). f^{(n+1)}(\xi_x).$$

Now

$$E_n^{(p)} \leq \left[\int_a^b |f(x) - q(x)|^p dx \right]^{\frac{1}{p}} \tag{3.14}$$

and, replacing the right side of (3.14) by the right side of (3.12) without the stars ($\ast$), we have

$$E_n^{(p)}(f) \leq$$

$$\frac{1}{(n+1)!} \max_{a \leq x \leq b} |f^{(n+1)}(x)| . \left[\int_a^b |(x-x_0) \cdots (x-x_n)|^p dx \right]^{\frac{1}{p}}.$$

That is,

$$E_n^{(p)}(f) \leq \frac{\delta_n^{(p)}}{(n+1)!} \left(\frac{b-a}{2} \right)^{n+1+\frac{1}{p}} . \max_{a \leq x \leq b} |f^{(n+1)}(x)| . \tag{3.15}$$

From (3.12), we also have

$$E_n^{(p)}(f) \geq \frac{\delta_n^{(p)}}{(n+1)!} \left(\frac{b-a}{2} \right)^{n+1+\frac{1}{p}} . \min_{a \leq x \leq b} |f^{(n+1)}(x)| . \tag{3.16}$$

It follows from these last two inequalities, by continuity of $f^{(n+1)}(x)$, that

$$E_n^{(p)}(f) = \frac{\delta_n^{(p)}}{(n+1)!} \left(\frac{b-a}{2}\right)^{n+1+\frac{1}{p}} . \left|f^{(n+1)}(\xi)\right|, \tag{3.17}$$

for some $\xi \in [a,b]$. Letting $p \to \infty$ in (3.17), we obtain (3.9).

3.3 The case $p=2$

Let us now consider further the special case where $p=2$. As in Definition 10 of Chapter 1, let $Q_j(x)$ denote the Legendre polynomial of degree j. Then, by Theorem 13, the infimum (3.13) with $p=2$ is attained by the polynomial

$$2^{n+1} \, Q_{n+1}(y) \Big/ \binom{2n+2}{n+1}$$

and has the value

$$\delta_n^{(2)} = \left(\frac{2}{2n+3}\right)^{\frac{1}{2}} . \; 2^{n+1} \Big/ \binom{2n+2}{n+1}. \tag{3.18}$$

Timan, 1963, notes that for all values of n

$$\frac{1}{\sqrt{2}} \cdot \frac{1}{2^n} < \delta_n^{(2)} < \frac{\sqrt{2}}{2^n} . \tag{3.19}$$

3.4 The case $p=1$

When $p=1$, the two polynomials $q(x)$ and $q^*(x)$ of Section 3.2 coincide. This is shown in Timan, 1963. That is, the best L_1 approximating polynomial of degree at most n to $f(x)$ on $[a,b]$ is simply the interpolating polynomial for $f(x)$ constructed at the abscissae which minimise

$$\inf_{(x_j)} \int_a^b |(x-x_0)\cdots(x-x_n)|\, dx.$$

We will return to this result in Section 3.6.

Meanwhile, we note that from Theorem 10 we have

$$\delta_n^{(1)} = 1/2^n. \qquad (3.20)$$

3.5 General values of p

Not so much appears to be known about $\delta_n^{(p)}$ for values of p other than $p = 1, 2, \infty$. See for example Nikolskii, 1964. However, we may write

$$\delta_n^{(p)} \leq \left[\int_{-1}^{1} \left| \frac{1}{2^n} T_{n+1}(x) \right|^p dx \right]^{\frac{1}{p}}, \tag{3.21}$$

where $T_{n+1}(x)$ is the Chebyshev polynomial (Definition 6). Thus

$$\delta_n^{(p)} \leq 1 \, / \, 2^{n - \frac{1}{p}} . \tag{3.22}$$

Equality holds in (3.22) only for $p = \infty$.

Also, from Theorem 19,

$$\left[\int_{-1}^{1} |(x - x_0) \cdots (x - x_n)|^p \, dx \right]^{\frac{1}{p}} \geq 2^{\frac{1}{p} - 1} \int_{-1}^{1} |(x - x_0) \cdots (x - x_n)| \, dx$$

for $p \geq 1$. Therefore

$$\delta_n^{(p)} \geq 2^{\frac{1}{p} - 1} \min_{(x_i)} \int_{-1}^{1} |(x - x_0) \cdots (x - x_n)| \, dx .$$

This gives the inequality

$$\delta_n^{(p)} \geq 1 / 2^{n+1-\frac{1}{p}} . \tag{3.23}$$

In this case, equality holds only for $p=1$.

Combining (3.22) and (3.23) gives

$$1/2^{n+1-\frac{1}{p}} \leq \delta_n^{(p)} \leq 1/2^{n-\frac{1}{p}}, \tag{3.24}$$

which generalises Timan's inequalities (3.19). Further, let us write

$$\theta_n^{(p)} = 2^{n+1+\frac{1}{p}} \delta_n^{(p)} . \tag{3.25}$$

Then (3.17) may be rewritten as

$$E_n^{(p)}(f) = \frac{\theta_n^{(p)}}{(n+1)!} \left(\frac{b-a}{4}\right)^{n+1+\frac{1}{p}} . |f^{(n+1)}(\xi)|, \tag{3.26}$$

where

$$2^{\frac{2}{p}} \leq \theta_n^{(p)} \leq 2^{1+\frac{2}{p}} \tag{3.27}$$

From the previous sections, we see that

$$\theta_n^{(1)} = 4 \qquad \text{and} \qquad \vartheta_n^{(\infty)} = 2 .$$

For other values of p, $1 < p < \infty$, both inequalities (3.27) hold strictly. It may be noted that $\theta_n^{(p)}$ depends only on n and p, and not on the function $f(x)$ nor on the interval $[a,b]$.

3.6 Further remarks on L_1 approximations

For completeness, and for the sake of an application to be described later in this section, we now consider two theorems on L_1 approximation. The first of these is stated explicitly and the second is implicit in the account of Timan, 1963.

Theorem 32. For $k = 0, 1, \ldots, n$,

$$\int_{-1}^{1} x^k . \operatorname{sign}\left[\sin (n+2) \cos^{-1} x\right] dx = 0. \qquad (3.28)$$

Proof. Let us put

$$x = \cos \theta .$$

Then the integral in (3.28) is

$$\int_0^{\pi} \sin\theta \ \cos^k\theta \ . \ \mathrm{sign}\,[\sin(n+2)\theta]\ d\theta$$

$$= \sum_{j=1}^{n+2} (-1)^{j-1} \int_{(j-1)\pi/(n+2)}^{j\pi/(n+2)} \sin\theta \ \cos^k\theta \ d\theta$$

$$= \frac{1}{(k+1)} \sum_{j=1}^{n+2} (-1)^j \ \Big[\cos^{k+1}\theta\Big]_{(j-1)\pi/(n+2)}^{j\pi/(n+2)}$$

$$= \frac{1}{2(k+1)} \sum_{j=0}^{n+2}{}'' (-1)^j \ \cos^{k+1} \frac{j\pi}{(n+2)} \ , \qquad (3.29)$$

where $\sum''$ denotes a sum whose first and last terms are halved. Now $\cos(k+1)\theta$ may be expressed as a polynomial of degree $k+1$ in $\cos\theta$, thatis as $T_{k+1}(\cos\theta)$. Conversely, we can find α_j such that

$$\cos^{k+1}\theta \ = \ \sum_{j=0}^{k+1} \alpha_j \cos j\theta .$$

Hence we can show that (3.29) vanishes for $k=0, 1, \dots, n$ if we can show that

$$\sum_{j=0}^{n+2}{}'' \ (-1)^j \quad \cos \frac{j(k+1)\pi}{(n+2)} \ = \ 0 \qquad (3.30)$$

for $k=0, 1, \ldots, n$. This is easily verified by expressing $\cos\theta$ as the real part of $e^{i\theta}$ and summing the geometric series then obtained from (3.30). This completes the proof of Theorem 32.

This now enables us to prove the most interesting result :

<u>Theorem 33.</u> The best L_1 approximation to a continuous function $f(x)$ on $[-1,1]$ by a polynomial of degree not greater than n is simply the interpolating polynomial for $f(x)$ constructed at the zeros of $\sin(n+2)\cos^{-1}x$ in the interior of $[-1,1]$.

<u>Proof.</u> Let $q(x) \in P_n$ be the interpolating polynomial for $f(x)$ constructed at the zeros of $\sin(n+2)\cos^{-1}x$ in the interior of $[-1,1]$. It follows that $f(x) - q(x)$ changes sign on $[-1,1]$ at the same points as $\sin(n+2)\cos^{-1}x$ changes sign. Thus from Theorem 32

$$\int_{-1}^{1} x^k . \ \operatorname{sign}\left[f(x) - q(x)\right] dx \ = \ 0 \qquad (3.31)$$

for $k = 0, 1, \ldots, n$. It follows from Theorem 18 that $q(x)$ is the best approximation.

It may be noted also that Theorem 10 follows from this result. For

$$\inf_{(c_j)} \int_{-1}^{1} | x^n + c_{n-1} x^{n-1} + \cdots + c_0 | \, dx \tag{3.32}$$

is attained when

$$c_{n-1} x^{n-1} + \cdots + c_1 x + c_0$$

is the interpolating polynomial for x^n constructed at the zeros of $\sin(n+1)\cos^{-1}x$ in the interior of $[-1,1]$, which are

$$x_j = \cos \frac{j\pi}{(n+1)}, \qquad j = 1, 2, \ldots, n.$$

These are the zeros of $U_n(x)$, the Chebyshev polynomial of the second kind, as in Definition 8. The integrand in (3.32) may be replaced by the modulus of the error formula for the interpolating polynomial,

$$\frac{1}{n!}\,(x-x_1)\cdots(x-x_n)\,\frac{d^n}{dx^n}\,(x^n)_{x=\xi_x}\quad , \tag{3.33}$$

where

$$x_j \quad = \quad \cos\frac{j\pi}{n+1} \quad .$$

That is, the integrand in (3.32) is a multiple of $U_n(x)$. It is easily checked from the definition that the leading coefficient of $U_n(x)$ is 2^n, which completes the proof of Theorem 10.

An application to quadrature.

A common class of quadrature formulae is obtained by making the approximation

$$\int_a^b f(x)\,dx \quad \simeq \quad \int_a^b q(x)\,dx \quad ,$$

where $q(x)$ is an interpolating polynomial for $f(x)$ at certain points $x_0, x_1, \ldots, x_n$ on $[a,b]$. These are called interpolatory quadrature formulae. Well known examples are the (closed) Newton - Cotes formulae in which $q(x)$ interpolates $f(x)$ at equally spaced points including the end points, the open Newton - Cotes formulae where the interpolating points are equally spaced but exclude the end points, and the Gauss - Legendre formulae where the interpolating points are the zeros of the Legendre polynomials. (See Davis and Rabinowitz, 1967).

Having studied L_1 polynomial approximation above, it is natural to suggest transforming the range of integration $[a,b]$ onto $[-1,1]$ and taking as $q(x)$ the polynomial which interpolates $f(x)$ at the zeros of $\sin(n+2)\cos^{-1}x$ in the interior of $[-1,1]$. That is, interpolating $f(x)$ at the zeros of $U_{n+1}(x)$. For, by Theorem 33, this will give the polynomial for which

$$\inf_{q(x)\in P_n} \int_{-1}^{1} |f(x) - q(x)|\, dx \qquad (3.34)$$

is attained.

For the first few values of n we obtain the following formulae as approximations to

$$\int_{-1}^{1} f(x)\,dx.$$

(A linear transformation will give the appropriate formulae when the range of integration is $[a,b]$).

(A) $f(\tfrac{1}{2}) + f(-\tfrac{1}{2})$

(B) $\tfrac{2}{3}\left[f(-\tfrac{1}{\sqrt{2}}) + f(0) + f(\tfrac{1}{\sqrt{2}})\right]$

(3.35)

(C) $$\left(\tfrac{1}{2} - \tfrac{\sqrt{5}}{30}\right)\left[f\left(\tfrac{\sqrt{5}+1}{4}\right) + f\left(-\tfrac{\sqrt{5}-1}{4}\right)\right] + \left(\tfrac{1}{2} + \tfrac{\sqrt{5}}{30}\right)\left[f\left(\tfrac{\sqrt{5}-1}{4}\right) + f\left(-\tfrac{\sqrt{5}+1}{4}\right)\right]$$

(D) $\tfrac{1}{45}\left[26 f(0) + 18\left(f(\tfrac{1}{2}) + f(-\tfrac{1}{2})\right) + 14\left(f(\tfrac{\sqrt{3}}{2}) + f(-\tfrac{\sqrt{3}}{2})\right)\right]$

These formulae are not new. They have already been studied by Filippi, 1964, who was led to their discovery by a different route to that described above. Filippi's starting point is the integration formula of Clenshaw and Curtis, 1960. This evaluates

$$\int_{-1}^{x} f(t)\, dt$$

by expressing the integrand as a Chebyshev series and integrating this, using (1.8), to give the integral also as a Chebyshev series. Filippi modifies this idea slightly to cut out the integration step. He writes

$$F(x) = F(-1) + \int_{-1}^{x} f(t)\, dt \tag{3.36}$$

and expresses

$$F(x) = \sum_{j=0}^{\infty}{}' A_j T_j(x) . \tag{3.37}$$

Thus

$$A_j = \frac{2}{\pi} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} F(x) T_j(x)\, dx \tag{3.38}$$

for $j=0, 1, 2, \ldots$. Using the identity

$$\frac{d}{dx}\left((1-x^2)^{\frac{1}{2}}\, T_n'(x)\right) = -n^2\, (1-x^2)^{-\frac{1}{2}}\, T_n(x)$$

and integrating (3.38) by parts, he obtains

$$A_j = \frac{2}{\pi j^2} \int_{-1}^{1} (1-x^2)^{\frac{1}{2}} f(x)\, T_j'(x)\, dx. \tag{3.39}$$

Now from (1.11) and Theorem 12 we note that the polynomials $T_j'(x)$ are orthogonal on $[-1,1]$ with respect to $(1-x^2)^{\frac{1}{2}}$. Now let

$$r_n(x) = \sum_{j=0}^{n} A_{j+1}\, T_{j+1}'(x), \tag{3.40}$$

which from (3.36) and (3.37) is seen to be a truncated series for $f(x)$. From (3.39) and Theorem 14 we see that $r_n(x)$ is the least squares approximation for $f(x)$ with respect to $(1-x^2)^{\frac{1}{2}}$. That is,

$$\min_{q(x)\in P_n} \int_{-1}^{1} (1-x^2)^{\frac{1}{2}} \left[f(x) - q(x)\right]^2 dx$$

is attained for the choice

$$q(x) = r_n(x) .$$

Filippi then considers $\bar{r}_n(x)$, the interpolating polynomial for $f(x)$ constructed at the zeros of $T'_{n+2}(x)$, that is at the zeros of $U_{n+1}(x)$. We have

$$f(x) - \bar{r}_n(x) = \frac{1}{(n+1)!} (x-x_1)\cdots(x-x_{n+1}) f^{(n+1)}(\xi_x),$$

so that

$$\begin{aligned} &\int_{-1}^{1} (1-x^2)^{\frac{1}{2}} [f(x) - \bar{r}_n(x)]^2 dx \\ &= \left[\frac{1}{(n+1)!} f^{(n+1)}(\bar{\xi}) \right]^2 \int_{-1}^{1} (1-x^2)^{\frac{1}{2}} (x-x_1)^2 \cdots (x-x_{n+1})^2 dx. \end{aligned} \tag{3.41}$$

From Theorem 15, we see that this choice of $x_1, \ldots, x_{n+1}$ minimises the integral on the right side of (3.41). Filippi argues from this that $\bar{r}_n(x)$ will usually give a close approximation to $r_n(x)$. Thus he suggests using the approximation

$$\int_{-1}^{1} f(x)\, dx \simeq \int_{-1}^{1} \bar{r}_n(x)\, dx .$$

For $n=1, 2, 3, 4$, this gives the formulae (3.35).

Filippi's integration formulae are compared numerically with both the Clenshaw - Curtis and the Gauss - Legendre formulae by Wright, 1966. In this comparison, none of these integration formulae can be dismissed, since each emerges as superior to the others for certain integrals.

The observation made here that the Filippi formulae satisfy the minimum L_1 property, as in (3.34), appears to be new.

3.7 Piecewise approximations

In Section 2.3, we considered minimax piecewise polynomial approximations. Here we shall consider L_p piecewise polynomial approximations. Let us write

$$E_{n,k}^{(p)}(f, a, b) = \inf \left[\sum_{j=1}^{k} \int_{x_{j-1}}^{x_j} | f(x) - q_j(x)|^p dx \right]^{\frac{1}{p}}, \tag{3.42}$$

where the infimum is over all partitions of $[a,b]$ into k sub-intervals, with sub-dividing points

$$a = x_0 < x_1 < \cdots < x_{k-1} < x_k = b,$$

and over all polynomial approximations $q_j(x) \in P_n$ on the j^{th} sub-interval, $j = 1, 2, \ldots, k$. An argument similar to the one used in Section 2.3 shows that there exist points x_j^* and polynomials $q_j^*(x)$ for which the infimum (3.42) is attained. Also, it is clear that for the best piecewise approximation, $q_j^*(x)$ must be the best L_p approximation for $f(x)$ on the interval $[x_{j-1}^*, x_j^*]$. Therefore

$$E_{n,k}^{(p)}(f, a, b) = \left[\sum_{j=1}^{k} \left[\epsilon_n^{(p)}(f, x_{j-1}^*, x_j^*)\right]^p\right]^{\frac{1}{p}}. \tag{3.43}$$

That is,

$$E_{n,k}^{(p)}(f, a, b) = \left[\sum_{j=1}^{k} \left[\frac{\theta_n^{(p)}}{(n+1)!}\left(\frac{x_j^* - x_{j-1}^*}{4}\right)^{n+1+\frac{1}{p}} \cdot \left|f^{(n+1)}(\xi_j)\right|\right]^p\right]^{\frac{1}{p}}.$$

Now on the right side of (3.43) each of the k summands must be equal, otherwise we could obtain a better approximation by taking another partition of $[a,b]$. It follows that, for $j=1, 2, \dots, k$,

$$\left(\frac{x_j^* - x_{j-1}^*}{4}\right)^{n+1+\frac{1}{p}} \cdot |f^{(n+1)}(\xi_j)| \leq \left(\frac{b-a}{4k}\right)^{n+1+\frac{1}{p}} \cdot \max_{a\leq x\leq b} |f^{(n+1)}(x)|$$

and so

$$E_{n,k}^{(p)}(f, a, b) \leq \frac{\theta_n^{(p)}}{(n+1)!} \left(\frac{b-a}{4k}\right)^{n+1+\frac{1}{p}} \cdot k^{\frac{1}{p}} \cdot \max_{a\leq x\leq b} |f^{(n+1)}(x)|. \tag{3.44}$$

A lower bound for $E_{n,k}^{(p)}(f,a,b)$ is obtained by arguing similarly that, for $j=1, 2, \dots, k$,

$$\left(\frac{x_j^* - x_{j-1}^*}{4}\right)^{n+1+\frac{1}{p}} \cdot |f^{(n+1)}(\xi_j)| \geq \left(\frac{b-a}{4k}\right)^{n+1+\frac{1}{p}} \cdot \min_{a\leq x\leq b} |f^{(n+1)}(x)|.$$

Therefore we have the inequality

$$E_{n,k}^{(p)}(f,\ a,\ b) \quad \geq$$

$$\frac{\theta_n^{(p)}}{(n+1)!}\left(\frac{b-a}{4k}\right)^{n+1+\frac{1}{p}} .\ k^{\frac{1}{p}} .\ \min_{a \leq x \leq b} \left| f^{(n+1)}(x) \right| . \tag{3.45}$$

By continuity of $f^{(n+1)}(x)$, we may combine the two inequalities (3.44) and (3.45) to give

$$E_{n,k}^{(p)}(f,\ a,\ b) \quad =$$

$$\frac{\theta_n^{(p)}}{(n+1)!}\left(\frac{b-a}{4k}\right)^{n+1+\frac{1}{p}} .\ k^{\frac{1}{p}} .\ \left| f^{(n+1)}(\xi) \right| , \tag{3.46}$$

for some $\xi \in [a,b]$. This generalises (2.36), remembering that $\theta_n^{(\infty)} = 2$.

We can also generalise the asymptotic result (2.33) as follows. Since all k summands on the right side of (3.43) are equal, we have

$$E_{n,k}^{(p)}(f,\ a,\ b) \;=\; k^{\frac{1}{p}}\ E_n^{(p)}(f,\ x_{j-1}^*, x_j^*),$$

for $1 \leq j \leq k$. That is,

$$E^{(p)}_{n,k}(f, a, b) = \frac{\theta^{(p)}_n}{(n+1)!} \left(\frac{x_j^* - x_{j-1}^*}{4} \right)^{n+1+\frac{1}{p}} . k^{\frac{1}{p}} . \left| f^{(n+1)}(\xi_j) \right| . \tag{3.47}$$

We can now proceed as in Section 2.3, writing

$$k \left[\frac{(n+1)!}{\theta^{(p)}_n} E^{(p)}_{n,k}(f, a, b) \right]^{\frac{1}{n+1+\frac{1}{p}}} = \frac{1}{4} k^{\frac{1}{np+p+1}} \sum_{j=1}^{k} (x_j^* - x_{j-1}^*) . \left| f^{(n+1)}(\xi_j) \right|^{\frac{1}{n+1+\frac{1}{p}}} . \tag{3.48}$$

Let us assume that $f(x)$ cannot be represented exactly by a polynomial of degree n on any sub-interval of $[a,b]$. Then we have that, as $k \to \infty$ in (3.48), the length of the largest sub-interval $[x^*_{j-1}, x^*_j]$ will tend to zero. As $k \to \infty$, we may therefore replace the summation on the right hand side of (3.48) by the Riemann integral. This gives us the asymptotic result,

$$\lim_{k\to\infty} k^{n+1} E^{(p)}_{n,k}(f, a, b) =$$

$$\frac{\theta^{(p)}_n}{(n+1)!}\left[\frac{1}{4}\int_a^b |f^{(n+1)}(x)|^{\frac{1}{n+1+\frac{1}{p}}} dx\right]^{n+1+\frac{1}{p}}. \qquad (3.49)$$

The special case of (3.49) with $n=1$, $p=2$ is given by Ream, 1961, in addition to the other special case $n=1$, $p=\infty$ already noted in Section 2.3.

To facilitate comparison with Ream's result for $n=1$, $p=2$ let us recall that

$$\theta^{(p)}_n = 2^{n+1+\frac{1}{p}} \delta^{(p)}_n ,$$

and from (3.18) that

$$\delta^{(2)}_n = \left(\frac{2}{2n+3}\right)^{\frac{1}{2}} . \; 2^{n+1} \Big/ \binom{2n+2}{n+1} .$$

For $n=1$, $p=2$ we then obtain from (3.49)

$$\lim_{k\to\infty} k^2 E^{(2)}_{1,k}(f, a, b) = \frac{1}{12\sqrt{5}} \left[\int_a^b |f''(x)|^{0.4} dx \right]^{2.5} . \tag{3.50}$$

This agrees with Ream's result.

Chapter 4

APPROXIMATION OF CONVEX DATA

4.1 Introduction

The last two chapters were devoted mainly to the problem of estimating the error in best L_p polynomial approximations. This chapter is of a different nature. It is concerned with a much more practical problem.

One is sometimes presented with a discrete set of data which is convex, or where physical reasons suggest that the data would be convex but for experimental error. In these circumstances, it seems unsatisfactory to use the standard least squares method, which may result in approximations with undesired inflexions. It would appear preferable to make use of the knowledge of the convexity of the data, in order to produce better approximations to the hidden convex function, say $g(x)$, and its first few derivatives, if these are required.

Let us suppose that $g''(x) \geq 0$ on a finite interval $[a,b]$. If we can find a sequence of functions

$$\psi_0(x), \quad \psi_1(x), \quad \psi_2(x), \quad \cdots,$$

with

$$\psi_j''(x) \geq 0, \qquad j = 0, 1, 2, \cdots,$$

on $[a,b]$, we can consider using functions of the form

$$\Phi_n(x) = \sum_{j=0}^{n} c_j \psi_j(x), \qquad c_j \geq 0, \tag{4.1}$$

for approximating to $g(x)$. In (4.1), by using a sum of non-negative multiples of the component functions $\psi_j(x)$, we also have that

$$\Phi_n''(x) \geq 0, \qquad a \leq x \leq b.$$

It still remains to determine two things: what Rice, 1964, calls the 'norm and form' of the approximation. The first of these, the choice of norm, will decide on the 'best' values for the coefficients c_j in (4.1). The second task is to make a suitable choice of the component functions $\psi_j(x)$, which is more difficult.

At this stage, it may be helpful to recall why polynomials have been so extensively and successfully used, particularly in approximating to discrete data. This is partly because polynomials are easily evaluated, but mainly because of Weierstrass' theorem, which shows that linear combinations of the monomials x^j are good enough for approximating arbitrarily closely to any continuous function. What we require here is a choice of the functions $\psi_j(x)$ for which we can state a similar theorem. That is, any function $g(x)$ such that $g''(x) \geqslant 0$ has to be approximable with arbitrary accuracy, on a finite interval, by a sum of non-negative multiples of the functions $\psi_j(x)$.

4.2 Choice of the component convex functions

It should be mentioned immediately that the Bernstein polynomials themselves (Definition 1) provide an apparent solution to our problem. For if $g(x)$ is convex, so also is $B_n(g;x)$. This is proved in Davis, 1963. However, it is also well known that the rate of convergence of $B_n(g;x)$ to $g(x)$ is given by :

<u>Theorem 34.</u> Let $g(x)$ be bounded on $[0,1]$ and let x_0 be a point of $[0,1]$ at which $g''(x_0)$ exists. Then

$$\lim_{n\to\infty} n\left[B_n(g;x_0) - g(x_0)\right] = \tfrac{1}{2}x_0(1-x_0)\, g''(x_0). \qquad (4.2)$$

(Proof in Davis, 1963).

That is, asymptotically, to halve the error we have to double the degree of the approximating Bernstein polynomial, which is clearly a poor practical proposition. Besides, as has already been remarked, the data may not actually be convex, due to experimental error. Therefore the Bernstein polynomials may also not be convex. We therefore discard the possibility of using the Bernstein polynomials directly. Instead, we prove :

<u>Theorem 35.</u> There exists a sequence of component functions

$$\psi_0(x), \quad \psi_1(x), \quad \psi_2(x), \quad \ldots ,$$

with $\psi_j''(x) \geq 0$ such that any function $g(x)$, with $g''(x) \geq 0$ and continuous, may be approximated with arbitrary accuracy on a finite interval by a sum of non-negative multiples of the component functions.

Proof. Let us suppose that we wish to approximate to $g(x)$ over the interval $[0,1]$. We can make a linear change of variable, if necessary, to transform any finite interval $[a,b]$ onto $[0,1]$. We use the Bernstein polynomials indirectly and write

$$B_n(g'';x) = \sum_{j=0}^{n} \binom{n}{j} x^j (1-x)^{n-j} g''\left(\tfrac{j}{n}\right). \tag{4.3}$$

Let us observe that $x^j(1-x)^{n-j} \geq 0$ on $[0,1]$ and that in (4.3) $g''(x)$ is being approximated by a sum of non-negative multiples of the polynomials $x^j(1-x)^{n-j}$.

For $n \geq 2$, define $q_n(x)$ by

$$\begin{aligned} q_n''(x) &= B_{n-2}(g'';x), \\ q_n'(0) &= g'(0), \\ q_n(0) &= g(0). \end{aligned} \tag{4.4}$$

Also define $\beta_{j,n}(x)$, for $2 \leq j \leq n$, by

$$\beta''_{j,n}(x) = x^{j-2}(1-x)^{n-j},$$

$$\beta'_{j,n}(0) = \beta_{j,n}(0) = 0. \tag{4.5}$$

To complete the definition of the polynomials $\beta_{j,n}(x)$, we define

$$\beta_{0,n}(x) = \operatorname{sign}[g(0)],$$

$$\beta_{1,n}(x) = x \,.\, \operatorname{sign}[g'(0)]. \tag{4.6}$$

We then have that

$$q_n(x) = \sum_{j=0}^{n} c_j \beta_{j,n}(x), \tag{4.7}$$

where $c_j \geqslant 0$ and $\beta''_{j,n}(x) \geqslant 0$ on $[0,1]$. Now, given any $\epsilon > 0$, it follows from Bernstein's theorem (Theorem 2) that there exists an integer n for which

$$| B_{n-2}(g''; x) - g''(x) | < \epsilon$$

on $[0,1]$. That is ,

$$| q_n''(x) - g''(x) | < \epsilon$$

on $[0,1]$ and therefore, for $0 \leq x \leq 1$,

$$\left| \int_0^x (q_n''(t) - g''(t))\, dt \right| \leq \int_0^x | q_n''(t) - g''(t) |\, dt$$

$$\leq \epsilon . x \leq \epsilon . \tag{4.8}$$

Using (4.4), the inequalities (4.8) give

$$| q_n'(x) - g'(x) | \leq \epsilon , \tag{4.9}$$

for $0 \leq x \leq 1$. Similarly, another integration shows that

$$| q_n(x) - g(x) | \leq \epsilon , \tag{4.10}$$

for $0 \leq x \leq 1$. Recalling the definition of $q_n(x)$ in (4.7), this last inequality (4.10) completes the proof.

Note that the polynomials $\beta_{j,n}(x)$ may be enumerated, say in the order

$$\begin{tikzcd}
\beta_{0,2} \arrow[r] & \beta_{1,2} \arrow[dl] & \beta_{2,2} \arrow[d] \\
\beta_{0,3} \arrow[d] & \beta_{1,3} \arrow[ur] & {} \\
\beta_{0,4} \arrow[ur] & &
\end{tikzcd}$$

and re-labelled $\psi_0(x),\ \psi_1(x),\ \psi_2(x), \dots$.

4.3 'Best' convex approximations

In practice, we may obtain convex approximations to $g(x)$ in the following way. First, choose a value of $n \geqslant 2$. This may be increased subsequently, if necessary. Then set,

$$\psi_j(x) = \beta_{j,n}(x), \qquad 2 \leq j \leq n. \tag{4.11}$$

Also, let us put

$$\psi_0(x) = \operatorname{sign}[g(0)],$$
$$\psi_1(x) = x \cdot \operatorname{sign}[g'(0)]. \qquad (4.12)$$

The signs of $g(0)$ and $g'(0)$ may be inferred from the data. The polynomials $\beta_{j,n}(x)$ are easily computed, since for $j \geqslant 2$,

$$\beta''_{j,n}(x) = x^{j-2}(1-x)^{n-j}$$
$$= x^{j-2}\sum_{i=0}^{n-j}(-1)^i\binom{n-j}{i}x^i.$$

Therefore, from (4.5),

$$\beta_{j,n}(x) = x^j \sum_{i=0}^{n-j}(-1)^i\binom{n-j}{i}\frac{x^i}{(i+1)(i+2)}. \qquad (4.13)$$

It may be noted that the polynomials $\psi_j(x)$ are linearly independent.

Theorem 35 depended on the convergence of $B_n(g;x)$ to $g''(x)$ on $[0,1]$. Although this convergence is slow, we can still hope to obtain a good (convex) approximation to $g(x)$ of the form

$$\Phi_n(x) = \sum_{j=0}^{n} c_j \psi_j(x)$$

because we still have the choice of the non-negative coefficients c_j at our disposal.

We use the least squares norm. Having made a linear change of variable so that $[a,b]$ is transformed onto $[0,1]$, we seek to minimise

$$\sum_{i=1}^{N} \left[f(x_i) - \sum_{j=0}^{n} c_j \psi_j(x_i) \right]^2, \tag{4.14}$$

subject to the constraints $c_j \geqslant 0$, $j=0, 1, \dots, n$. The data we are concerned with here is the set of points with co-ordinates $(x_i, f(x_i))$, for $i=1, 2, \dots, N$. The set of values $f(x_i)$ are regarded as perturbed values of some 'hidden' convex function $g(x)$ at the points $x=x_1, \dots, x_N$ on $[0,1]$. We take $n<N$. The expression (4.14) may be

written as a function of the (column) vector

$$c = \{ c_0, c_1, \ldots, c_n \}^T .$$

Then (4.14) becomes

$$\Psi(c) = K - v^T c + c^T M c , \qquad (4.15)$$

where $K = \sum_{i=1}^{N} [f(x_i)]^2$, v is a vector whose $(j+1)$ element is

$$v_{j+1} = 2 \sum_{i=1}^{N} f(x_i) \psi_j(x_i)$$

and M is a matrix whose $(j+1, k+1)$ element is

$$m_{j+1,k+1} = \sum_{i=1}^{N} \psi_j(x_i) \psi_k(x_i) .$$

Thus the problem is that of finding

$$\inf_{c \geq 0} \Psi(c), \qquad (4.16)$$

the infimum being over all vectors c with non-negative elements. From (4.15), this is a quadratic programming problem whose only constraints are the non-negativity constraints on the c_j. For methods of solving such a problem, see for example Hadley, 1964.

If A denotes the matrix whose $(i, j+1)$ element is $\psi_j(x_i)$, we have

$$M = A^T A$$

so that in (4.15)

$$c^T M c = (Ac)^T (Ac) > 0 \tag{4.17}$$

unless $c=0$. For if $c \neq 0$, $Ac \neq 0$ by the linear independence of the functions $\psi_j(x)$. Therefore the matrix M in (4.15) is positive definite. This entails that $\Psi(c)$ is a convex function of c, because with $0 < \lambda < 1$ and any two vectors $c^{(1)} \neq c^{(2)}$, the expression

$$\lambda \Psi(c^{(1)}) + (1-\lambda) \Psi(c^{(2)}) - \Psi(\lambda c^{(1)} + (1-\lambda) c^{(2)})$$

simplifies to give the inequality

$$\lambda(1-\lambda)\,(c^{(1)} - c^{(2)})^T \; M \; (c^{(1)} - c^{(2)}) \; > 0 . \qquad (4.18)$$

Thus any method which finds a local minimum of $\Psi(c)$ will have found a global minimum. This result is shown in Hadley, 1964, where it is also shown that an appropriate method for solving this problem is Wolfe's quadratic programming algorithm.

4.4 Numerical examples

In order to allow an objective assessment of this method for deriving convex approximations, a number of numerical experiments were performed. These begin by calculating sets of specimen 'nearly convex' data of the form

$$f(x_i) \;=\; g(x_i) \;+\; \delta . R_i \qquad (4.19)$$

where $g(x)$ denotes some convex function, the R_i are numbers in the range $[-1,1]$ produced by a random number generator, and δ is a scaling factor. Various choices were made of $g(x)$, δ, the R_i, the degree of the approximating polynomial n, and the number of data points.

A comparison was made between the 'best' convex approximation, say $\Phi(x)$, and the conventional least squares approximation of the same degree, say $Q(x)$. The numbers

$$E(\Phi^{(s)}) = \left[\sum_{i=1}^{N} [g^{(s)}(x_i) - \Phi^{(s)}(x_i)]^2 \right]^{\frac{1}{2}} \tag{4.20}$$

were calculated as a measure of the error in approximating to the s^{th} derivative of $g(x)$ by the s^{th} derivative of the convex polynomial $\Phi(x)$. These were compared with the corresponding numbers $E(Q^{(s)})$ obtained by estimating similarly the error in approximating $g^{(s)}(x)$ by $Q^{(s)}(x)$, the s^{th} derivative of the conventional least squares polynomial approximation of the same degree.

The accompanying tables show some typical results. Table 1 (page 98) was obtained for the function $g(x) = 1 - \sin \pi x$ on the eleven points x $0(0.1)1$, with

$n=5$ and $\delta = 0.2$. The table shows the results obtained for six sets of calculations performed using different sets of random numbers. The mean values of $E(\Phi^{(s)})$ and $E(Q^{(s)})$ for the six sets of values are shown in the last row of the table. Table 2 (page 99) shows the last set of results from Table 1 in more detail. Table 3 (page 100) gives results, as in Table 1, for the function $g(x) = 1/(1+x)$ and the same choice of the x_i, n and δ.

4.5 Discussion

These experiments suggest that the convex approximations, $\Phi(x)$, are advantageous in smoothing crude convex data, and are particularly useful in approximating to derivatives. This is borne out most strongly in the comparison of second derivatives in Table 2.

The use of the L_2 norm leads to a problem which is comparatively simple to solve. The amount of computation depends chiefly on the degree, n, of the approximation required, which specifies the size of the matrix M in (4.15), and not on the number of points, N, in the data.

One might consider using the L_1 norm instead of least squares. (The L_∞ or minimax norm is not usually recommended for the approximation of discrete data, since it takes undue regard of 'wild' points). The L_1 approximation problem leads to a linear programming problem, but in this case the size of the problem depends on the number of points, N. Algorithms for calculating L_1 approximations are given by Barrowdale and Young, 1966. However, the latter are concerned merely with polynomial (not convex polynomial) approximations. We could consider relaxing the convexity conditions on the component functions $\psi_j(x)$ and also relax the non-negativity constraints on the coefficients c_j. We could then solve the L_1 problem, with additional constraints, such as

$$\sum_{j=0}^{n} c_j \, \psi_j''(x_i) \geq 0, \qquad i = 1, \cdots, N$$

to try to impose convexity on the approximation. Even so, it does not seem that this will guarantee that we will always obtain a convex approximation. This approach does not seem worth pursuing.

Lastly, it may be noted that the method described here for least squares convex approximations applies equally to approximations on the interval $[0,1]$ as well as on a finite point set. For again we have a quadratic programming problem. This time, we have to minimise, for $c \geqslant 0$,

$$\Psi(c) = K - v^T c + c^T M c .$$

This time,

$$K = \int_0^1 [f(x)]^2 \, dx ,$$

the vector v has $(j+1)$ element

$$v_{j+1} = 2 \int_0^1 f(x) \, \psi_j(x) \, dx$$

and the matrix M has $(j+1, k+1)$ element

$$m_{j+1, k+1} = \int_0^1 \psi_j(x) \, \psi_k(x) \, dx .$$

<u>Table 1</u>

	$E(Q)$	$E(\Phi)$	$E(Q')$	$E(\Phi')$	$E(Q'')$	$E(\Phi'')$
	(function)		(1^{st} derivative)		(2^{nd} derivative)	
	0.007	0.006	0.090	0.058	1.48	0.34
	0.010	0.010	0.081	0.073	1.23	0.31
	0.008	0.007	0.109	0.042	1.56	0.26
	0.010	0.009	0.146	0.080	2.78	0.57
	0.010	0.010	0.099	0.100	1.02	0.98
	0.012	0.010	0.214	0.060	3.82	0.40
Mean	0.010	0.009	0.123	0.069	1.99	0.48

A comparison of convex approximations, $\Phi(x)$, with the conventional least squares approximations, $Q(x)$. The data was obtained by perturbing the function $1 - \sin \pi x$. (See page 94).

Table 2

x	$g(x)$	$Q(x)$	$\Phi(x)$	$g'(x)$	$Q'(x)$	$\Phi'(x)$	$g''(x)$	$Q''(x)$	$\Phi''(x)$
0.0	1.000	0.996	0.992	-3.14	-3.54	-3.16	0.00	7.33	0.53
0.1	0.691	0.675	0.684	-2.99	-2.91	-2.96	3.05	5.55	3.51
0.2	0.412	0.411	0.410	-2.54	-2.37	-2.48	5.80	5.57	5.94
0.3	0.191	0.203	0.195	-1.85	-1.76	-1.79	7.98	6.72	7.80
0.4	0.049	0.063	0.058	-0.97	-1.01	-0.94	9.39	8.35	9.01
0.5	0.000	0.007	0.010	0.00	-0.10	-0.01	9.87	9.77	9.53
0.6	0.049	0.047	0.057	0.97	0.92	0.94	9.39	10.34	9.31
0.7	0.191	0.189	0.196	1.85	1.92	1.83	7.98	9.38	8.29
0.8	0.412	0.424	0.417	2.54	2.72	2.57	5.80	6.23	6.44
0.9	0.691	0.718	0.702	2.99	3.07	3.08	3.05	0.22	3.69
1.0	1.000	1.012	1.023	3.14	2.65	3.28	0.00	-9.31	0.00

This table lists in more detail the approximations referred to in the last entry of Table 1. In the approximation of the second derivative, the last three columns of the table show that the convex approximation $\Phi(x)$ is very much superior to the conventional least squares approximation $Q(x)$.

Table 3

	$E(Q)$	$E(\Phi)$	$E(Q')$	$E(\Phi')$	$E(Q'')$	$E(\Phi'')$
	(function)		(1^{st} derivative)		(2^{nd} derivative)	
	0.007	0.005	0.080	0.036	1.31	0.17
	0.010	0.010	0.076	0.047	1.00	0.39
	0.008	0.007	0.098	0.058	1.30	0.69
	0.010	0.007	0.147	0.073	2.79	0.49
	0.010	0.008	0.110	0.076	1.26	0.62
	0.012	0.009	0.213	0.040	3.79	0.19
Mean	0.010	0.008	0.121	0.055	1.91	0.43

A comparison of convex approximations, $\Phi(x)$, with the conventional least squares approximations, $Q(x)$. The data was obtained by perturbing the function $1/(1+x)$. (Similar to Table 1).

REFERENCES

Ahlberg, J.H., Nilson, E.N. and Walsh, J.L. (1967). 'The Theory of Splines and Their Applications', Academic Press, New York.

Barrowdale, I. and Young, A. (1966). 'Algorithms for Best L_1 and L_∞ Linear Approximations on a Discrete Set', Numerische Mathematik, Vol. 8, p. 295.

Clenshaw, C.W. and Curtis, A.R. (1960). 'A method for numerical integration on an automatic computer', Numerische Mathematik, Vol. 2, p. 197.

Davis, P.J. (1963). 'Interpolation and Approximation', Blaisdell, New York.

Davis, P.J. and Rabinowitz, P. (1967). 'Numerical Integration', Blaisdell, New York.

Elliott, D.E. (1963). 'A Chebyshev method for the numerical solution of Fredholm integral equations', Computer Journal, Vol. 6, p. 102.

Filippi, S. (1964). 'Angenäherte Tschebyscheff-Approximation einer Stammfunktion - eine Modifikation der Verfahrens von Clenshaw und Curtis, Numerische Mathematik, Vol. 6, p. 320.

Goodwin, E.T. et al. (1960). 'Modern Computing Methods', H.M.S.O., London.

Hadley, G. (1964). 'Nonlinear and Dynamic Programming', Addison-Wesley, Reading, Massachusetts.

Hardy, G.H., Littlewood, J.E. and Pólya, G. (1934). 'Inequalities', Cambridge.

Meinardus, G. (1967). 'Approximation of Functions : Theory and Numerical Methods', Springer-Verlag, Berlin. (Translated from German).

Nikolskii, S.M. (1964). 'Quadrature Formulae', Hindustan Publishing Corporation (India), Dehli. (Translated from Russian).

Ream, N. (1961). 'Note on ''Approximation of Curves by Line Segments'' ', Mathematics of Computation, Vol. 15, p. 418.

Rice, J.R. (1964). 'The Approximation of Functions', Volume 1 : Linear Theory', Addison-Wesley, Reading, Massachusetts.

Stone H. (1961). 'Approximation of Curves by Line Segments', _Mathematics of Computation_, Vol. 15, p. 40.

Timan, A.F. (1963). _'Theory of Approximation of Functions of a Real Variable'_, Pergamon, Oxford. (Translated from Russian).

Wright, K. (1966). 'Series methods for integration', _Computer Journal_, Vol. 9, p. 191.

INDEX OF THEOREMS