1963

Tchebycheff approximation

BOSTON UNIVERSITY
GRADUATE SCHOOL

Thesis

TCHEBYCHEFF APPROXIMATION

by

JOSEPH W. GAULD
(A. B. Bowdoin College, 1951)

Submitted in partial fulfillment of the requirements for the degree of

Master of Arts
1963

Approved
by

First Reader . . . . Francis Scheid
Professor of

Second Reader . . . . Donald W. Blackett
Professor of

TCHEBYCHEFF APPROXIMATION

I. *Tchebycheff Approximation of Continuous Functions* 1

A. *Definition of Tchebycheff Approximation* 1

B. *Tchebycheff Polynomials* 2

C. *Theorems and Proofs* 4

D. *Examples of Application* 7

1. *Application of Alternating Signs Theorem* 7

2. *Power Series Economization* 9

3. *Lanczos Tau Method* 12

E. *Solutions of Examples* . . . . 14 –20

II. *Tchebycheff Approximation of Overdetermined Systems of Linear Equations* 21

A. *Relationship of the overdetermined problem to the discrete point problem* 21

B. *Theorem and Proof* 23

C. *Examples of Applications* 29

D. *Solutions of Examples* 30- 32

III. *Tchebycheff Approximation of Discrete Point Sets by a Polynomial of Degree k* 33

A. *Theorem and Proof* 33

B. *Examples of Application* 36

C. *Solutions of Examples* 38

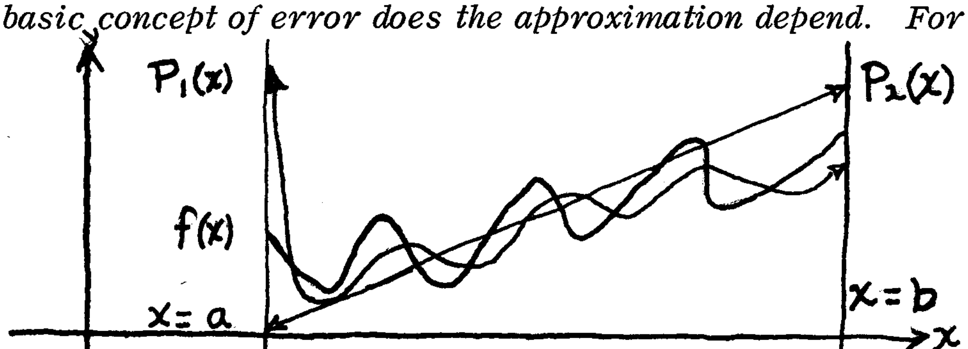
IV. *Equivalence of the Problem of Linear Programming to the System of Overdetermined Equations* 44

TCHEBYCHEFF APPROXIMATION

I. *Tchebycheff Approximation to continuous functions*

A. *Definition of Approximation*

In approximation theory the question immediately arises "What do we consider a good approximation?" This question is asked, not only from the standpoint of the restriction of error, but also upon what basic concept of error does the approximation depend. For example:



$P_1(x)$ resembles $f(x)$ more than $P_2(x)$ for the most part, and certainly the area between the two curves seems to be less; however, $P_2(x)$ has the advantage over $P_1(x)$ of being closer to $f(x)$ over the entire domain. In other words, it would seem that the maximum of $|P_2(x) - f(x)|$ in $a \leq x \leq b$ is less than the maximum of $|P_1(x) - f(x)|$. In this paper, that will be our basic requirement of approximation, we will try to minimize the maximum error of $|P(x) - f(x)|$ in $a \leq x \leq b$. In general, polynomials are better known

and as a result usually easier to deal with than other functions, so that we will assume that $P_n(x)$*, the approximation, is always a polynomial of the form*

$$a_0 x^n + a_1 x^{n-1} + \ldots\ldots + a_n x + a_n .$$

B. Tchebycheff Polynomials.

One of the most useful set of polynomials in approximation is known as the Tchebycheff polynomial.

$$T_n(x) \quad (n=0, 1, 2. \ldots\ldots)$$

While they can be defined in many ways, depending upon the problem, we will use the following definition:

$$T_n(x) = \cos(n \cos^{-1} x) \text{ when } |x| \leq 1.$$

It has the following properties:

1) *it is a polynomial of degree n.*
2) *It is odd or even in x according as n is odd or even.*
3) *It obeys the recursion relation* $T_{n+1} = 2x\, T_n - T_{n-1}$.
4) *For* $n > 0$ *its leading term is* $2^{n-1} x^n$.

Since $|x| \leq 1$ *we can let* $\cos\theta = x$ $\;0 \leq \theta \leq \pi$ *then* $T_n(x) = \cos n\theta$.

Thus:

$$T_0 = \cos 0 = 1$$

$$T_1 = \cos\theta = x$$

$$T_2 = \cos 2\theta = 2\cos^2\theta - 1 = 2x^2 - 1.$$

From here we can use the recursion relation to build the set. Further values are:

$T_3 = 4x^3 - 3x$

$T_4 = 8x^4 - 8x^2 + 1$

$T_5 = 16x^5 - 20x^3 + 5x$

$T_6 = 32x^6 - 48x^4 + 18x^2 - 1$ *and so on.*

Since $T_n(x) = \cos n\theta$, $0 \leq \theta \leq \pi$, $T_n(x)$ *has* n *roots at* $\theta_j = (2_j - 1/_{2n})$ $j = 1, 2, \ldots n$. *In addition, from* $T_n = \cos n\theta$, *we learn that* $|T_n| \leq 1$ *in* $-1 \leq x \leq 1$ *and attains its maximum* $n + 1$ *times at* $e_j\ (T_n) = \cos {}^{j\pi}/n$ $j = 0, 1 \ldots n$ *at which points* $T_n(e_j) = (-1)^j$. *No other polynomials have this property of attaining their maximum magnitude* $n + 1$ *times in* $-1 \leq x \leq +1$ *so that any* $P_n(x)$ *that does must be a multiple of* $T_n(x)$.

We now return to our definition of approximation and apply it to Tchebycheff polynomials. Let $\{\phi^\wedge(x)\}$ *represent a set of functions, say* $\{\phi^\wedge(x)\} = \{P(x)\}$ *is the set of all polynomials of degree n or less (in keeping with our note in section A). The approximation which is optimum in the Tchebycheff sense to f(x) with respect to the set* $\{\phi(x)\}$ *is that member* $\phi^*(x)$ *for which:*

$E = \max|f(x) - \phi(x)|$ $-1 \leq x \leq +1$ *is minimized.*

($\phi^*(x)$ *will be used in this sense throughout.)*

While we have used the interval $-1 \leq x \leq 1$, *any finite interval can be mapped on to this interval so there is no loss in generality.*

C. *Theorems and Proofs.* [11]

The door opener to our concept on approximation is a theorem of Weierstrass.

Theorem A. If f(x) is continuous in $a \leq x \leq b$ *then given any* $e > 0$ *there is a polynomial* $P = Pe(x)$ *such that:*

$$/f(x) = P(x)/ \leq e \qquad a \leq x \leq b$$

(There is a beautiful proof of this Theorem by Bernstein that[11] *uses a consideration of probability without depending on probability ideas.*

The optimum approximation to a continuous f(x) with respect to polynomials of a prespecified degree n or less has an error E(x) which actually achieves its maximum magnitude at least $n + 2$ *in* $a \leq x \leq b$ *and furthermore does so in an oscillatory manner. In addition, no other polynomial has this property.*

This statement allows us to recognize whether a given $\emptyset$ *is* $\emptyset^*$; *if it achieves its error less than* $n + 2$ *times it is not* $\emptyset^*$, *if it achieves its maximum error* $n + 2$ *or more times it is* $\emptyset^*$ *and is unique. Any change in its coefficients increases its maximum error.*

This idea will be used throughout our examples on Tchebycheff approximation to continuous functions, so it would be wise to give it some justification. We will assume the following two theorems:

Theorem B. *If $f(x)$ is continuous in $a \le x \le b$, then for each n there exists a polynomial of degree $\le n$ of best approximation.*

Theorem B guarantees a best approximation to $f(x)$ regardless of what n we choose for $\phi_n^*(x)$; however, the degree of $\phi_n^*(x)$ may turn out to be less than n.

Now define E equal to the minimum of the maximum (abbreviated minimax) of $|f(x) - \phi(x)|$; that is, $E = \max |f(x) - \phi^*(x)|$,

Theorem C. *If $\phi_n^*(x)$ is the best approximation to $f(x)$, then there are at least $n + 2$ E points which are alternately + and - .*

Our problem in applications will deal more with the converse of this theorem.

Theorem I. *Given $\phi_n(x)$, if $E = \max\ f(x) - \phi_n(x)$ is achieved at least $n + 2$ times alternately + and - , then $\phi_n(x) = \phi_n^*(x)$.*

Theorem II. *$\phi_n^*(x)$ is unique.*

We prove Theorem II first.

Proof: Suppose there were two $\phi_n^*(x)$'s, say ϕ', ϕ'', then we would have

$$-E \le f - \phi' \le +E \qquad E > 0$$
$$-E \le f - \phi'' \le +E.$$

Construct a third polynomial $\phi''' = \frac{1}{2}(\phi' + \phi'')$ which would also have to be a polynomial of best fit since $-E \le f - \phi''' \le E$.

Consider a minus point at which $f - \phi = -E$. Then at this point:

$$f - \phi''' = -E$$
$$f - [\tfrac{1}{2}\phi' + \phi''] = -E$$
$$2f - \phi' - \phi'' = -2E$$
$$f - \phi' + f - \phi'' = -2E.$$

Since $|f - \phi'| \le E$ and $|f - \phi''| \le E$ the last statement can only be true if $f - \phi' = -E$ and $f - \phi'' = -E$, which means $\phi' = \phi''$ at this point. We can give the same argument at a plus point. But by Theorem B there are $n + 2$ such points; hence $\phi' = \phi''$ at $n + 2$ points so they must be the same polynomial. Q. E. D.

Proof of Theorem I.

We are given a $\phi(x)$ such that $\max |f(x) - \phi(x)| = M$ is achieved $n + 2$ times. We know that $M \ge E$ $(E = \mathrm{Max}|f(x) - \phi^*(x)|)$ and since we want to prove $M = E$ we will assume $M > E$.

Then $\phi_n(x)$ is not the best approximation.

Assume $\emptyset'_n(x)$ *is (Theorem B), and is unique (Theorem I).*

Then $\emptyset' - \emptyset = \emptyset' - f + f - \emptyset$.

Since $|\emptyset' - f| \leq E$ *and* $|f - \emptyset| > E$ *then the signs of* $\emptyset$ *and* $\emptyset'$ *coincide at the* $n + 2$ *extrema. (At the extrema* $|f - \emptyset|$ *will always be larger than* $|\emptyset' - f|$ *and hence determine the sign.) Therefore the polynomial* $\emptyset' - \emptyset$ *will have* $n + 2$ *sign changes and this means there are* $n + 1$ *roots. Since* $\emptyset' - \emptyset$ *is at most degree n it must vanish identically. This gives:*

$$M = Max|f - \emptyset| = Max|f - \emptyset'| = E.$$

This contradicts our assumption that $M > E$, *therefore* $M = E$ *and* $\emptyset = \emptyset^*$. *Q. E. D.*

D. Examples of Applications

1. Application of Alternating Signs Theory.

To see the Tchebycheff polynomials in action, let us consider the simple problem of approximating $f(x) = x^n$ *by a polynomial of degree n or less. The solution is*

$$\emptyset^*(x) = x^n - \frac{T_n(x)}{2^{n-1}},$$ *and since* $T_n / 2^{n-1}$ *has leading coefficient unity* $\emptyset^*(x)$ *is a polynomial of degree* $n - 2$.

While we did not derive this solution, it is apparently correct by observing in the diagrams of page 14 that the maximum *error is achieved the required number of times.*

Example 1.

$$f(x) = x^2 \quad \phi^*_2(x) = x^2 - \left(\frac{2x^2 - 1}{2}\right) = 1/2.$$

$$f(x) = x^3 \quad \phi^*_3(x) = x^3 - \left(\frac{4x^3 - 3x}{4}\right) = 3/4 \; x.$$

$$f(x) = x^4 \quad \phi^*_4(x) = x^4 - \left(\frac{8x^4 - 8x^2 + 1}{8}\right) = x^2 - 1/8.$$

$$f(x) = x^5 \quad \phi^*_5(x) = x^5 - \left(\frac{16x^5 - 20x^3 + 5x}{16}\right) = 1/16\,(20x^3 - 5x).$$

Diagram on Page 14

Example 2.

Let us go on to approximate the sine curve from 0 to $\pi/2$ *by a straight line. While we are not positive that our diagram is correct we have no worries, for if the maximum error is achieved three times or more we have found* $\phi^*(x)$ *by Theorem II; if not we try again.*

Solution on page 15

Example 3.

Since the cosine curve is parabolic in nature, it might be interesting to approximate it by a parabola in the interval $-\pi/2$ *to* $+\pi/2$ *to see how much variation there is between the two curves. We run into a problem as the maximum error must* be achieved at least four times in an oscillatory fashion, and the symmetry of the cosine curve indicates that this has to be done an odd number of times. Therefore we assume it is achieved five times and proceed with the confidence explained in Example 2. Note in this problem we were unable to find a closed formula for ϕ^* as we did in Example 1. In finding the derivative, we had to approximate the sine function.

Solution on page 16-17 ,

2. Power Series Economization. [4]

When f(x), the function to be approximated, is an arbitrary continuous function, in general no closed formulas are known for $\phi^*(x)$. It is true that f(x) can be expanded in a Tchebycheff series, i.e.

$$f(x) = \sum_{k=0}^{\infty} C_k T_k(x) \quad -1 \leq x \leq +1.$$

Suppose we truncate the Tchebycheff series after the nth term so that

$$f(x) = \sum_{k=0}^{n} C_k T_k(x) + E(x), \quad \text{where}$$

$$E_n(x) = \sum_{k=n+1}^{\infty} C_k T_k(x).$$

When the series is rapidly converging, the error E(x) will be approximately the first term $C_{n+1}\ T_{n+1}(x)$. The question arises, can such a truncation be used as a useful approximation to $\phi^*(x)$?

Consider the function

$$f(x) = 1 - x + x^2 + x^3 - x^4 + x^5 - x^6, \qquad -1 \leq x \leq +1.$$

At first glance approximating $f(x)$ looks quite discouraging. If we drop x^6, we would be commiting a very large error, at least as x neared the ends of the interval. However, returning to the Tchebycheff polynomials we note:

$$T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1.$$

Solving for x^6

$$x^6 = \frac{48x^4 - 18x^2 + 1}{32} + \frac{T_6(x)}{32}$$

This is a rather remarkable result. Though x^6 is independent of the lower powers and is not a linear combination of them, it is nearly so in the range we have chosen as

$$T_6(x)/_{32} \leq 1/32. \qquad -1 \leq x \leq 1$$

Suppose further we replace x^4 and x^2 in a similar fashion and keep doing this until the right hand side is devoid of powers of x, then we would have $x^6 = \sum_{k=1}^{6} a_k\ T_k(x)$. Continuing the process for other powers of x, we would then have a method for expressing a polynomial $\sum_{k=0}^{n} a_k x^k$ in terms of $\sum_{k=0}^{n} b_k\ T_k(x)$.

Since we have an idea of the accuracy we want, we would know where to truncate the series $\sum_{k=0}^{n} b_k T_k(x)$. *The first few translations of* x^n *look like this:*

$$1 = T_0$$

$$x = T_1$$

$$x^2 = 1/2\,(T_0 + T_2)$$

$$x^3 = 3T_1 + T_3$$

$$x^4 = 3T_0 + 4T_2 + T_4$$

This method is known as power series economization.

Example 4.

Find the best parabolic fit to $y = f(x) = e^x$.

Since e^x *is polynomial of infinite order we will have to truncate* $T_n(x)$ *using power series economization.*

Solution on page 18

Example 5 .

Approximate $y = \cos x$ $\; -\frac{\pi}{2} \le x \le \frac{\pi}{2}$ *using power series economization.*

This will give us a good opportunity to observe the error in economization, as we have already solved this problem in pure (at least in theory) fashion.

Comparing the two results:

$$\phi_1^*(x) = -.405x^2 + .97 \qquad E = .07$$

$$\phi_2^*(x) = -.47x^2 + .99 \qquad E = .03$$

We see that economization has done a remarkably good job for such an early truncation.

Solution on page 19

[4]

3. *The Lanczos Tau Method.*

Since we are discussing truncation of series, we might take a side trip and examine the Lanczos Tau Method which applies to the solution of differential equations. Consider a function that is defined by a linear differential equation with rational coefficients. Generally this problem yields a solution that is a polynomial of infinite order. Briefly, the Tau method then is to assume a truncated series as a solution. We set the differential equation to some multiple of a Tchebycheff polynomial of our choosing, and then, by equating coefficients, evaluate the coefficients of our assumed solution. The multiple will give us an idea of the error involved as $T_n(x)$ *is a bounded function.*

Example 6.

Approximate the solution of $xy'' + y = 0$, subject to initial conditions $y(0) = 0$, $y'(0) = 1$, $0 \leq x \leq 1$.

Since the range is $0 \leq x \leq 1$, we will need to normalize $T_n(x)$ from $-1 \leq x \leq 1$ to $0 \leq x \leq 1$. To do this is quite easy. We put $\cos\theta = 2x - 1$ so $x = \frac{\cos\theta + 1}{2} = \cos^2\frac{\theta}{2}$. Now as θ varies from 0 to π, x varies from 0 to 1. We again define these Tchebycheff polynomials by $T_n^*(x) = \cos n\theta$ so that

$T_n^*(x) = T_n(2x - 1)$.

With this relationship we again build a table. The first few are:

$$T_0^*(x) = 1$$

$$T_1^*(x) = -1 + 2x$$

$$T_2^*(x) = 1 - 8x + 8x^2$$

$$T_3^*(x) = -1 + 18x - 48x^2 + 32x^3$$

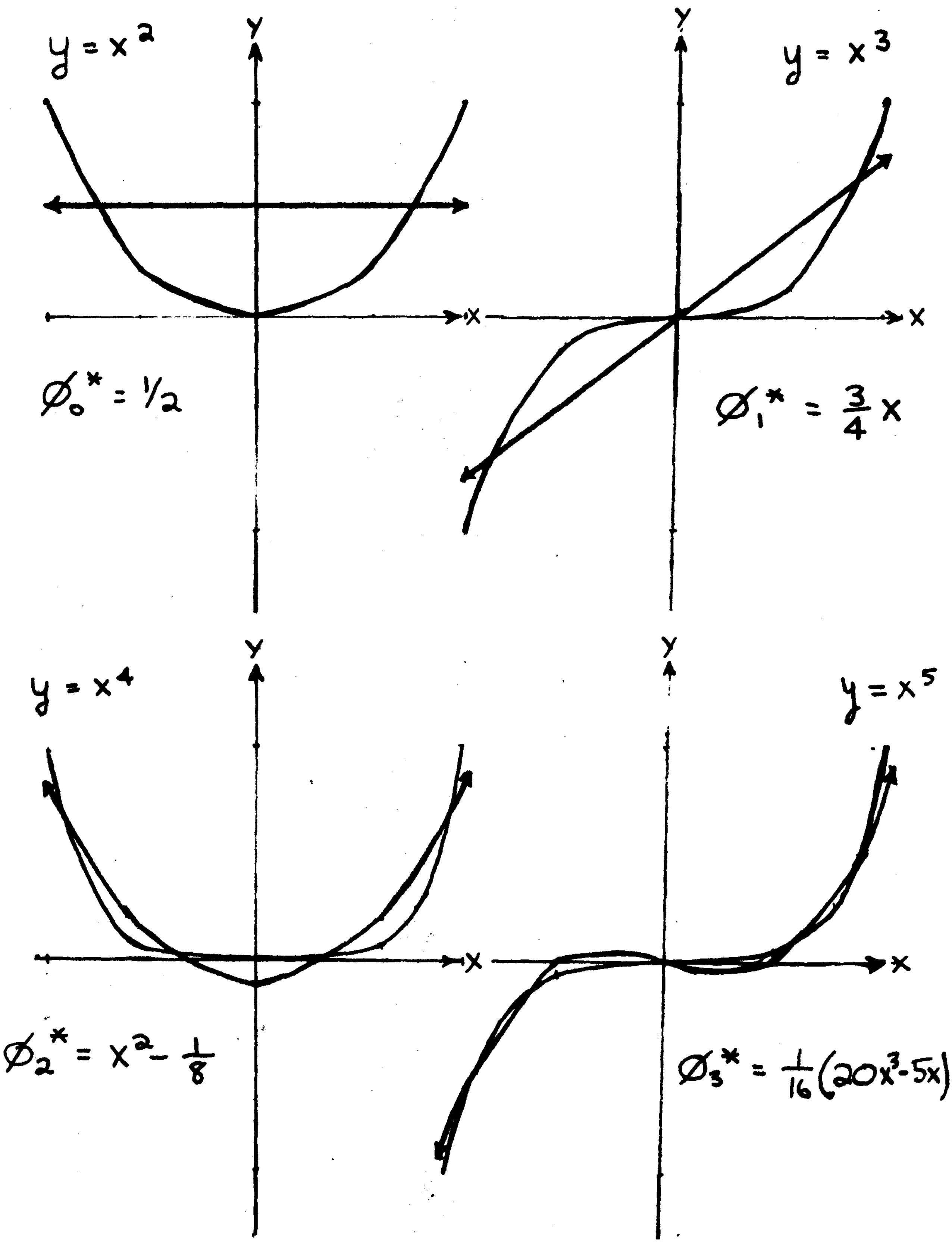
$$T_4^*(x) = 1 - 32x + 160x^2 - 256x^3 + 128x^4$$

These are known as the "shifted" Tchebycheff polynomials.

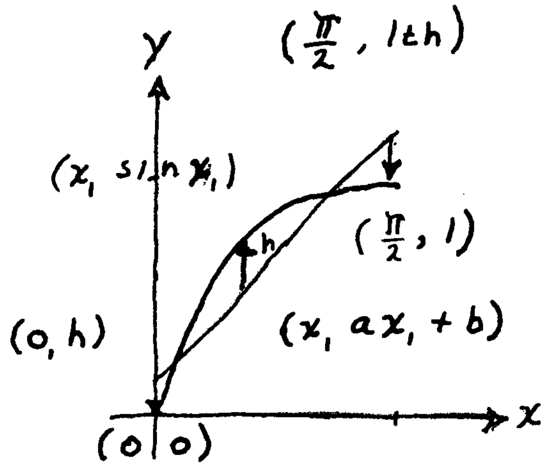
Solution on page 20

This ends our brief survey of methods of Tchebycheff approximation to a continuous $f(x)$. We now turn our attention to approximation regarding discrete point sets.

Example 1



__Example 2:__ *Find the best linear approximation to $y = f(x) = \sin x$ from $x=0$ to $x=\pi/2$.*



Solution: Let $\phi^*(x) = ax + b$

$E = \max \left| f(x) - \phi^*(x) \right|$ has

$n+2=3$ extrema and $n+1=2$ roots

$$1+h = a \cdot \frac{\pi}{2} + b$$

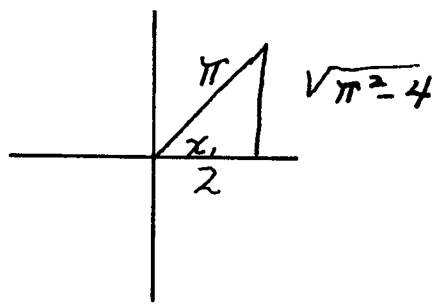
$$\frac{h = \qquad b}{1 = a \cdot \frac{\pi}{2}}$$

$$\frac{2}{\pi} = a$$

at x_1, $\left| \phi^*(x) - f(x) \right|$ is a maximum

$$h = \sin x_1 - (a x_1 + b)$$

$$\frac{dh}{dx_1} = \cos x_1 - a = 0 \qquad \cos x_1 = a = \frac{2}{\pi}$$



$$b = h = \frac{\sqrt{\pi^2-4}}{\pi} - \frac{2}{\pi} \cos^{-1} \frac{2}{\pi} - h$$

$$b = h = \frac{\sqrt{\pi^2-4}}{2\pi} - \frac{1}{\pi} \cos^{-1} \frac{2}{\pi}$$

$$\therefore \phi^*(x) = \frac{2}{\pi} x + \frac{\sqrt{\pi^2-4}}{2\pi} - \frac{1}{\pi} \cos^{-1}\left(\frac{2}{\pi}\right)$$

Example 3: *Find the best quadratic approximation to*

$y = f(x) = \cos x \quad |x| \leq \pi/2$

Solution: We guess there are five extrema and four roots to the equation

$E(x) = \phi^*(x) - f(x)$

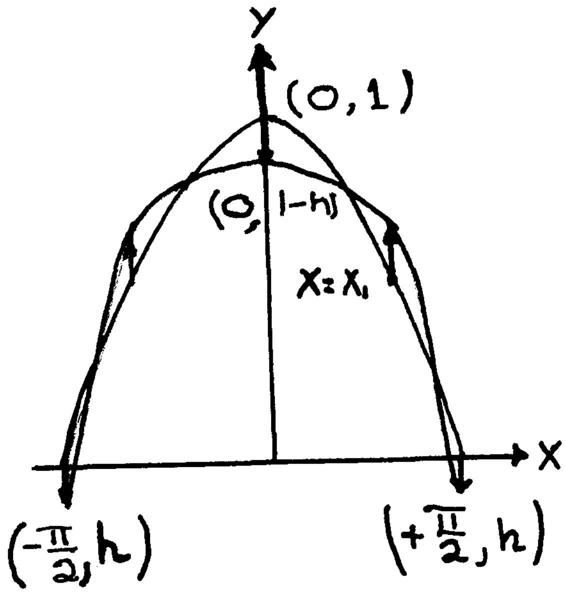
Let $\phi^*(x) = ax^2 + bx + c$

at $(0,\ 1-h)$ $\quad 1-h=c$

at $(\pm \frac{\pi}{2},\ -h)$ $\quad -h = \frac{\pi^2}{4}a + \frac{\pi}{2}b + c$

$$-h = \frac{\pi^2}{4}a + \frac{\pi}{2}b + c$$

$$b = 0 \quad a = -\frac{4}{\pi^2}$$



$\phi^*(x_1) = \cos x_1 + h$

$-\frac{\pi}{4}x_1^2 + 1 - h = \cos x_1 + h$

$2h = -\frac{4}{\pi^2}x_1^2 = \cos x_1 + 1$

$2\frac{dh}{dx_1} = \sin x_1 - \frac{8}{\pi^2}x_1$

at x_1 $\frac{dh}{dx_1} = 0$

Since $\sin x_1 \cong x_1 - \frac{x_1^3}{6}$ *

$x_1 - \frac{x_1^3}{6} - \frac{8}{\pi^2}x_1 = 0$

$x_1 = \pm\sqrt{6(1 - 8/\pi^2)}$

$h = -\frac{2}{\pi^2}x_1^2 - \frac{\cos x_1}{2} + \frac{1}{2}$

$c = 1-h = \frac{1}{2} + \frac{2}{\pi^2}x_1^2 + \frac{\cos x_1}{2}$

Hence $\phi^*(x) = -\frac{\pi^2}{4}x^2 + \left(\frac{1}{2} + \frac{2}{\pi^2}x_1^2 + \frac{\cos x_1}{2}\right)$

or $\phi^*(x) = -.405x^2 + .97$

**Because of this approximation, our answer has an error*

$|e| \quad x^5/5! \leq 1/120$

(Completed next page)

This value for ϕ^* is valid <u>only if</u> the five extrema to $\phi^*(x) - f(x)$ are equal with alternating signs.

at $x = 0$ $\quad \phi^* - f(x) = +.03$

at $x = \pm 1.07$ $\quad \phi^* - f(x) = -.03$

at $x = \pm \pi/2$ $\quad \phi^* - f(x) = +.03$ Q. E. D.

Example 4. *Approximate e^x in $-1 \le x \le +1$ by power series economization.*

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots\ldots .$$

$$= T_0 + T_1 + \frac{\frac{1}{2}(T_0 + T_2)}{2} + \frac{3T_1 + T_3}{4 \cdot 6} + \frac{3T_0 + 4T_2 + T_4}{24 \cdot 8} \quad *$$

$$= \frac{243}{192} T_0 + \frac{27}{24} T_1 + \frac{52}{192} T_2 + \frac{3}{24} T_3 + \frac{1}{192} T_4$$

Dropping T_3 and T_4 the error $< 3/24 + 1/192 = 25/192$ since $|T_k(x)| \le 1$. In addition we see our error is heaviest at $|x| = 1$.

converting back:

$$\frac{243}{192}(1) + \frac{27}{24}(x) + \frac{52}{192}(2x^2 - 1)$$

$$\phi^*(x) = \frac{1}{192}(104x^2 + 216x + 191)$$

at $x = -1$ $\quad \phi^* - f(x) = .07$

at $x = +1$ $\quad \phi^* - f(x) = -.15$

**In pure form, the truncation should not be made here as the next term involves T_1, T_3, T_5, and we eventually use T_1, while T_3 helps determine the error. However in this particular example the next denominator is $5! \cdot 16$ so that this term would contribute nothing to our level of approximation.*

<u>Example 5.</u> *Using Tchebycheff Polynomials Approximation*

$f(x) = y = \cos x, \qquad |x| \leq \frac{\pi}{2}$ *with power series economization.*

Given $y = \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \ldots\ldots\ldots \qquad |x| \leq \frac{\pi}{2}$

To use Tchebycheff Polynomials we must normalize the range from

$-1 \leq u \leq +1 \quad \therefore \quad u = 2/\pi\, x$

$$y = \cos u = 1 - \frac{\pi^2}{8} u^2 + \frac{\pi^4}{388} u^4 \qquad |u| \leq 1.$$

$$C.P. = \emptyset(u) = T_0 - \frac{\pi^2}{8} \cdot \frac{1}{2} (T_0 + T_2) + \frac{\pi^4}{388} \cdot \frac{1}{8} (3T_0 + 4T_2 + T_4) \quad *$$

$$= T_0 \left(1 - \frac{\pi^2}{16} + \frac{3\pi^4}{388.8}\right) + \left(\frac{4\pi^4}{388.8} - \frac{\pi^2}{16}\right) T_2 + \frac{\pi^4 T_4}{388.8}$$

the last term is less than .01.

$$= .41\, T_0 - .58\, T_2$$

converting back

$$\emptyset(u) = .41 - .58\,(2u^2 - 1) = .41 - 1.16u^2 + .58$$

$$= .99 - 1.16u^2 \quad |u| \leq 1$$

replacing $u = \frac{2}{\pi} x$

$$= .99 = .47x^2$$

x	$\emptyset(x)$	$f(x)$	e
0	.99	1	.01
± 1	.52	.54	.02
$\pm \frac{\pi}{2}$	-.07	0	.07

In keeping with Tchebycheff Polynomials - the error is smallest near the middle of the interval and heaviest at the end points.

**again we use the truncation explained in the note on example 4.*

Example 6. Use the Lanczos Tau Method to $xy' + y = 0$

$y(0) = 0 \quad y'(0) = 1$ for $0 \leq x \leq 1$ using $n = 4$.

Let $xy'' + y = tT_4^*(x)$ *

Assume $y = a + bx + cx^2 + dx^3 + ex^4$

then $2cx + 6dx^2 + 12\,ex^3 + y \equiv t\,(128x^4 + 256x^3 + 160x^2 - 32x + 1)$

$y(0) = 0 \;\therefore\; \boxed{a = 0}$ $\qquad$ $y'(0) = 1 \;\therefore\; \boxed{b = 1}$

$b + 2c = -32t$ $\qquad$ $c = (-32t - 1)/2$

$c + 6d = 160t$ $\qquad$ $d = (352t + 1)/12$

$d + 12e = -256t$ $\qquad$ $e = (-3072t - 1)/144$

$e = 128t$

$144 \cdot 128t = -3072t - 1 \qquad 21504t = -1 \qquad t = -1/21504$

$$\therefore\; y = x + \frac{\left(\frac{32}{21504} - 1\right)}{2} x^2 + \frac{\left(-\frac{352}{21504} + 1\right)}{12} x^3 - \frac{128}{21504} x^4$$

$$y = x - \frac{21472}{43008} x^2 - \frac{21152}{258048} x^3 - \frac{128}{21504} x^4$$

with an error $\dfrac{T_4^*(x)}{21504} \leq \dfrac{1}{21504}$

with regard to the original differential equation.

*Here again we use the "shifted" Tchebycheff Polynomial because of the interval $0 \leq x \leq 1$.

II. Tchebycheff Solution of an Overdetermined System of Linear Equations.

A. Relationship of the overdetermined problem to the discrete point problem.

As we mentioned in the first section, in general there is no known finite method of finding $\emptyset^*(x)$ for any continuous $f(x)$ $a \leq x \leq b$. It would seem logical then to replace the interval with a discrete set of n points $a \leq x_1 < x_2 \ldots\ldots < x_n \leq b$ where $y_k = f(x_k)$ $(k = 1, 2, \ldots. n)$ and then find the best polynomial approximation $\emptyset^*_n(x)$ for the n set of points. As a matter of fact, for suitable selections of the n points, it can be shown that $\emptyset^*_n(x)$ tends uniformly to $\emptyset^*(x)$ of $f(x)$ as $n \rightarrow \infty$*.

Let us analyze the discrete problem further through a particular example.

Assume we have five points $(x_1\ y_1)$ $(x_2\ y_2) \ldots. (x_5\ y_5)$ that we wish to approximate with a quadratic $ax^2 + bx + c = y$ of best fit. Therefore the points will define five equations in three unknowns:

$$
\begin{aligned}
ax_1^2 + bx_1 + c &= y_1 \\
ax_2^2 + bx_2 + c &= y_2 \\
&\vdots \\
ax_5^2 + bx_5 + c &= y_5
\end{aligned}
$$

* J. R. Rice: On the convergency of an algorithm for best Tchebycheff approximations. JS. IA. M. 7 pp 133-142 (1959)

Certainly there is no solution to this system (assuming the points are distinct.) The best we can hope for is the set (a^*, b^*, c^*) that tends to level and minimize the errors in all five equations. We will define a system of n equations in m unknowns with $n>m$ as an overdetermined system of linear equations. Clearly we can generalize this special case to say that if we have n distinct points $(x_n y_n)$ in a plane such that $x_1 < x_2 < \ldots\ldots < x_n$ and wish to find the best fit with a polynomial of degree m, then we define an overdetermined system of n equations in m unknowns; that is

$$\sum_{j=1}^{m} a_{ij}\, x_j + b_i = 0 \qquad (i = 1, 2, \ldots n).$$

If we study the known coefficients in our particular example we see they are inter-related (i.e. coefficient a = (coefficient b)^2, (coefficient c = 1). so that the point problem is a subset of overdetermined systems in general. With this thought in mind we will turn our attention to the study of the overdetermined system, for certainly anything we prove would be true for the point problem.

There are basically three questions we must answer:

1) *Does every overdetermined system have a Tchebycheff solution?*
2) *Is it unique?*
3) *How do we find it?*

B. *Theorems and Proofs.* [7]

Our basic problem is this: Given $\sum_{j=1}^{m} a_{ij} x_j + b_i = 0 \; (i=1, \ldots n)$

If $m < n$ *and if the* n *equations are linearly independent, we have no solution. We set*

$$h_i(x) = \sum_{j=1}^{m} a_{ij} x_j + b_i = 0 \quad (i = 1 \ldots n) \qquad \text{where}$$

$x = (x_1, x_2 \ldots x_m)$. *Then we wish to find an* $x = x^*$ *which minimizes*

$h(x^*) = \max h_i(x) \quad i = 1, 2, \ldots n$

In order to prove the existence of $h(x^*)$, *we will have to assume that every* m *by* m *matrix is non singular. Later we will examine this restriction, at least in the discrete problem, and get some idea of what happens in the more general case.*

In order to have a clearer, more continuous paper we will prove the hypothesis with only two variables as most of our analysis will deal with the straight line problem. There should be no difficulty with the general case if the reader remembers we have put $m = 2$.

Theorem III. *There exists a unique Tchebycheff Solution to every* n *set of overdetermined equations in two unknowns, if every* 2×2 *matrix* $= 0$.

Proof: First we choose three of the n equations, with no loss in generality by choosing the first three since they are not ordered. Given:

$x = (x_1, x_2)$ we have:

$$a_{11} x_1 + a_{12} x_2 + b_1 - h_1 (x) = 0$$
$$a_{21} x_1 + a_{22} x_2 + b_2 - h_2 (x) = 0$$
$$a_{31} x_1 + a_{32} x_2 + b_3 - h_3 (x) = 0$$

We say that $(x_1, x_2, 1)$ is a solution (non-zero by assumption) of the three equations which can only be if the determinants

$$\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & h_1(x) \\ a_{21} & a_{22} & h_2(x) \\ a_{31} & a_{32} & h_3(x) \end{vmatrix}$$

Expanding these by elements of the last column, and denoting the cofactors by $B_1, B_2, B_3 \neq 0$ we have:

$$B_1 b_1 + B_2 b_2 + B_3 b_3 = B_1 h_1 (x) + B_2 h_2 (x) + B_3 h_3 (x).$$

Define C equal to both of these.

We now use a Lemma by de La Vallée-Poussin*.

If $c_1 u_1 + c_2 u_2 + c_3 u_3 = c$

u_i $(i = 1, 2, 3)$ variable $c_1, c_2, c_3 \neq 0$

then the maximum of $|u_i|$ $(i = 1, 2, 3,)$ has a minimal

*we have taken $(i = 1, 2, 3)$ to fit our particular case; the Lemma is actually true for $(i = 1, 2, \ldots n)$.

value $$p = \frac{|c|}{c_1 + c_2 + c_3} \tag{2}$$

which is attained for just one set (u_1^*, u_2^*, u_3^*) of the u_i's (i=1, 2, 3) given by

$$u_i = u_i^* = p \text{ sgn } (c_i c) \qquad i = 1, 2, 3 \tag{3}$$

("sgn" means sign of).

Briefly the lemma says that if we allow for free variation of the u_i's in (1), the the minimax of the u_i's occurs for only one set — when they are all equal in absolute value.

Proof: First notice that the u_i's in (3) does indeed satisfy (1).

$$c_1 \frac{|c| \text{ sgn } (c\, c_1)}{|c_1| + |c_2| + |c_3|} + c_2 \frac{|c| \text{ sgn } (c\, c_2)}{|c_1| + |c_2| + |c_3|} + c_3 \frac{|c| \text{ sgn } c\, c_3}{|c_1| + |c_2| + |c_3|} = c$$

by factoring the left hand side and observing that w sgn w is positive.

Next suppose there is a set of u_i's satisfying (1) such that $\max_{i=1,2,3} |u_i| \leq p$

then for $|d_i| \leq 1$ $\quad u_i = d_i p \qquad i = 1, 2, 3.$

Substituting in (1) we get:

$$c_1 d_1 p + c_2 d_2 p + c_3 d_3 p = c \tag{4}$$

but $c = \text{sgn } c \cdot |c|$ and using (2) to substitute for $|c|$

$$\text{sgn } c \cdot |c| = \text{sgn } c \left[p\, |c_1| + p\, |c_2| + p\, |c_3| \right] . \tag{5}$$

Equating (4) and (5) and dividing by p we get

$$c_1 d_1 + c_2 d_2 + c_3 d_3 = \text{sgn } c \; (|c_1| + |c_2| + |c_3|).$$

This can hold only if $d_i = \text{sgn } c \cdot \text{sgn } c_i = \text{sgn }(c\, c_i)$ (by equating coefficients); therefore $u_i = u_i^*$ $(i = 1, 2, 3)$ uniquely.

Now return to

$$B_1 b_1 + B_2 b_2 + B_3 b_3 = B_1 h_1(x) + B_2 h_2(x) + B_3 h_3(x) = c. \tag{6}$$

Using $c_i = B_i$ in the Lemma, it tells us that the minimax of $|u_i|$ $(i = 1, 2, 3)$ where

$B_1 u_1 + B_2 u_2 + B_3 u_3 = c$, $(u_i = b_i)$, is attained for $u_i = u_i^* = p \text{ sgn }(B_i c)$ $(i = 1, 2, 3)$ where p has the value (2).

Therefore we will have proved that x^* is the Tchebycheff Solution to the three equations if we can show that there is an x^* such that $h_i(x^*) = u_i^* = p \text{ sgn }(B_i c)$ $i = (1, 2, 3)$.

$$\text{where } p = \frac{|c|}{|B_1| + |B_2| + |B_3|} = \frac{|B_1 b_1 + B_2 b_2 + B_3 b_3|}{|B_1| + |B_2| + |B_3|} \tag{7}$$

Let x^* be the solution of the two linear equations

$$h_1(x) = p \text{ sgn }(c\, B_1)$$

$$h_2(x) = p \text{ sgn }(c\, B_2)$$

which exists (and is unique) by our assumption that all 2 x 2 matrices $\neq 0$. Then by (6)

$$\begin{aligned} B_3 h_3(x^*) &= c - B_1 h_1(x^*) - B_2 h_2(x^*) \\ &= c - B_1 p \text{ sgn }(c\, B_1) - B_2 p \text{ sgn }(c\, B_2) \\ &= c - \left[p \text{ sgn } c\, |B_1| + p \text{ sgn } c\, |B_2|\right] \\ &= c - \left[c - p \text{ sgn } c\, |B_3|\right] \\ &= p \text{ sgn } c\, |B_3|. \end{aligned}$$

Dividing both sides by B_3

$h(x^*) = sgn(c\ B_3)$. *Since we started with any three equations of the system, we have established the existence of a unique Tchebycheff Solution to any three of our n equations. We now extend it to all n equations.*

There are ${}_nC_3$ *combinations of three equations. Now there is a theorem by Helley* [6] *which states: if* $k_1 \ldots\ldots k_p$ *are convex sets such that every k = 1 of them have a point in common, then all p also have a point in common. Roughly translated for our purposes, it says this: any three of our n equations define three lines in a plane that form a triangle. This must be so, as no two of the lines are parallel (2 x 2 matrices* $\neq 0$*). (The Tchebycheff solution is a point interior to this triangle.) Therefore by the theorem, since every three lines have a Tchebycheff Solution in common, then all n lines have a Tchebycheff Solution in common. This proves the existence. To show the uniqueness, let M be the largest minimal deviation p (defined by (7) of all the* ${}_nC_3$ *sets of three equations. Choose any one of the three, say j. By our Lemma:*

$a_{j1}x_1 + a_{j2}x_2 + b_j = h_j(x)$ *where* $h_j(x)$ *is fixed.*

Then setting $c = h_j(x) - b_j$ *(*$h_j(x)$ *is constant)* $a_{j1}x_1 + a_{j2}x_2 = c$ *which is, by our Lemma, obtained for just one set of values* $x = (x_1, x_2)$*. Hence our Tchebycheff Solution is unique.*

Now that we have proven the existence of a unique Tchebycheff Approximation for any set of n overdetermined equations in two unknowns, we now set up a method for solving the system.

The principle of the method is this: we know from the proof that there is a unique Tchebycheff Solution to any $m + 1 = 3$ of the n equations. Also in the proof we used the concept of a maximal - minimal deviation that existed for at least one of the ${}_nC_3$ sets of three equations. This triple would therefore be $\geqq$ to any other set and hence be the Tchebycheff Solution for the entire system.

<u>The Method:</u> To find the minimal deviation M of any set of $m + 1 = 3$ equations.

1) $$M = \frac{|D|}{\sum B_j} \quad (i = 1, 2, 3)$$ *where*

D is an $(m+1) \times (m+1)$, (3×3) determinant built by adjoining the column vector b to the column vectors made by the coefficients of x_1, x_2. B_1, B_2, B_3 represent the cofactors of the elements of the adjoined column b_1, b_2, b_3.

2) *Set each of the three equations $h_i(x) = e_i M$, with $e_i = \pm 1$*

The sign of e_i is determined by the equation $\operatorname{sgn} D \cdot e_i = B_i$ $\quad i=1, 2, 3$.

3) Compute M for all the $_nC_3$ sets of $m + 1 = 3$ equations. The largest M is the Tchebycheff Solution for the entire system.

This method becomes overbearing for large systems; for example with ten equations, $_{10}C_3 = 120$ 3×3 determinants would have to be evaluated before we determined the Tchebycheff Solution for the ten equations. However, it should be pointed out that while this method does not constitute an algorithm, at least it proves that the problem can be handled in a finite number of steps.

Example 7. *Solving an overdetermined system of four equations in two unknowns. I have chosen four, for to increase it by one would mean solving ten determinants. There are algorithms for solving this problem; one by Stiefel will be indicated later in IV. Stiefel also indicated a method of simultaneously raising the lower bound and lowering the upper bound on the minimax value. At this writing, work is being done on this by Cheney and Goldstein.*

Solution on page 30-31

Example 8. *Solving four equations in three unknowns. I included this problem to indicate the method is the same for larger systems of unknowns.*

Solution on page. 32

Problem 7 . *Overdetermined System in Two Unknowns.*

		$h^*(x)$	
1)	$3x_1 - x_2 - 2 = 0$	$-7/4$	where $x_1^* = 1/2$
2)	$x_1 + x_2 = 0$	$+7/4$	$x_2^* = 5/4$
3)	$x_1 - 2x_2 + 1 = 0$	-1	$h^*(x) = \pm 7/4$
4)	$2x_1 - x_2 + 2 = 0$	$+7/4$	

Equations

$$\begin{matrix} 1) \\ 2) \\ 3) \end{matrix} \begin{vmatrix} 3 & -1 & -2 \\ 1 & 1 & 0 \\ 1 & -2 & +1 \end{vmatrix} = 10. \quad M_1 = \frac{10}{|-3| + |+5| + |+4|} = \frac{5}{6}$$

$$\begin{matrix} 1) \\ 2) \\ 4) \end{matrix} \begin{vmatrix} 3 & -1 & -2 \\ 1 & 1 & 0 \\ 2 & -1 & 2 \end{vmatrix} = 14. \quad M_2 = \frac{14}{|-3| + |+1| + |+4|} = \frac{14}{8} = \frac{7}{4}$$

$$\begin{matrix} 1) \\ 3) \\ 4) \end{matrix} \begin{vmatrix} 3 & -1 & -2 \\ 1 & -2 & 1 \\ 2 & -1 & 2 \end{vmatrix} = -15. \quad M_3 = \frac{15}{|+3| + |+1| + |-5|} = \frac{15}{9} = \frac{5}{3}$$

$$\begin{matrix} 2) \\ 3) \\ 4) \end{matrix} \begin{vmatrix} 1 & 1 & 0 \\ 1 & -2 & 1 \\ 2 & -1 & 2 \end{vmatrix} = 1. \quad M_4 = \frac{1}{|+3| + |+3| + |-3|} = \frac{1}{9}$$

Since M_2 *is the largest of the* M_i*'s, this is minimax value for the system. Now using* $e_i \operatorname{sgn} b_i = \operatorname{sgn} D$ *(i=1, 2, 4) we find* $e_1 = -1$, $e_2 = e_4 = +1$.

Solving:

1) $3x_1 - x_2 - 2 = -7/4$

2) $x_1 + x_2 = +7/4$

4) $2x_1 - x_2 + 2 = +7/4$

$$4x_1 - 2 = 0$$

$$x_1 = 1/2$$

$$x_2 = 7/4 - 1/2 = 5/4$$

Problem 8 . Overdetermined System on 3 Unknowns.

1) $x_1 + x_2 + x_3 - 2 = 0$ $\qquad \frac{8}{3} + \frac{1}{3} - 2 - 2 = -1$

2) $x_3 + 1 = 0$ $\qquad -2 + 1 = -1$

3) $2x_1 - x_2 + 3x_3 + 2 = 0$ $\qquad \frac{16}{3} - \frac{1}{3} - 6 + 2 = +1$

4) $-x_1 + 2x_2 - x_3 - 1 = 0$ $\qquad -\frac{8}{3} + \frac{2}{3} + 2 - 1 = -1$

Where

$x_1^* = 8/3$

$x_2^* = 1/3$

$x_3^* = -2$

$|h^*(x)| = 1$

$$M = \begin{vmatrix} 1 & 1 & 1 & -2 \\ 0 & 0 & 1 & 1 \\ 2 & -1 & 3 & 2 \\ -1 & 2 & -1 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 3 & -2 \\ 0 & 0 & 0 & 1 \\ 2 & -1 & 1 & 2 \\ -1 & 2 & 0 & -1 \end{vmatrix} = 2 \begin{vmatrix} 1 & 1 & 3 \\ 2 & -1 & 1 \\ -1 & 2 & 0 \end{vmatrix} = 2 \cdot 6 = 12$$

$$\overset{-B}{\begin{vmatrix} 0 & 0 & 1 \\ 2 & -1 & 3 \\ -1 & 2 & -1 \end{vmatrix}} = -3. \qquad \overset{+B}{\begin{vmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ -1 & 2 & -1 \end{vmatrix}} = -3. \qquad \overset{-B}{\begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ -1 & 2 & -1 \end{vmatrix}} = +3. \qquad \overset{+B}{\begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ -1 & 2 & -1 \end{vmatrix}} = -3.$$

$$M = |D| \Big/ \sum_{i=1}^{4} |B_i| = +1, \qquad e_1 = e_2 = e_4 = -1, \quad e_3 = +1.$$

$x_1 + x_2 + x_3 - 2 = -1$

$x + 1 = -1$

$2x_1 - x_2 + 3x_3 + 2 = +1$

$-x_1 + 2x_2 - x_3 - 1 = -1$

$x_3 = -2$

$x_1 + x_2 = 3$

$2x - x = 5$

$3x_1 = 8$

$x_1 = 8/3$

$x_2 = 1/3$

III. *Tchebycheff Approximation of Discrete Point Sets by a Polynomial of degree k.*

A. *Theorem and Proof.*

Now we will examine the special subset of the overdetermined system that we mentioned in the last section, namely finding the Tchebycheff Approximation to any n set of discrete points (x_i, y_i) $(i=1 \ldots n)$ *by* $P_k(x)$. *Again for the sake of continuity we will examine the simple case, which in this section is approximation by a linear equation* $y = mx + b$. *(For convenience we will consider m and b the variables to be evaluated, and x and y the constants.) It should be pointed out that our results will hold for the general case* $P_k(x)$, *and we will do an approximation by a quadratic to help emphasize the point.*

One nice feature of this special subset is the well ordering of one set of the constants, in this case the abscissas

$$x_i \ (i = 1, 2, \ldots \ldots . n)$$

This will help return us to the concept of best approximation of section one, and give us an analog of Theorem C.

<u>*Theorem IV*</u>. *If* $\emptyset_k^*$ *is of the form* $y = mx + b$, *then to any given point set* (x_i, y_i) $(i = 1, 2, \ldots . n)$, $x_1 < x_2 < \ldots \ldots < x_n$, *there are*

at least $k + 2 = 3$ *E points which are alternately + and -. This seems rather wordy; however, as we shall see, if the points are not well-ordered (i.e.,* (x_1, y_1) (x_2, y_2) *) then we would not be guaranteed a unique Tchebycheff line. The well-ordering continues our assumption of Section two that no 2 x 2 matrix equals 0.*

The proof is simple. We have already established in Section two existence, uniqueness, and the fact that there is one set of three equations that has equal deviation E and satisfies the entire system. It remains to prove that E is of alternating sign. There is no loss in generality by assuming the triple is (x_1, y_1) (x_2, y_2) (x_3, y_3) *with* $x_1 < x_2 < x_3$. *The three equations defined are:*

$$x_1 m + 1 \cdot b - y_1 = 0$$

$$x_2 m + 1 \cdot b - y_2 = 0$$

$$x_3 m + 1 \cdot b - y_3 = 0$$

Using the method of Section two:

$$\text{sgn } B_i \cdot e_i = \text{sgn } D, \qquad (i = 1, 2, 3) \quad e_i = \pm 1$$

then

$$B_1 = +\begin{vmatrix} x_2 & 1 \\ x_3 & 1 \end{vmatrix} = x_2 - x_3 < 0 \qquad (x_3 > x_2).$$

$$B_2 = -\begin{vmatrix} x_1 & 1 \\ x_3 & 1 \end{vmatrix} = x_3 - x_2 > 0$$

$$B_3 = +\begin{vmatrix} x_1 & 1 \\ x_2 & 1 \end{vmatrix} = x_1 - x_2 < 0$$

so regardless of sgn D, the E points alternate in sign. *Q. E. D.*

Another nice feature about the order in this special subset is we will no longer have to compute the ${}_nC_3$ errors as we did in the more general case of overdetermined equations. The new method, known as the exchange method, is explained in the following Algorithm.[8]

<u>Step one</u>. Choose any triple from the data points. The triple defines three equations of the form $x_j\,m + b = y_j + h(x)$, where $h(x) = e_j M$, defined in Section two, alternates in sign according to the ordering. (A new method to find $h(x)$, m, b, in this special case is: add adjacent equations to eliminate $h(x)$, subtract the two remaining equations to eliminate b. Solve for m and substitute to find b, $h(x)$.

<u>Step two</u>. Calculate the errors at all remaining points. If $h(x) \geq$ to all of these, we have found the best line. If not:

<u>Step three.</u> This is the exchange set. Select a new triple by adding to the old triple a data point at which the largest error occurs, dropping one of the former points in such a way that the new triple has errors of alternating sign. (This is always possible). Continue this process until you have found the best line.

B. Examples of Application

<u>Example 9.</u> Finding the best line to a discrete set of points.

These points came from an overdetermined system of linear equations which I normalized to the point problem by dividing through by the coefficients of b. (This accounts for the odd point set.) In case you are wondering, the solution will not solve the original. Dividing through by unequal numbers provides weights to the errors.

Solution on page 38-39.

<u>Example 10.</u> In this example we study the effect on the point problem when we drop the restriction that every 2 x 2 matrix ≠ 0.

Solution and Discussion on page 40-42

<u>Example 11.</u> Here we approximate a point set with a quadratic. I have tried to use a set of points that might reasonably represent a frequency distribution, so that the quadratic might have a possible prediction value. In first doing this problem, I made an oversight on the first exchange, taking my points in + - + + order. It is interesting to note that two exchanges later Tchebycheff returned me to the quad where the error was made. In addition, the reference error was lowered, an impossibility with the exchange method as the absolute value of each reference error forms a

lower bound to the absolute value of $E = h^*(x)$. *I should also mention that I had spent so much time on this problem I abandoned the exchange method after the first exchange and chose the largest errors in alternating order, and happened to hit the right quad.*

Solution on page 43.

<u>Problem 9</u>. *Find $\phi^* = ax + b$ for the given set of points.*

	x	y		h_1	h_2	h_3
1)	-4	2	$-4m + b - 2 = 0$	+ 1/6	- 39/11	- 35/24
2)	-2	0	$-2m + b - 0 = 0$	+15/6	+3/11	+25/24
3)	-5/3	- 1/3	$-5/3\ m + b + 1/3 = 0$	+26/9	+10/11	+35/24
4)	-1	0	$-m + b - 0 = 0$	+16/6	+13/11	+31/24
5)	- 1/2	- 1/2	$-1/2(m) + b - 1/2 = 0$	+27/12	+25/22	+22/24
6)	0	3	$b - 3 = 0$	-1/6	-10/11	-35/24
7)	1/5	1	$+1/5(m) + b - 1 = 0$	+56/30	+14/11	+14.1/24
8)	3/4	3/2	$3/4\ (m) + b - 3/2 = 0$	+35/24	+14/11	+11/48
9)	1	1	$m + b - 1 = 0$	+ 2	+ 2	+19/24
10)	2	3	$2m + b - 3 = 0$	+1/6	+10/11	-23/24

1) $-4m + b = 2 + h_1$

6) $b = 3 - h_1$

10) $2m + b = 3 + h_1$

$-4m + 2b = 5$

$2m + 2b = 6$

$-6m = -1$

$m = 1/6$

$b = 17/6$

$h_1 = -1/6$

3) $-5/3(m) + b = -1/3 + h_2$

6) $+ b = 3 - h_2$

10) $2m + b = 3 + h_2$

$-5/3(m) + 2b = 8/3$

$2m + 2b = 6$

$-11/3(m) = -10/3$

$m = +10/11$

$b = +23/11$

$h_2 = +10/11$

continued next page

$$
\begin{array}{lrcl}
1) & -4m + b & = & 2 + h_3 \\
3) & -\tfrac{5}{3}m + b & = & -\tfrac{1}{3} - h_3 \\
6) & b & = & 3 + h_3 \\
\hline
 & -\tfrac{17}{3}(m) + 2b & = & \tfrac{5}{3} \\
 & -\tfrac{5}{3}(m) + 2b & = & \tfrac{8}{3} \\
\hline
 & -4m & = & -1
\end{array}
$$

$$m = 1/4 = 6/24$$

$$b = 37/24$$

$$h_3 = -35/24$$

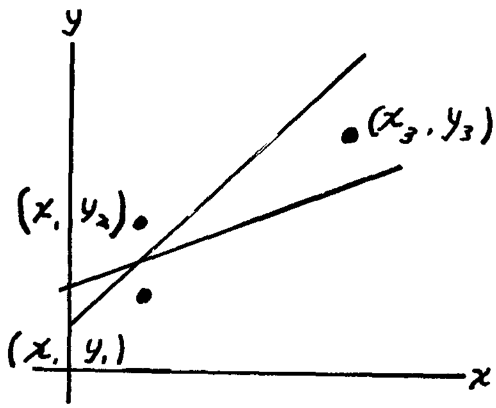
$\therefore$ $m^* = 1/4$ $b^* = 37/24$ $|h^*| = 35/24$

Example 10. If we drop the restriction that every 2 x 2 matrix $\neq 0$ then we allow, say, the equations

$$x_1 m + b = y_1$$

$$x_1 m + b = y_2 \qquad y_1 \neq y_2$$

which define the points (x_1, y_1) (x_1, y_2). If we consider any third point (x_3, y_3) then there is ambiguity in terms of the Tchebycheff line.



As the diagram shows, either line would satisfy the Tchebycheff conditions. As a matter of fact, any line passing through the mid-point of the line joining (x_1, y_1) and (x_1, y_2) and passing within an h vertical distance $[h = (y_2 - y_1)/2]$ of (x_3, y_3) would satisfy the conditions, since we are only concerned with minimizing the maximum error.

Herein lies the clue to whether or not there is ambiguity for a given set of discrete points. Suppose h^* is the minimax value for the system, and we again have (x_1, y_1) and (x_1, y_2) with $h = (y_2 - y_1)/2$. Then if $h^* > h$ there is no ambiguity, if $h^* = h$ there is the possibility of ambiguity. (If $h = h^*$ there would be no ambiguity if only one of the points (x_1, y_1), (x_1, y_2) were used in the triple that generated h^*.) Of course $h^* < h$ is impossible by the definition of h^*.

We include the following problem only as an attempt to gain some insight into the generalities of the discrete point problem. The reader interested in the effect on any overdetermined system by dropping the restriction that every 2 x 2 ≠ 0 is referred to reference [7].

Example

x	y			h_1	h_2
-2	+2		$-2m + b = 2$	$+7/4$	$+3/6$
-1	-1		$-1m + b = -1$	$-7/4$	$-13/6$
0	3	3	$b = 3$	$+7/4$	$+13/6$
0	1		$b = 1$	$-1/4$	$+1/6$
0	0		$b = 0$	$-5/4$	$-5/6$
1	2		$m + b = 2$	$+1/4$	$+9/6$
2	-1	1	$2m + b = -1$	$-13/4$	$-7/6$
2	-2		$2m + b = -2$	$-17/4$	$-13/6$

$$-2m + b = 2 + h_1$$
$$-m + b = -1 - h_1$$
$$b = 3 + h_1$$

$$-3m + 2b = 1$$
$$-m + 2b = 2$$

$$m = 1/2 \quad b = 5/4 \quad h = 7/4$$
$$y = 1/2\,(x) + 5/4$$

$$- m + b = -1 - h_2$$

$$b = 3 + h_2$$

$$2m + b = -2 - h_2$$

$$-m + 2b = 2$$

$$2m + 2b = 1$$

$m^* = -1/3 \quad b = 5/6 \quad h_2 = -13/6$

$\therefore \quad y = -1/3(x) + 5/6$

with $m^* = -1/3 \quad b^* = 5/6 \quad |h^*| = 13/6$

We were guaranteed with the first set of points that there would be no ambiguity as $h_1 = 7/4$ and $1/2$ of the largest difference of the multiple points $= 3/2$. The exchange method guarantees all future errors will be greater than $7/4$.

Problem 11. *Find parabola of best approximation for this set of discrete points.*

Assume $\phi^*(x) = ax^2 + bx + c.$

Given

x	y	h_1	h_2	h_3
0	0	+ .5	- 2.4	- 1.75
1	3	- .5	- .9	- .7
2	6	+ .5	+ .9	+1.0
3	5	- .5	- .9	- .8
4	8	+4.5	+1.8	+1.75
5	6	+6.5	- .3	- .1
6	7	+13.5	+1.4	+1.5
7	4	+18.5	- .6	- .4
8	1	+25.5	-2.4	-1.75
9	2	+38.5	+ .6	+1.2
10	0	+50.5	+ .9	+1.75

$\phi_1(x) = 1/2\ (-2x^2 + 10x - 1)$

$\phi_2(x) = 1/64\ (-14x^2 + 120x + 143)$

$\phi_3^*(x) = 1/144\ (-35x^2 + 298x + 260)$

Exchange values :

0	1	0
1 →	2 →	4
2	3	8
3	10	10

$h_1 = -0.50$

$h_2 = +0.89$

$h_3 = +1.75$

IV. Equivalence of the Problem of Linear Programming to the System of Overdetermined Equations. [10]

The basic problem of linear programming is to maximize a linear form $z = \sum_{k=1}^{m} p_k x_k$ subject to the linear in equalities $y_j = \sum_{k=1}^{m} a_{jk} x_k + b_j = 0 \qquad j = 1, 2, \ldots\ldots n$

If $m = 2$, this problem has the same dimension of our overdetermined problem, further if $n = m+1$ we have

maximize $z = p_1 x_1 + p_2 x_2$ subject to

$$y_1 = a_{11} x_1 + a_{12} x_2 + b_1 \geq 0$$
$$y_2 = a_{21} x_1 + a_{22} x_2 + b_2 \geq 0$$
$$y_3 = a_{31} x_1 + a_{32} x_2 + b_3 \geq 0$$

We now attempt to doctor our overdetermined system to fit this problem. We have

$$h_1(x) = a_{11} x_1 + a_{12} x_2 + b_1$$
$$h_2(x) = a_{21} x_1 + a_{22} x_2 + b_2$$
$$h_3(x) = a_{31} x_1 + a_{32} x_2 + b_3$$

Where we need to minimize $h^*(x) = \max \left| h_j(x) \right| \qquad j = 1, 2, 3.$

We introduce the equation $z = x_{m+1} = x_3 \quad (p_1 = p_2 = 0)$,

and the equations $y_j = h_j(x) + x_3 \qquad y_j = -h_j(x) + x_3 \qquad j = 1, 2, 3.$

Our problem then is minimize $z = x_3$ *by appropriate choice of the* $m + 1 = 3$ *variable subject to the constraints*

$$y_j = h_j(x) + x_3 \geq 0$$
$$y_j = -h_j(x) + x_3 \geq 0 \qquad (j = 1, 2, 3) \qquad (1)$$

This matches the two problems. We must now show that the value (x_1, x_2) *corresponding to this minimum is the solution of the Tchebycheff problem.*

Let $z = x_3$ *be the minimum of the program. Rewrite (1) so that* $x_3 \geq h_j(x)$, $x_3 \geq -h_j(x)$ $j = 1, 2, 3.$
Suppose we fix (x_1, x_2) *and allow* x_3 *to vary.*
Then $h_j(x) = a_{j1} x_1 + a_{j2} x_2 + b_j$ *is also fixed so that for our particular choice of* (x_1, x_2), $z = x_3$ *is the larger of the two numbers* max $h_j(x)$, max $-h_j(x)$ $(j = 1, 2, 3)$, *or equivalently* $z = \max |h_j(x)| = k$. *Since k is a restricted minimum for our arbitrary choice it certainly is not smaller than the minimum, say* z_0, *corresponding to the* _free_ *variation of all the variables* (x_1, x_2, x_3), *therefore* $k \geq z_0$. *This establishes the Tchebycheff conditions with* $z_0 = h^*(x)$, *the minimax of* $k = |h_j(x)|$ $j = 1, 2, 3.$

This establishes that the linear programming problem can be solved by using an algorithm that is adaptable to a computer. If we return to the examples of Section II, (we solved these by a method that had obvious limitations), the equivalent problems become:

Problem 7. Minimize $z = x_3$ *subject to the restrictions*

$$3x_1 - x_2 + x_3 - 2 \geq 0 \qquad -3x_1 + x_2 + x_3 + 2 \geq 0$$

$$x_1 + x_2 + x_3 \geq 0 \qquad -x_1 - x_2 + x_3 \geq 0$$

$$x_1 - 2x_2 + x_3 + 1 \geq 0 \qquad -x_1 + 2x_2 + x_3 - 1 \geq 0$$

$$2x_1 - x_2 + x_3 + 2 \geq 0 \qquad -2x_1 + x_2 + x_3 - 2 \geq 0$$

Problem 8. Minimize $z = x_4$ *subject to the restrictions*

$$x_1 + x_2 + x_3 + x_4 - 2 \geq 0 \qquad -x_1 - x_2 - x_3 + x_4 + 2 \geq 0$$

$$x_3 + x_4 + 1 \geq 0 \qquad -x_3 + x_4 - 1 \geq 0$$

$$2x_1 - x_2 + 3x_3 + x_4 + 2 \geq 0 \qquad -2x_1 + x_2 - 3x_3 + x_4 - 2 \geq 0$$

$$-x_1 + 2x_2 - x_3 + x_4 - 1 \geq 0 \qquad x_1 - 2x_2 + x_3 + x_4 + 1 \geq 0$$

REFERENCES

1. *Cheney, W. and Goldstein, A.*

   Note and paper by Zuhovickii concerning The Tchebycheff Problem for Linear Equations.

   Journal Soc. Indust. Appl. Math 1958.

2. *Goldstein, A. and Cheney, W.*

   A Finite Algorithm for the Solution of Consistent Linear Equations and Inequalities and for the Tchebycheff Approximation of Inconsistent Linear Equations.

   Pacific Journal of Mathematics, 1958.

3. *Hastings, E.*

   Approximations for Digital Computers.

   Princeton University Press, 1955.

4. *Lanczos, C.*

   Applied Analysis

   Prentice Hall, Inc. 1956.

5. *Langer, R.*

On Numberical Approximation

University of Wisconsin Press, 1959.

6. *Rademacher, H. and Shoenberg, I. J.*

Helly's Theorems on Convex Domains and Tchebycheff's Approximation Problem.

Canadian Journal of Mathematics, 1950. Pg 245-256.

7. *Rivlin, T. S.*

Overdetermined Systems of Linear Equations.

S. I. A. M. Review, January, 1963. Pg 52-66.

8. *Scheid, E.*

The Under-Over-Under Theorem.

American Math. Monthly. Vol. 68 No. 9, November, 1961.

9. *Stiefel, E.*

Jordan Elimination, Linear Programming and Tchebycheff Approximation.

Numerische Mathematik (2) 1960. pp 1-17.

10. *Stiefel, E.*

Numerical Methods of Tchebycheff Approximation

(In Ref. [5]) pp 217-232.

11. *Todd, J.*

Survey of Numerical Analysis

McGraw-Hill Book Co., Inc. 1962.