1962

Least squares approximations

BOSTON UNIVERSITY

GRADUATE SCHOOL

Thesis

LEAST SQUARES APPROXIMATIONS

by

MARVIN E. WIENER

(A.B. Yeshiva University, 1959)

Submitted in partial fulfillment of the requirements for the degree of

Master of Arts

1962

AM
Copy 1

Approved
by

First Reader ...Donald W. Blackett...
Professor of

Second Reader ......Gerhard S. Norton
Professor of

TABLE OF CONTENTS

INTRODUCTION

The problem of determining the best approximation to a function arbitrarily defined on some interval is a most fascinating subject. A thorough investigation into it would involve higher analysis. If the approximation is to be a polynomial, as is the case in this paper, the Lebesgue Theory would be an indispensable tool.

It is not my intention to use higher analysis in this endeavor. I shall be mainly concerned with an approach to the following problem: To determine that polynomial of degree n which offers the best approximation in the sense of "least squares" to a function which is continuous or has, at most, a finite number of discontinuities in some interval.

The approach is through vector spaces. The function and its approximation are to be considered elements of a linear normed vector space.

CHAPTER I

APPROXIMATION IN LINEAR VECTOR SPACES

1.1 The Linear Normalized Vector Space

The purpose of this chapter is to develop and show how algebraic procedure may be used in the determination of an approximation of a function in the finite interval (a,b). We begin by defining a linear vector space since it will be common to all phases in the development of the theory of approximation.

Definition[1]

A Linear Normalized Vector Space E, with real scalars, is a vector space such that for any two vectors x and y in E there corresponds an inner product (x,y) which is symmetric, bilinear, and positive.

The norm or length of a vector x in E will be defined as $(x,x)^{\frac{1}{2}} = \|x\|$.

We introduce the metric in E by defining the distance between two vectors x and y in E as $d(x,y) = \|x - y\|$. This distance function has the following properties:

$$d(x,y) > 0 \qquad \text{unless } x = y$$

$$d(x,y) = d(y,x)$$

$$d(x,z) \leqq d(x,y) + d(y,z)$$

---

[1] Birkhoff & Mac Lane, A SURVEY OF MODERN ALGEBRA (New York: The Macmillan Company, 1953), p. 189.

1.2 An Approximation in a Vector Space

Let γ be any element in the Linear Normalized Vector Space E and let $\alpha_1, \alpha_2, \ldots\ldots, \alpha_n$ be n linearly independent elements in E. In order to obtain the best approximation of γ in terms of the α_i's, it is necessary to determine the numbers $c_1, c_2, \ldots\ldots, c_n$ such that

$$\varphi(c_1, c_2, \ldots\ldots, c_n) = \| \gamma - c_1\alpha_1 - c_2\alpha_2 - \cdots - c_n\alpha_n \|$$

shall be a minimum.

We will first prove that the numbers $c_1, c_2, \ldots\ldots, c_n$ having the required property, exists. Then, we will determine their representation in terms of γ and α_i's, $i = 1, \ldots, n$

Proof:[2]

$\varphi(c'_1, \ldots\ldots, c'_n)$ is a continuous function.

By the triangular inequality, we have:

$$|\varphi(c'_1, c'_2, \ldots, c'_n) - \varphi(c_1, c_2, \ldots, c_n)| =$$

$$\left| \| \gamma - \sum_{k=1}^{n} c'_k\alpha_k \| - \| \gamma - \sum_{k=1}^{n} c_k\alpha_k \| \right| \leqq$$

---

[2] N. I. Achieser, THEORY OF APPROXIMATION (New York: Frederick Ungar Publishing Co., 1956), p. 10.

$$\left\| \sum_{k=1}^{n} (c_k' - c_k)\alpha_k \right\| \leqq \sum_{k=1}^{n} |c_k' - c_k| \, \|\alpha_k\| \leqq$$

$$\operatorname{Max}_j |c_j' - c_j| \sum_{k=1}^{n} \|\alpha_k\|$$

If we now introduce another continuous function

$$\psi(c_1, c_2, \ldots, c_m) = \| c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m \|$$

according to one of Weierstrass's theorems, the function ψ has a minimum μ upon the "sphere"

$$|c_1|^2 + |c_2|^2 + \cdots + |c_m|^2 = 1$$

which is a bounded closed point set in ordinary space of a finite number of dimensions.

The non-negative number μ must be greater than zero since $\alpha_1, \alpha_2, \ldots, \alpha_m$ are linearly independent.

Let $\rho \geqq 0$ denote the greatest lower bound of the values of the function ϕ. If

$$\sqrt{\sum_{k=1}^{n} |c_k|^2} > \frac{1}{\mu}(\rho + 1 + \|\gamma\|) = R$$

then

$$\phi(c_1, c_2, \ldots, c_m) \geqq \| c_1\alpha_1 + \cdots + c_m\alpha_m \| - \|\gamma\| \geqq$$

$$\sqrt{\sum_{k=1}^{m} |c_k|^2} \; \mu - \|\gamma\| > \rho + 1$$

Therefore, in seeking the minimum of the function $\phi(c_1, \cdots, c_m)$ we can restrict ourselves to those values of c_k for which $\sqrt{\sum_{k=1}^{m} |c_k|^2} \leqq R$

In other words, to the consideration of the function in a bounded closed region. In such a region, a continuous function has a minimum.

We now wish to determine these c_k's.

Let γ, α, β be vectors in E, and let the representation of γ take the form

$$\gamma = \alpha + \beta$$

where the vector α is considered to be the approximation of γ with the remainder being β.

Let E' be an n-dimensional subspace of E and let γ be a vector which is, in general, not an element in E'.

Then, if $\alpha_1, \alpha_2, \dots, \alpha_n$ are any n linearly independent vectors in E',

$$\gamma = a_1\alpha_1 + a_2\alpha_2 + \cdots a_n \alpha_n + \beta$$

Taking the scalar product of γ and α_k, we get

$$(\gamma, \alpha_k) = a_1(\alpha_1, \alpha_k) + \cdots a_k(\alpha_k, \alpha_k) + \cdots a_n(\alpha_n, \alpha_k) + (\beta, \alpha_k)$$

If it is assumed that β is orthogonal to E' so that $(\beta, \alpha_k) = 0$ for k = 1, ..., n, this results in a system of n simultaneous equations from which the coefficients a_i may be determined.

This, of course, would be laborious. It would be simple to determine the a_i's if the α_i's were an orthogonal basis of E'. With the aid of the Gram Schmidt Orthogonalization Procedure, it is possible to obtain an orthonormal basis in E' such that $\delta_1, \delta_2, \ldots, \delta_m$ may be found where

$$\gamma = c_1\delta_1 + c_2\delta_2 + \cdots + c_m\delta_m + \beta$$

and $\quad C_i = (\gamma, \delta_i) \quad$ when β is orthogonal to γ

Theorem:

The remainder β is a minimum if α is selected so that β is orthogonal to E', that is, so that $\alpha = \sum_{i=1}^{n} c_i\delta_i$ where $c_i = (\gamma, \delta_i)$

Proof:

We will now show that the norm of the squared error $\|\gamma - \alpha\|^2$ is a minimum. Let λ be another approximation to γ

$$\lambda = \sum_{i=1}^{n} d_i\delta_i$$

$$\begin{aligned}
\|\gamma - \alpha\| &= (\gamma,\gamma) - 2(\gamma,\lambda) + (\lambda,\lambda) \\
&= \|\gamma\|^2 - 2(\gamma, (\Sigma d_i\delta_i)) + \|\lambda\|^2 \\
&= \|\gamma\|^2 - 2\Sigma d_i(\gamma,\delta_i) + \Sigma d_i^2 \\
&= \|\gamma\|^2 - 2\Sigma d_i c_i + \Sigma d_i^2 \\
&= \|\gamma\|^2 - \Sigma c_i^2 + \Sigma(c_i - d_i)^2
\end{aligned}$$

Therefore, in order that β be a minimum, $c_i = d_i$

The process used in determining α may result in a perfect representation of γ when both γ and α are in the same subspace of E.

CHAPTER II

THE BEST APPROXIMATION IN THE SENSE OF LEAST SQUARES

2.1 Introduction

If a function f(x) is to be approximated by a polynomial, it is natural to ask to what extent it is legitimate to attempt polynomial approximation. The answer was given by Weierstrass.

A function, continuous in a finite closed interval, can be approximated with a preassigned accuracy by polynomials. A function of a real variable which is continuous and has a period of 2π can be approximated by trigonometric polynomials.[3]

There are many formulas which may be used to approximate functions. For example, suppose f(x) is to be approximated over a domain D. If f(x) has enough derivatives, we may expand it about the point x in the interval (a,b) in the form

$$f(x) \approx f(x_0) + f'(x_0)(x-x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n.$$

This is, of course, Taylor's Approximation Theorem. We may use some of the many interpolation formulas as f(x)'s approximation. However, these approximation formulas

---

[3] D. Jackson, THE THEORY OF APPROXIMATION, American Mathematical Society Colloquium Publications, Vol. XI, 1930, p. 1.

may not be practical. For example, in the case where the given data is subject to error or where the function to be approximated is known to be fairly smooth and the approximation polynomial must be of high degree in order to be accurate.

A method of finding a polynomial approximation which takes into consideration the above as well as other problems is the method of least squares.

2.2 Function Space - A Normed Linear Vector Space

F is a linear vector space over the real field R if its elements comply to the operations of addition and scalar multiplication that define any vector space.

The scalar product of two real functions f(x) and g(x) in F with respect to a positive weighting function $\omega(x)$ where x ranges over the real interval (a,b), is defined by the Riemann Integral

$$(f,g) = \int_a^b \omega(x)\, f(x)\, g(x)\, dx$$

The real-valued function $\|f\| = (f,f)^{1/2}$ defined in F is said to be a norm (or length) since:

$\|f\| \geqq 0$ and $\|f\| = 0$ if, and only if, f = 0

$\|\alpha f\| = |\alpha|\,\|f\|$

$\|f+g\| \leqq \|f\| + \|g\|$

We introduce the metric in F by defining

$$d(f,g) = \|f-g\|$$

This is the distance function in the following sense:

$d(f,g) > 0$ unless $f = g$

$d(f,g) = d(g,f)$

$d(f,g) \leqq d(f,h) + d(h,g)$

The resulting metric space is known as a Normed Linear Vector Space.

The following general problem of approximation will have meaning in a real normed linear vector space F in the interval (a,b).

Given a $f(x) \in F$ and a set of linearly independent elements $\varphi_1(x), \varphi_2(x), \ldots\ldots, \varphi_n(x)$ all in F

Find the linear combination $\sum_{i=1}^{n} a_i \varphi_i(x)$ which is the best approximation in the sense of the given norm $\|f\|$ to f, i.e. determine

$$\min_{\{a_i\}} \left\| f - \sum_{i=1}^{n} a_i \varphi_i(x) \right\|$$

This is essentially the problem solved in Chapter I. The problem of approximating a vector in a vector space as a linear combination of n independent, but otherwise arbitrary vectors, is analogous to the problem of approximating an arbitrary function in some fundamental domain as a linear combination of members of some given set of n functions.

Whether or not f(x) is defined analytically or is given by a set of discrete data points, the solution $\sum a_i \varphi_i(x)$ that will be obtained will be based on the principle of least squares.

2.3 The Principle of Least Squares

Since we are dealing in a linear normed vector space, we will let $P_0(x), P_1(x), \ldots\ldots, P_n(x)$ be a sequence of orthonormal functions of degree n over an interval (a,b), that is, a sequence of functions satisfying the conditions:

$$\int_a^b \omega(x) P_i(x) P_j(x)\, dx = 0 \qquad \int_a^b \omega(x)\, P_n^2(x)\, dx = 1$$

where w(x) is a positive weighting function.

Let f(x) be another function defined over the same interval. The functions $P_k(x)$ and f(x) may be continuous or, more generally, bounded and integrable in the sense of Riemann, or merely summable together with their squares. The problem of least squares is that of determining a set of coefficients $c_i = \int_a^b \omega(x) f(x) P_i(x)\, dx$ so that when

$$\varphi_n(x) = c_0 P_0(x) + c_1 P_1(x) + \cdots + c_n P_n(x)$$

the integral

$$\int_a^b \omega(x) \left[f(x) - \varphi_n(x) \right]^2 dx$$

regarded as a function of its coefficients, shall be a minimum.

2.4 The Discrete Case

When f(x) is not defined in an interval (a,b) explicitly, but is given by a set of N observations, it is still possible to use the properties of vector spaces to approximate it. We will be dealing with vectors of dimension N possessing N components $f_1, f_2, \ldots\ldots, f_N$. Each component f_k is the value of a function f at the point x_k ; $f_k = f(x_k)$. If a vector f has components $f_1, f_2, \ldots, f_N$ and a vector g has components $g_1, g_2, \ldots, g_N$ we will call the expression

$$(f,g) = \sum_{k=1}^{N} w_k f_k g_k$$

an inner product with respect to a positive weighting function w_k . The norm of a vector f will take the form

$$\| f \| = (f,f)^{1/2}$$

In terms of a set of orthogonal functions $\phi_1(x), \phi_2(x), \ldots, \phi_m(x)$ $m \leq N$, a given vector f possesses an orthonormal (or a Fourier) expansion

$$f \approx \sum_{k=1}^{n} (f, \phi_k)\, \phi_k$$

for each $n \leq N$, the Fourier Expansion possesses the least squares property

$$\| f - \sum_{k=1}^{n} (f, \phi_k)\phi_k \| \quad \text{be a minimum}$$

from among all possible approximations of f of the form

$$y_n(x) = \sum_{k=1}^{m} a_k \phi_k$$

2.5 Existence and Uniqueness of the Best Approximation

In the case of least squares, we now state two theorems regarding the existence and uniqueness of a best approximation to a given function.

2.5.1 Bessel's Inequality[4]

Let $\phi_1, \phi_2, \dots$ be an orthonormal system and let f be any function. The Fourier coefficients are defined $c_i = (f, \phi_i), i = 1, 2, \dots$. Now $\int (f - \sum_{i=1}^{n} c_i \phi_i)^2 dx \geq 0$ (1)

$$0 \leqq \int f^2 dx - 2 \sum_{i=1}^{n} c_i \int f \phi_i dx + \sum_{i=1}^{n} c_i^2 = Nf - 2 \sum_{i=1}^{n} c_i^2 + \sum_{i=1}^{n} c_i^2$$

$$\sum_{i=1}^{n} c_i^2 \leqq Nf$$

This proves that the sum of the squares of the expansion coefficients always converges.

The significance of the integral (1) is that it occurs in the problem of approximating the given function f(x) by a linear combination $\sum_{i=1}^{m} \gamma_i \phi_i$ with the constant

[4] Courant & Hilbert, METHODS OF MATHEMATICAL PHYSICS (New York: Interscience Publishers, Inc., 1953), Vol. I, p. 51.

coefficients γ_i and fixed n, in such a way that the mean square error

$$M = \int (f - \sum_{i=1}^{n} \gamma_i \varphi_i)^2 dx$$

is as small as possible. Since

$$M = \int (f - \sum_{i=1}^{n} \gamma_i \varphi_i)^2 dx = \int f^2 dx + \sum_{i=1}^{n} (\gamma_i - c_i)^2 - \sum_{i=1}^{n} c_i^2$$

it follows that M takes on its least value for $\gamma_i = c_i$. We may now state the following theorem on uniqueness.

Let f(x) be an arbitrary function subject to conditions of integrability. If we choose a sequence of polynomials $\{P_n(x)\}$ where $P_n(x)$ has degree n such that they are orthonormal and if we let $S_n(x)$ denote the sum

$$c_0 P_0(x) + c_1 P_1(x) + \cdots\cdots + c_n P_n(x)$$

where
$$c_k = \int_a^b f(x) P_k(x)\, dx$$

the polynomial $S_n(x)$ can be characterized among all polynomials of the nth or lower degree as the one for which the integral of the weighted square of the error is a minimum.

CHAPTER III

PRACTICAL LEAST SQUARES APPLICATIONS

3.1 Introduction

In this chapter we will derive approximating polynomials by the method of least squares to functions defined analytically or given as a set of discrete points over finite, semi-infinite, and infinite intervals.

3.2 The Legendre Approximation Theory

The best approximation in the sense of least squares to a function f(x) over the interval (a,b) is $\sum_{r=0}^{n} a_r P_r(x)$ where $P_i(x)$ satisfies the equalities

$$\int_a^b \omega(x)\, P_m(x) P_n(x)\, dx = 0 \qquad \omega(x) \geqq 0$$

$$\int_a^b \omega(x)\, P_n^2(x)\, dx = \frac{2}{2n+1}$$

The approximation in the least squares sense requires

$$\int_a^b \omega(x) \left[f(x) - \sum_{r=0}^{n} a_r P_r(x) \right]^2 dx$$

to be a minimum. If the interval (a,b) is considered to be (-1,1), and w(x) = 1 for all values of x in the interval, the above polynomials $P_i(x)$ are called the Legendre Polynomials of degree n.

Milne* derives these polynomials by a method using the orthogonality property $\int P_n P_m dx = 0$. Hildebrand** also uses this property in determining these polynomials. We shall now show how the solutions of the Legendre Differential Equation are these polynomials.

$$\frac{d}{dx}\left\{(1-x^2)\frac{dy}{dx}\right\} + n(n+1)y = 0$$

The actual solution to the Legendre Differential Equation results in two independent solutions; one solution is an infinite series and the other a polynomial solution. These polynomials are called Legendre Polynomials.

The polynomial solution is

$$Y_n(x) = x^n - \frac{n(n-1)}{2(2n-1)}x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2^2\cdot 2!\,(2n-1)(2n-3)}x^{n-4}$$
$$- \frac{n(n-1)(n-2)(n-3)(n-4)(n-5)}{2^3\cdot 3!\,(2n-1)(2n-3)(2n-5)}x^{n-6} + \cdots$$

$P_n(x) = \frac{(2n)!}{2^n(n!)} y_n(x)$ is known as the Legendre Coefficient of the nth order.

3.2.1 The Orthogonality of $P_n(x)$:

We will now show the orthogonality properties of these polynomials.

---

* See Bibliography

** See Bibliography

The Legendre Polynomials $P_n(x)$ satisfy the differential equation

$$\frac{d}{dx}\left[(1-x^2)P'_n(x)\right] + n(n+1)P_n(x) = 0$$

Multiply by $P_m(x)$ and integrate by parts.

$$\int_{-1}^{1} P_m(x)\frac{d}{dx}\left[(1-x^2)P'_n(x)\right]dx + n(n+1)\int_{-1}^{1} P_m(x)P_n(x)\,dx = 0$$

$$\int_{-1}^{1} P_m(x)\frac{d}{dx}\left[(1-x^2)P'_n(x)\right]dx = P_m\left[(1-x^2)P'_n(x)\right]_{-1}^{1} - \int_{-1}^{1}(1-x^2)P'_n(x)P'_m(x)\,dx$$

so

$$-\int_{-1}^{1}(1-x^2)P'_n(x)P'_m(x)\,dx + n(n+1)\int_{-1}^{1}P_m(x)P_n(x)\,dx = 0$$

interchanging n and m, we get

$$-\int_{-1}^{1}(1-x^2)P'_m(x)P'_n(x)\,dx + m(m+1)\int_{-1}^{1}P_n(x)P_m(x)\,dx = 0$$

subtracting

$$\left[n(n+1) - m(m+1)\right]\int_{-1}^{1}P_m(x)P_n(x)\,dx = 0 \quad \text{qed}$$

Similarly, it may be shown

$$\int_{-1}^{1}P_n^2(x)\,dx = \frac{2}{2n+1}$$

3.2.2 The Roots of $P_m(x)$

An important property of $P_m(x)$ is that all the roots are contained in the interval $(-1,1)$. By the orthogonality of $P_m(x)$,

$$\int_{-1}^{1} Q(x) P_m(x)\, dx = 0 \tag{2}$$

where $Q(x)$ is a polynomial of degree $n < m$. Since $P_m(x)$ is of degree m, it has m roots, real or complex. Each root is counted a number of times equal to its multiplicity. Let $r_1, r_2, \ldots\ldots, r_m$ denote the real roots of odd order lying in the interval $-1 \leq x \leq 1$. Then the product

$$V(x) = (x-r_1)(x-r_2)\cdots(x-r_k)\cdot P_m(x)$$

does not change sign in $(-1,1)$. Since all its zeros in this interval are of even order

$$\int_{-1}^{1} V(x)\, dx \neq 0$$

so $k = m$; otherwise, it would vanish due to (2). It follows that all the roots of $P_m(x) = 0$ are real, distinct, and lie in the interval $(-1,1)$.

3.2.3 The Space Spanned By The Legendre Polynomials

The orthogonal set of continuous functions $\{P_k(x)\}_{k=0,\ldots m}$ of degree k form a basis of the subspace G of dimension n in the vector space of all continuous functions F over the interval (a,b). Therefore, any continuous function f(x) defined analytically in F over the interval (a,b) may be approximated to any desired degree of accuracy by a linear combination of the polynomials $P_0(x), P_1(x), \ldots, P_n(x)$ where n is not fixed.

When we speak of a Legendre Approximation in the sense of least squares, we shall mean an approximation of an analytic function defined over the interval (-1,1) by a polynomial approximation using the Legendre Polynomials as the orthogonal set of elements.

It will be possible to approximate f(x) over any interval (a,b) as long as it is finite since all that is needed to transform (a,b) to (-1,1) is an appropriate linear transformation of the independent variable. To illustrate, consider the function f(x) to be approximated to be a polynomial defined in the finite interval (-1,1). If f(x) is a polynomial and the degree of the approximating polynomial is equal to the degree of f(x), then the approximation will be an exact representation of the function itself since f(x) lies in the subspace spanned by the Legendre Polynomials.

<u>Example</u>:

Find a third degree polynomial approximation to a function $f(x) = x^3 + x^2 + x + 1$.

$$f(x) \approx y(x) = \sum_{i=0}^{3} a_i P_i(x), \quad a_r = \frac{2r+1}{2}\int_{-1}^{1} f(x)P_r(x)\,dx$$

$$a_0 = 4/3, \quad a_1 = 8/5, \quad a_2 = 2/3, \quad a_3 = 2/5$$

$$y(x) = 4/3\,P_0(x) + 8/5\,P_1(x) + 2/3\,P_2(x) + 2/5\,P_3(x)$$

$$= 4/3 + 8/5\,x + 2/3\left(1/2\,[3x^2 - 1]\right) + 2/5\left(1/2\,[5x^3 - 3x]\right)$$

$$= x^3 + x^2 + x + 1 \equiv f(x)$$

If the function f(x) cannot be assigned a degree, that is, it is not a polynomial, then it cannot be considered an element of a finite subspace of F. Hence, it cannot have an exact representation and must, in the true sense of the word, be represented by an approximation.

3.3 The Chebyshev Polynomial Approximation

The Chebyshev Approximation in the sense of least squares is used in approximating a function over a finite interval. This interval may be considered to be (-1,1) without any loss of generality since any finite interval may be transformed to the interval (-1,1) by a linear transformation. This approximation is used when the errors near the end points of the interval under consideration are of importance. To keep extreme errors down near the end points, a weighting function of the form $1/\sqrt{1-x^2}$ is used. The error associated with a Chebyshev Approximation will tend to oscillate with uniform amplitude over (-1,1). The error associated with the Legendre series resulting from least squares approximation with a uniform weighting function w(x) will tend to oscillate with an amplitude which increases toward the ends of the interval (-1,1) on the average.[5] The Legendre Approximation keeps the average square error down, but in doing so, isolated extreme errors are permitted. The error associated with a Chebyshev approximation allows a larger average square error to exist. The Chebyshev polynomials may be derived

---

[5] F. B. Hildebrand, INTRODUCTION TO NUMERICAL ANALYSIS (New York: McGraw-Hill Book Company, Inc., 1956), p. 392.

from the Chebyshev Differential Equation

$$(1-x^2)T_n'' - xT_n' + n^2 T_n$$

The resulting polynomial solution satisfies the conditions

$$\int_a^b \frac{T_i(x)\,T_j(x)}{\sqrt{(x-a)(x-b)}}\,dx = 0 \quad , \quad i \neq j$$

$$= \pi/2 \quad , \quad i = j$$

Hence $T_1(x), T_2(x) \; - - - - , T_n(x)$ may be considered to be the basis of a subspace of a linear normed vector space.

The recurrence relation of these polynomials

$$T_{n+1}(x) + T_{n-1}(x) = 2x\,T_n(x) \qquad n \geq 1$$

may be obtained from the trigonometric relation

$$\cos(n+1)\theta + \cos(n-1)\theta = 2\cos\theta\cos n\theta$$

where

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

3.4 The Laguerre Approximation

We are now concerned with approximating a function f(x) continuous and defined over a semi-infinite range (a,∞). A simple linear transformation of the independent variable will transform the interval over to the half-line $(0,\infty)$.

The series solution to the Laguerre Differential Equation

$$xL_n'' + (\alpha+1-x)L_n' + nL_n = 0$$

are

$$L_n(x) = x^{-\alpha} e^{x} \frac{d^n}{dx^n}\left(x^{n+\alpha} e^{-x}\right) \quad , \alpha > -1 \tag{3}$$

and are called the Laguerre Polynomials.

These polynomials have the following properties:

$$\int_0^\infty x^\alpha e^{-x} L_m(x) L_n(x)\, dx = 0$$

$$\int_0^\infty x^\alpha e^{-x} L_n^2(x)\, dx = \Gamma(n+1)\Gamma(n+\alpha+1)$$

The polynomials may be normalized and we get

$$\Pi_n(x) = \frac{L_n(x)}{\sqrt{\Gamma(n+1)\Gamma(n+\alpha+1)}}$$

so

$$\int_0^\infty x^\alpha e^{-x} \left[\Pi_n(x)\right]^2 dx = 1$$

Now an arbitrary function f(x) may be expressed in a series of Laguerre Polynomials if the function f(x) satisfies the following conditions:[6]

1. That the integral $\int_{\gamma}^{\infty} x^{\alpha} e^{-x} [f(x)]^2 dx$ exists for a certain γ.

2. That in case $-1 < \alpha \leq -1/2$ the integral $\int_0^{\beta} x^{\alpha} e^{-x} [f(x)]^2 dx$ exists for a certain β, while in case $\alpha > -1/2$ only the existence of the integral $\int_0^{\beta} x^{\alpha/2 - 1/4} |f(x)| dx$ is required.

3. That f(x) is of limited variation in a certain interval $x-\delta$, $x+\delta$ and absolutely integrable in any finite interval, then the series

$$A_0 \Pi_0(x) + A_1 \Pi_1(x) + \cdots \qquad x > 0$$

and

$$A_k = \int_0^{\infty} x^{\alpha} e^{-x} \Pi_k(x) f(x) dx$$

converges and has for its sum $\dfrac{f(x+0)+f(x-0)}{2}$

In approximating a function f(x) which is not a polynomial of finite degree by a finite Laguerre series,

$$f(x) \approx y_n(x) = \sum_{k=0}^{n} \left[\frac{1}{\Gamma(n+1)\Gamma(n+\alpha+1)} \int_0^{\infty} x^{-\alpha} e^{x} F(x) L_k(x) dx \right] L_k(x)$$

---

[6] J. V. Uspensky, "On the development of arbitrary functions in a series of Hermite's and Laguerre's polynomials," ANNALS OF MATH, (2) Vol. 28, (1927), pp. 593-619.

it will be noticed that the error in the approximation increases in magnitude very rapidly for values of x greater than the nth zero of $L_n(x)$. A technique that may be employed to extend the range of the approximation is to extend the largest zero of the nth Laguerre Polynomial.[7] This may be accomplished by replacing x with tx in Equation (3). The new Laguerre Polynomial will now be

$$L_n(tx) = (tx)^{-\alpha} e^{tx} \frac{d^n \left[(tx)^{n+\alpha} e^{-tx}\right]}{d(tx)^n}, \quad \alpha > -1$$

The value of t may be determined in the following manner. Suppose we wish to extend an mth degree approximation of f(x) so that it would be valid in an interval $0 \leqq x \leqq a$ where a is greater than the largest root of $L_n(x)$. If $x_{m,m}$ denotes the largest root of the mth Laguerre Polynomial, then we want to extend the mth root so that it exceeds a, that is,

$$\frac{x_{m,m}}{t} \geqq a$$

where $x_{m,m} \approx (4m+2) - 3.1715(4m+2)^{1/3} + 2.7551(4m+2)^{-1/3}$

---

[7] V. E. Spencer, "Zeros of the Generalized Laguerre Polynomial," DUKE MATH. JOURNAL, Vol. 3, (1937), pp. 667-675.

3.5 The Hermite Approximation

An approximation formula based on the principle of least squares that may be used to approximate functions defined over the entire space (- ∞, ∞) is called a Hermite Approximation.

A solution to the Hermite Differential Equation

$$\frac{d^2 H_n(x)}{dx^2} - 2x \frac{d H_n(x)}{dx} + 2n H_n(x) = 0$$

is the series

$$H_n(x) = (2x)^n - \frac{n(n-1)}{1!}(2x)^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2!}(2x)^{n-4} - \cdots$$

When n is equal to zero or a positive integer, we get the Hermite polynomials. They may be expressed in the form

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx}(e^{-x^2})$$

and may also be obtained from the recurrence relation

$$H_{n+1}(x) - 2x H_n(x) + 2n H_{n-1}(x)$$

These polynomials have the following properties :

$$\int_{-\infty}^{\infty} e^{-x^2} H_m H_n \, dx = 0 \qquad n \neq m$$

$$\int_{-\infty}^{\infty} e^{-x} H_n^2 \, dx = 2 \cdot 4 \cdot 6 \cdots 2n \sqrt{\pi}$$

To simplify the work in using these polynomials, we normalize them.

$$V_n(x) = \frac{H_n(x)}{\sqrt{2\cdot 4\cdot 6\cdots 2n\sqrt{\pi}}}$$

$$\int_{-\infty}^{\infty} e^{-x^2}(V_n(x))^2 dx = 1$$

$$\int_{-\infty}^{\infty} e^{-x^2}(V_n(x)V_m(x))\,dx = 0$$

The approximation to f(x) may now be written as

$$f(x) \simeq y_n(x) = \sum_{k=0}^{n}\left[\frac{1}{2^k k!\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-x^2} V_k(x) f(x)\,dx\right] V_k(x)$$

The approximation to a function f(x) will be of the form

$$S_n(x) = A_0V_0(x) + A_1V_1(x) + \cdots + A_nV_n(x)$$

The conditions assuring convergence of this series are:[8]

1. f(x) is absolutely integrable in any finite interval
2. f(x) is of limited variation in a certain interval $x-\delta,\ x+\delta$
3. the integrals $\int_{-\infty}^{\infty} e^{-x^2}[f(x)]^2 dx$ and $\int_{-\infty}^{-\alpha} e^{-x^2}[f(x)]^2 dx$ exist for certain α.

$S_n(x)$ will then converge and have for its sum

$$f(x+0) + f(x-0)\,/\,2$$

In case f(x) is a polynomial of degree n, the resulting $S_n(x) = f(x)$[8]

[8] Uspensky, op. cit.

3.6 The Discrete Case

When a function is defined by a set of observations, taken at equal intervals, and the data is biased when it is recorded, a least squares approximation may be used to represent this function. In order to obtain the best polynomial approximation, we must know how the function behaves in the interval of interest. This enables us to choose the degree of the polynomial approximation which would fit the function between the data points.

The requirement that the sum of the squared errors be a minimum results in a least squares fit which will give preference to many smaller errors rather than to a few larger ones.

Corresponding to the orthogonal set of functions used in the previous approximations of an analytically defined function are a set of orthogonal functions for the discrete case. These functions are orthogonal under summation over the discrete range. One of the most common sets is the set corresponding to the Legendre Polynomials.** Another set is the set analogous to the Hermite Polynomials.* These polynomials are used if we have data whose values tend to zero at both ends.

---

* See Bibliography
** See Bibliography

3.7 Fourier Series Approximation

If a function f(x) is such that $\int_{-\pi}^{\pi} f(t)\,dt$ exists as a Riemann integral, it may be approximated by a Fourier Series, provided that f(x) is defined and bounded in the range $(-\pi, \pi)$ and if f(x) has only a finite number of maxima and minima and a finite number of discontinuities in this range. The function must be also periodic outside the range $(-\pi, \pi)$. This is Dirichlet's Theorem.

The Fourier Series used in the above circumstances will be of the form

$$S_n(x) = \frac{1}{2}a_0 + \sum_{i=1}^{\infty}(a_n \cos nx + b_n \sin nx)$$

where $a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t) \cos nt\,dt$, $b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t) \sin nt\,dt$

The above series converges to the sum $\frac{1}{2}\left\{f(x+0) + f(x-0)\right\}$

The greatness of the Fourier Series is that with it we are able to represent a function which cannot be defined by means of a single, closed formula in terms of the so-called elementary functions alone. These functions may have a finite number of discontinuities in the interval of interest and yet, they may be represented by a single expression.

ABSTRACT

This paper, utilizing the properties of vector spaces, describes an approach to polynomial approximations of functions defined analytically or by a set of observations over some interval. If the function and its approximation are both considered to be elements of a linear normed vector space, a weighted sum or integral of the square of the discrepancy between the function and its approximation is to be a minimum. When this condition is satisfied, and depending upon the interval of interest, the polynomial approximation to the function becomes either the Legendre, Chebyshev, Laguerre, or Hermite approximation formulas.

An investigation into the properties and applications of these formulas is included, and it is shown that these formulas give the best polynomial approximation to certain functions in the sense of least squares.

BIBLIOGRAPHY

Achieser, N. I., THEORY OF APPROXIMATION. New York: Frederick Ungar Publishing Co., 1956.

Birkhoff & MacLane, A SURVEY OF MODERN ALGEBRA. New York: The Macmillan Co., 1953.

Courant & Hilbert, METHODS OF MATHEMATICAL PHYSICS, Vol. I. New York: Interscience Publishers, Inc., 1953.

**Hildebrand, F. B., INTRODUCTION TO NUMERICAL ANALYSIS. New York: McGraw-Hill Book Co., Inc., 1956.

Jackson, D., FOURIER SERIES & ORTHOGONAL POLYNOMIALS, The Carus Mathematical Monographs. The Mathematical Association of America, 1941.

-----, THE THEORY OF APPROXIMATION. American Mathematical Society Colloquium Publications, Vol. XI, 1930.

*Milne, W. E., NUMERICAL CALCULUS. New Jersey: Princeton University Press, 1949.

Rudin, W., PRINCIPLES OF MATHEMATICAL ANALYSIS. New York: McGraw-Hill Book Co., Inc., 1953.

Scarborough, J. B., NUMERICAL MATHEMATICAL ANALYSIS. Baltimore: The Johns Hopkins Press, 1955.

Shilov, G. E., AN INTRODUCTION TO THE THEORY OF LINEAR SPACES. New Jersey: Prentice Hall, 1961.

Spencer, V. E. "Zeros of the Generalized Laguerre Polynomial," DUKE MATH. JOURNAL, Vol. 3, 1937.

Todd, J., SURVEY OF NUMERICAL ANALYSIS. New York: McGraw-Hill Book Co., Inc., 1962.

Uspensky, J. V., "On the Development of Arbitrary Functions in a Series of Hermite's & Laguerre's Polynomials," ANNALS OF MATH (2) Vol. 28, 1927.