THE LEAST SQUARES APPROXIMATION

P. VANICEK
D. E. WELLS

October 1972

LECTURE NOTES
NO. 22

THE LEAST SQUARES APPROXIMATION

P. Vanicek
D. E. Wells

Department of Geodesy and Geomatics Engineering
University of New Brunswick
P.O. Box 4400
Fredericton, N.B.
Canada
E3B 5A3

October 1972

PREFACE

In order to make our extensive series of lecture notes more readily available, we have scanned the old master copies and produced electronic versions in Portable Document Format. The quality of the images varies depending on the quality of the originals. The images have not been converted to searchable text.

PREFACE

The authors would like to thank Mr. Charles Merry for providing the solved problems appearing at the end of some chapters.

The junior author would like to acknowledge that the geometrical interpretations of Chapter 4 had their genesis in discussions he has had with Professor William R. Knight of the UNB Departments of Mathematics and Computer Science.

CONTENTS

TABLE OF CONTENTS CONTINUED

1. INTRODUCTION

1.1 Notation Conventions

Since sets, vectors and matrices occur together in some of the discussions in these notes, the following convention is used (there are exceptions):

The elements of a SET are enclosed in BRACES, e.g. $A \equiv \{a,b,c\}$

The elements of a VECTOR are enclosed in PARENTHESES, e.g. $\vec{A} \equiv (a,b,c)$

The elements of a MATRIX are enclosed in BRACKETS, e.g. $A \equiv \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$

Other notation conventions, which may be unfamiliar to some readers are:

SYMBOLISM	MEANING
$\{X_1,\ldots X_n\}$	"The discrete set whose only elements are X_1, X_2,...X_n"
$\{X_i\}$	"The discrete set of all elements X_i" (The rule for i is specified elsewhere and understood here)
$A \equiv \{X_i\}$	"A is the discrete set of all elements X_i"
$A \equiv [a,b]$	"A is the compact set of all x such that $a \leq x \leq b$"
$A \equiv [a,b)$	"A is the compact set of all x such that $a \leq x < b$"
$A \equiv (a,b]$	"A is the compact set of all x such that $a < x \leq b$"
$A \equiv (a,b)$	"A is the compact set of all x such that $a < x < b$"
$x \in A$ or $A \ni x$	"x is an element of the set A" or equivalently "A contains the element x"
$y \notin A$ or $A \not\ni y$	"y is not an element of the set A" or equivalently "A does not contain the element y"
$A \subset B$ or $B \supset A$	"A is a subset of B" or equivalently "B contains A"

$\{A \to B\}$	"the set of all mappings from the set A to the set B" (A mapping relates one and only one element of B to each element of A).
$f \in \{A \to B\}$	"f is a mapping of A to B". (If A and B are numerical sets then f is called a function).
$f(x)$	"the functional value of the function f" ($x \in A$ is the argument, $f \in \{A \to B\}$ is the function, and $f(x) \in B$ is the functional value. A is the domain or definition set of f, and B is the range or image set of f).
E	"the set of all real numbers"
E^+	"the set of all positive real numbers"
E_n or E^n	"the set of all ordered n-tuples of real numbers"
${}_nA_m$	"the matrix A which has n rows and m columns"
$[A \vdots B]$	"the augmented matrix formed by the column vectors of the matrices A and B" (A and B must have equal numbers of rows).
A^T	"the transpose of the matrix A" (found by interchanging rows and columns of A).
A^{-1}	"the inverse of the matrix A" (the inverse exists if and only if A is nonsingular, that is its determinant is nonzero).
I	"the identity matrix" (having diagonal elements equal to unity, off-diagonal elements equal to zero).
$\Rightarrow$	"implies"
iff or $\leftrightarrow$	"if and only if"

Some of the functions which are used in these notes are:

$\sum_{x \in M} f(x)$	"the sum of all functional values of f, evaluated over the domain M"
$\prod_{x \in M} f(x)$	"the product of all functional values of f, evaluated over the domain M"
$\max_{x \in M} f(x)$	"the maximum functional value of f, evaluated in the domain M"
$\min_{x \in M} f(x)$	"the minimum functional value of f, evaluated in the domain M"

$\int_M f(x)dx$	"the integral of f over the domain M"
$\lim_{x \to a} f(x)$	"the limit of the functional value of f, evaluated as x approaches a"
$\|\| f \|\|$	"the norm of the function f" (see page 7)
$< f,g >$	"the scalar product of the functions f and g" (see page 17)
$\| f \|$	"the absolute value of the function f" (see page 6)
$\rho(f,g)$	"the distance between the functions f and g" (see page 6)
$\sqrt{f} = \sqrt{(f)} = (f)^{1/2}$	"the square root of the function f" (the functional value of the square root function is always non-negative)
$\underline{=}$	"equality in the mean sense" (see page 51)

1.2 Theory of Approximation

With the increasing use of computers, the importance of the theory of approximation increases, too. It has ceased to be a domain for pure/or applied mathematicians and has crept into all kinds of fields. It is not only the vital part of numerical analysis, but is used whenever we have to deal with functional relations and their numerical representation. Hence the theory of approximation has become an indivisible part of all experimental sciences and all branches of engineering.

The problem of approximation can be defined as follows: given a function F, defined on a set M (compact or discrete), find another function of a prescribed general form that would represent the given function in a specified way. The approximating function can be represented, without any loss of generality, as a "generalized polynomial":

$$P_n = \sum_{i=1}^{n} c_i \, \phi_i$$

where $c_i \varepsilon E$ (E is the set of real numbers) are known as the coefficients of the polynomial and ϕ_i are the prescribed functions. The degree of the approximating generalized polynomial may even grow beyond all limits if we want. The set of the prescribed functions

$$\Phi \equiv \{\phi_1, \phi_2, \ldots, \phi_n\}$$

has to have certain properties for certain approximations, as we shall see later. The individual functions ϕ_i, and therefore even the polynomial P_n, may be functions of one, two or n-variables.

There are three distinctly different categories of approximations, namely

i) point approximations;

ii) approximations with prescribed properties;

iii) best interval approximations.

i) The point approximations seek to approximate the given function F at a specified point(x_o say) in a prescribed way. The best known technique here is that of Taylor (McLaurin) that seeks the point approximation to an analytic F such as to have as many common derivatives with F as wanted. With infinitely many common derivatives, it represents F on any interval M, providing F is on M continuous. P_n in this case is an algebraic polynomial or a power series, i.e., $\phi_i = x^i$. We are not going to devote our attention to this type of approximation since it is sufficiently known.

ii) These are various techniques based on various ideas. The most widely used of these approximations is the approximation using "spline functions", known as spline approximation. The properties of the approximating function is that it approximates the given F in sections having common tangent (or even derivatives of higher degrees) in the points where the sections join. This kind of approximation has recently become quite popular because of its ability to lend itself to an easy physical interpretation. To venture into these approximations is not the purpose of this course.

iii) The last and by far the most widely used is the best approximation on an interval. Because of its importance we shall devote a whole section to it.

1.3 Best Approximation on an Interval

The problem can be generaly formulated as follows: for the given function F defined on M find such a P_n -- for a previously selected Φ -- that has the smallest distance from F.

What is the distance between F and P_n? In order to be able to answer this question we have first to define the space into which we are going to measure. It will be the set of all possible functions defined on the same set M of arguments as F. Such a set (let us denote it by G_m) is known as functional space. Note that $F \in G_m$, too.

Once we have got the space to measure the distance in, we can proceed to define the distance. It can be shown that any function

$$\rho(G,H)$$

that maps a two-tuple of functions G, H from our functional space G_m, onto the set of real numbers can be used to measure the distance, providing it satisfies the following three conditions:

i) $\rho(G,H) \geq 0 \qquad (\rho(G,G) = 0)$ (non-negativeness)

ii) $\rho(G,H) = \rho(H,G)$ (symmetry)

iii) $\rho(G,H) \leq \rho(G,E) + \rho(H,E)$ (triangle rule),

where $G,H,E \in G_m$. These conditions are known as the axioms for distance (or metric).

It can be seen, that the above axioms are satisfied for a large family of functions. In practice mainly the following two metrics are used:

i) $\rho(G,H) = \max_{X \in M} |G-H|$,

known as the uniform metric and

ii) $\rho(G,H) = ||G-H||$,

where $||F||$ is the norm of F, is known as the mean-quadratic or least-squares metric. The use of the first metric leads to uniform (minimax, Tchebyshev) approximation while the second leads to mean-quadratic or least-squares approximation.

1.4 Problems

1. Which sets are finite? Which sets are compact?

   The people living on the earth.

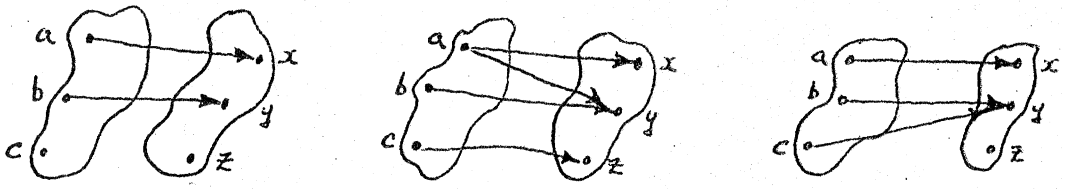
   All integers between 1 and 100.

   All integers.

   $A \equiv [0, 1]$.

   $A \equiv \{0, 1\}$.

2. Given sets $A \equiv \{a, b, c\}$ and $B \equiv \{x, y, z\}$ which diagrams define functions?



3. Given sets $A \equiv \{a, b, c\}$ and $B \equiv \{0, 1\}$ how many different functions are there from A into B, and what are they (diagrammatically)?

4. List the domain and range for the 24 goniometric functions (trigonometric, hyperbolic, and their inverses).

2. UNIFORM (TCHEBYSHEV) APPROXIMATION

Tchebyshev has proved that the best fitting P_n to any F in the uniform sense (the P_n that has the smallest uniform distance) always exists on any M. This proof is based on two Weierstrass' theorems telling us that there is always an algebraic or trigonometric polynomial P_n that approximates the given function F with an error $\varepsilon = |F - P_n|$ smaller than a required value. The degree n of P_n depends on the required value.

On the other hand, there is no satisfactory method that would solve the problem of finding such a P_n for any Φ and F. Tchebyshev himself has come up with a solution to the seemingly nonsensical problem: "given $F(\Phi) = 0$ on $M \equiv [-1,1]$ and $\Phi \equiv \{1,X,\ldots,X^n\}$, find the uniformly best fitting P_{n+1} under the restriction $c_{n+1} = 1$". (Note that without the restriction $c_{n+1} = 1$ the P_{n+1} (best fitting) would be identically 0). He found that such a P_{n+1} is given by following relations:

$$P_{n+1}(X) = T_n(X) = \frac{1}{2^{n-1}} \cos (n \arccos X) \quad n \geq 1$$

$$T_o(X) = 1 \ . \ *)$$

---

*) Note that the degree of an algebraic polynomial, as it is generally known, differs by 1 from the degree of a generalized polynomial. The algebraic polynomials with coefficients by the highest degree term equal to one are known as normalized algebraic polynomials. Note also that T_n for n even is even and for n odd is odd.

Let us agree, from now on that the algebraic polynomials containg X^n will be denoted by subscript n rather than n+1. These algebraic polynomials became consequently known as Tchebyshev polynomials (of the 1st kind).

The same problem can be formulated for $M \equiv [a,b]$ which, after the linear transformation

$$Z = \frac{1}{2}(a + b + (b-a)X)$$

(show that this linear transformation transforms the interval $[-1,1] \ni X$ to the interval $[a,b] \ni Z$) leads to a new system of polynomials

$$T_n(Z) = \frac{(b-a)^n}{2^n} T_n(X) \quad .$$

The quoted solution to the nonsensical looking problem permits to solve two different categories of problems:

i) Tchebyshev Economization

It can be shown that the best uniform approximation of the function $F(X) = X^n$ on $M \equiv [-1,1]$ by an algebraic polynomial $\tilde{\tilde{T}}_{n-2}$ (note that n-2 is the degree of the algebraic polynomial, i.e., $\tilde{\tilde{T}}_{n-2}$ contains n-1 terms) is given by

$$\tilde{\tilde{T}}_{n-2} = X^n - T_n(X) \quad .$$

The maximum error involved is 2^{1-n}. This fact can be used in lowering the degree of a given algebraic polynomial by one with the minimum uniform error.

Example: $P_4(X) = 1 - X + \frac{X^2}{2!} - \frac{X^3}{3!} + \frac{X^4}{4!}$ is to be lowered by one degree. The best uniform approximation to X^4, $\tilde{T}_2(X)$ is given by:

$$\tilde{T}_2(X) = X^4 - T_4(X) = X^4 - (X^4 - X^2 + \frac{1}{8}) = X^2 - \frac{1}{8} \quad .$$

Hence $$P_4(X) \simeq 1 - X + \frac{X^2}{2!} - \frac{X^3}{3!} + \frac{1}{4!}\left(X^2 - \frac{1}{8}\right)$$

$$= \left(1 - \frac{1}{8.4!}\right) - X + \left(\frac{1}{2!} + \frac{1}{4!}\right) X^2 - \frac{X^3}{3!} \quad .$$

The error involved in the approximation is smaller than $\frac{1}{4!}\, 2^{1-4} = \frac{1}{2^3 4!} =$

$= \frac{1}{192} \simeq 0.005$ on $M \equiv [-1,1]$.

Similar treatment can be designed even for $M \equiv [a,b]$ using the variable Z as described above.

This trick can be used when we want to represent a function in the form of truncated power series. The representation is more uniform when we truncate the series further on and lower the degree of such polynomial than if we truncate the series at the wanted place and leave it as it is.

i) Tchebyshev Algebraic Interpolating Polynomial

If we have a continuous function F and want to interpolate it on [a,b] using an algebraic interpolation polynomial of a prescribed degree n, the interpolation will give the least residual if we chose for interpolation nodes the roots of the Tchebyshev polynomial. The maximum error is then

$$R_{n+1} = \frac{(b-a)^{n+1}}{2^{2n+1}\,(n+1)\,!} \max_{X \in [a,b]} \left|f^{(n+1)}(X)\right|$$

2.1 Problems

Solved Problems

1. Tchebyshev polynomials. Compute the Tchebyshev polynomials of the first kind, $T_n(x)$, for degrees n from 0 to 6.

Solution

Tchebyshev polynomials of the first kind are given by

$$T_n(x) = \cos(n \arccos x).$$

Let $x = \cos\theta$, or $\theta = \arccos x$. Then

$$T_n(x) = T_n(\cos\theta) = \cos n\theta.$$

But

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

therefore

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

or

$$T_n(x) = 2x\,T_{n-1}(x) - T_{n-2}(x).$$

Now for n = 0 and n = 1

$$T_0(x) = \cos(0) = 1$$
$$T_1(x) = \cos(\arccos x) = x.$$

Hence using the recursion relation we obtain

$$T_2(x) = 2x^2 - 1$$
$$T_3(x) = 4x^3 - 3x$$
$$T_4(x) = 8x^4 - 8x^2 + 1$$
$$T_5(x) = 16x^5 - 20x^3 + 5x$$
$$T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1.$$

2. Roots of Tchebyshev Polynomials. Compute numerical values for the roots of the Tchebyshev polynomials of the first kind, $T_n(x)$, for degrees n from 2 to 6.

Solution.

For the roots x_i of $T_n(x)$

$$T_n(x_i) = \cos\ (n \arccos x_i) = 0, \quad x_i \in [-1, 1]$$

$$n \arccos x_i = \arccos 0 = (2i - 1)\frac{\pi}{2} \quad i = 1, 2, \ldots n$$

$$\arccos x_i = \frac{(2i - 1)}{2n}\ \pi$$

$$x_i = \cos\left(\frac{(2i - 1)\pi}{2n}\right) .$$

There will be n distinct roots x_i, for each of i = 1, 2, ..., n.

n	x_i
2	$\pm \cos 45°$
3	$0, \pm \cos 30°$
4	$\pm \cos 22\frac{1}{2}°$, $\pm \cos 67\frac{1}{2}°$
5	$0, \pm \cos 18°, \pm \cos 54°$
6	$\pm \cos 15°, \pm \cos 45°, \pm \cos 75°$

3. Tchebyshev Economization. Represent the function cos z, $z \in [-\pi, \pi]$, by the truncated power series $P_n(z) = \sum_{k=0}^{n} (-1)^k \frac{z^{2k}}{(2k)!}$

Represent the same function again by applying Tchebyshev economization to the last two terms of the truncated power series $P_{n+2}(z)$, obtaining a second polynomial of degree n. Compare these two approximating polynomials, taking $P_{n+2}(z)$ as the standard.

Solution.

The function cos z, $z \in [-\pi, \pi]$ is transformed to cos x, $x \in [-1,1]$ by $z = (1/2)(a + b + (b - a)x) = (1/2)(-\pi + \pi + (\pi + \pi)x) = \pi x$

and thus $P_n(x) = \sum_{k=0}^{n} (-1)^k \frac{(\pi x)^{2k}}{(2k)!}$, $x \in [-1,1] \equiv M$

and $$P_{n+2}(x) = P_n(x) + \frac{(-1)^{n+1} (x)^{2(n+1)}}{[2(n+1)]!} + \frac{(-1)^{(n+2)} (x)^{2(n+2)}}{[2(n+2)]!}$$

Thus $$\left| P_n(x) - P_{n+2}(x) \right| = \left| \frac{(-1)^{(n+1)} (\pi x)^{2(n+1)}}{[2(n+1)]!} + \frac{(-1)^{(n+2)} (\pi x)^{2(n+2)}}{[2(n+2)]!} \right|$$

$$= \frac{(\pi x)^{2(n+1)}}{[2(n+1)]!} \left(1 - \frac{(\pi x)^2}{(2n+4)(2n+3)} \right)$$

So that

$$\max_{x \in M} \left| P_n(x) - P_{n+2}(x) \right| = \frac{\pi^{2(n+1)}}{[2(n+1)]!} \left(1 - \frac{\pi^2}{(2n+4)(2n+3)} \right)$$

Let $\tilde{P}_n(x)$ be the n-th degree polynomial resulting from Tchebyshev economication of the (n+1) and (n+2) terms of $P_{n+2}(x)$. Then

$$\max_{x \in M} \left| x^{2(n+1)} - T_{2(n+1)}(x) \right| = \frac{1}{2^{2n+1}}$$

$$\max_{x \in M} \left| x^{2(n+2)} - T_{2(n+2)}(x) \right| = \frac{1}{-2^{2n+3}}$$

Hence $$\max_{x \in M} \left| \tilde{P}_n(x) - P_{n+2}(x) \right| =$$

$$= \left| \frac{(-1)^{(n+1)} \quad 2(n+1)}{[2(n+1)]\ !} \quad \frac{1}{2^{2n+1}} + \frac{(-1)^{n+2} \quad 2(n+2)}{[2(n+2)]\ !} \quad \frac{1}{2^{2n+3}} \right|$$

$$= \frac{\pi^{2(n+1)}}{[2(n+1)]\ !} \left(1 - \frac{\pi^2}{(2n+4)\ (2n+3)} \quad \frac{1}{4} \right) \left(\frac{1}{2^{2n+1}} \right)$$

And we conclude that $P_{n+2}(x)$ is better approximated by $\tilde{P}_n(x)$ than by $P_n(x)$ by a factor of about $2^{-(2n+1)}$.

4. Tchebyshev Algebraic Interpolating Polynomial. Design an algorithm for computing the Tchebyshev algebraic interpolating polynomial of degree n for a given function $F(z)$, $z \in [a,b]$.

Solution.

A function $F(z)$ defined on the domain $[a,b]$ can be approximated by algebraic polynomials of degree n

$$P_n(z) = \sum_{k=o}^{n} c_k z^k.$$

If the cofficients c_k are chosen such that for (n+1) given basepoint values z_i

$$P_n(z_i) = f(z_i)$$

then $P_n(z)$ is called an algebraic interpolating polynomial for $F(z)$. The Tchebyshev algebraic interpolating polynomial is the one in which the basepoints z_i are chosen to be distributed over $[a,b]$ in the same way that the (n+1) roots x_i of the Tchebyshev polynomial of the first kind, $T_{n+1}(x)$, are distributed over $[-1,1]$.

Given a function $F(z)$, its domain $[a,b]$, and the degree of polynomial desired, n, the algorithm to compute the Tchebyshev algebraic interpolating polynomial contains the following steps:

a) Compute the (n+1) roots x_i of $T_{n+1}(x)$ (see problem 2);

b) Compute the equivalent values for z_i from

$$z_i = (1/2)(a + b + (b - a)\, x_i);$$

c) Solve the system of (n+1) linear equations

$$P_n(z_i) = \sum_{k=o}^{n} c_k z_i^k = F(z_i), \quad i=1,2,\ldots n+1$$

for the (n+1) coefficients $c_o, c_1, c_2 \ldots c_n$;

d) The Tchebyshev algebraic interpolating polynomial is then

$$P_n(z) = \sum_{k=o}^{n} c_k z^k.$$

Unsolved Problems.

5. Express the powers of x in terms of the Tchebyshev polynomials of the first kind, from x^0 to x^6.

6. Apply the results of problem 3 to a specific case. Choose a value for n. Should n be even or odd or does it matter? Compute numerical values for the (n+1) coefficients of $P_n(z)$ and $\tilde{P}_n(z)$. Compare the "true" values of cos z to each of these approximations for several values of z. Which approximation has the maximum error for values in the interval $[-\pi, \pi]$? For values in the interval $[-\pi/2, \pi/2]$?

7. Apply the results of problem 4 to a specific case. Let $F(z) = \sin z$, $z \in [0, \pi/2]$. Should the algebraic polynomial of degree n for this function contain only odd powers of z, only even powers of z, all powers of z, or does it matter? Choose a value for n such that the Tchebyshev algebraic interpolating polynomial approximates $F(z)$ to better than 1.1×10^{-5}. Compute the coefficients for such a polynomial. Write a computer program to compute sin z using this polynomial, using the method of nested mutiplication (Horner's Method).

8. Choose any problem from chapters 2 or 3 of Cheney [1966] and solve it.

3. LEAST SQUARES APPROXIMATION

3.1. Norm And Scalar Product

We call the norm of $G \in G_m$ a function $||G||$ that maps the element of G_m on E and satisfies the following axioms:

i) $||G|| \geq 0$ ($||G|| = 0$ if and only if $G(x) \equiv 0$ on M)

ii) $||\lambda G|| = |\lambda| \, ||G||$ for $\lambda \in E$

iii) $||G+H|| \leq ||G|| + ||H||$, $G,H \in G_m$.

A space G_m in which a norm is defined is said to be a normed space.

The norm in the least-squares approximation is in practice defined in two different ways

$$||G|| = \sqrt{\sum_{X \in M} W(X) \, G(X)^2} \in E^+$$

for discrete M and

$$||G|| = \sqrt{\int_M W(X) \, G(X)^2 \, dX} \in E^+$$

for compact M and F integrable in Riemann's sense (all the integrals involved in our development are considered Riemann's). Conceivably, the integrals could be taken as defined in Lebesgue's sense or other just as well.) The function W that has to be non-negative on M is known as weight function. Note that the Tchebyshev's distance $\rho(G,H)$ can also be considered as a norm $||G-H||$.

Note that the distance $||H-G||$ on discrete M is Eucleidean:

$$\rho(H,G) = \sqrt{\sum_{X\in M} W(X)\cdot(H(X)-G(X))^2} \in E^+$$

i.e., describes the distance in the common Eucleidean geometric space with axes X_i scaled inversely proportionate to $W(X_i)$.

In order to simplify this notation in the forthcoming development, let us define one more operation on the functional space G_m, the scalar product of two functions. If $G,H \in G_m$ then

$$\langle G,H\rangle = \begin{cases} \sum_{X\in M} W(X)\cdot G(X)\cdot H(X) \\ \int_M W(X)\cdot G(X)\cdot H(X)\, dX \end{cases}$$

is known as scalar product for M discrete or compact respectively. The functional space on which scalar product is defined is known as Hilbertian functional space. (Generally the scalar product can be defined in infinitely many ways providing it satisfies the axioms for scalar product that are somewhat similar to the axioms of metric. In practice though the two definitions are used almost exclusively. They are closely related to the selected metric of the functional space.)

For any two $G,H \in G_m$ the Schwartz's inequality

$$\langle G,G\rangle\cdot\langle H,H\rangle \geq \langle G,H\rangle^2$$

is satisfied. This is, of course, true for both definitions of the scalar product.

If for two functions $G,H\in G_m$ the scalar product is zero, they are known as orthogonal. Consequently, if for a system of functions $\Phi \subset G_m$ the following equation is valid

$$\langle\phi_i,\phi_j\rangle = \begin{cases} K_i \neq 0 & i=j \\ 0 & i\neq j \end{cases} \qquad i,j = 1,2,\ldots,n$$

the system is known as a system of mutually orthogonal functions or simply an orthogonal system. Note that

$$K_i = \langle \phi_i, \phi_i \rangle = ||\phi_i||^2 \in E^+ ,$$

the square of the norm of ϕ_i. The above equation is often written using the Kronecker δ, a symbol defined thus

$$\delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j . \end{cases}$$

Using this notation, we can rewrite the condition for orthogonality of $\Phi \in G_m$ as

$$\langle \phi_i, \phi_j \rangle = ||\phi_i||^2 \delta_{ij} \qquad i,j = 1,2,\ldots,n.$$

If, in addition to the orthogonality, the norms of all the ϕ's equal to 1, the system $\underline{\Phi}$ is said to be orthonormal. For such a system we have

$$\langle \phi_i, \phi_j \rangle = \delta_{ij} \qquad i,j = 1,2,\ldots,n.$$

Note that the orthogonality (orthonormality) depends not only on Φ but M and W as well. Hence we may have orthogonal (orthonormal) systems on one M and not on another M'. We can also note that an orthogonal system can be always orthonormalized by dividing the individual functions of the system by their norms.

3.2 Base Functions

Providing the prescribed functions $\Phi \equiv \{\phi_1, \phi_2, \ldots, \phi_n\}$ are linearly independent on G_m, we talk about the base Φ. If and only if Φ is a base, the coefficients of the best fitting polynomial P_n can be uniquely determined. More will be said about it later.

The necessary and sufficient conditon for $\phi_1, \phi_2, \ldots, \phi_n$ to be linearly independent on G_m is

$$\sum_{i=1}^{n} \lambda_i \phi_i (X) = 0 \qquad X \in M, \quad \lambda_i \in E$$

if and only if all the λ's equal to zero. When the ϕ's create a base the following determinant, known as Gram 's determinant,

$$G(\Phi) = \det (\langle \phi_i, \phi_j \rangle) = \begin{vmatrix} \langle \phi_1, \phi_1 \rangle, \langle \phi_1, \phi_2 \rangle, \ldots, \langle \phi_1, \phi_n \rangle \\ \langle \phi_2, \phi_1 \rangle, \langle \phi_2, \phi_2 \rangle, \ldots, \langle \phi_2, \phi_n \rangle \\ \vdots \qquad \vdots \qquad \vdots \\ \langle \phi_n, \phi_1 \rangle, \langle \phi_n, \phi_2 \rangle, \ldots, \langle \phi_n, \phi_n \rangle \end{vmatrix}$$

is different from zero. It can be shown that the definition of linear dependence (independence) using the linear combination and using Gram 's determinant are equivalent. Let us point out here that if M is a discrete set of m points, the necessary condition (not sufficient!) for Φ to be linearly independent on G_m is that n be smaller or equal to n.

We may note at this point that the Gram 's determinant of an orthogonal system of functions is given by

$$G(\Phi) = \prod_{i=1}^{n} ||\phi_i||^2 .$$

Similarly, an orthonormal system has got its Gram 's determinant equal to

$$G(\Phi) = \prod_{i=1}^{n} 1 = 1 .$$

Hence we can see that an orthogonal (orthonormal) system cannot be linearly dependent. If one or more of the norms equal to zero then, according to defintion, the system ceases to be orthogonal. This will be the case when the number of functions is larger than the number of points in M, if M is a discrete set.

3.3 Normal Equations

According to the general formulation of the problem of the best approximation on an interval we shall now be looking for such a generalized polynomial P_n, i.e., its coefficients $c_1, c_2, \ldots, c_n$, that would make the distance $||F-P_n||$ the minimum.

To begin with let us take M to be discrete. This means that we should minimize the Eucleidean distance $\sqrt{\sum_{X\in M} W(X)\ (F(X) - P_n(X))^2}$ with respect to $c_1, c_2, \ldots, c_n$. Here, the minimum of the square root obviously occurs for the same argument as the minimum of the function under the square root sign. We may therefore write the condition as

$$\min_{c_1,\ c_2,\ldots c_n \in E} \rho^2(F,P_n) = \min_{c_1,c_2,\ldots c_n \in E} \sum_{X\in M} W(X)\ (F(X) - P_n(X))^2$$

$$= \min_{c_1,c_2,\ldots \varepsilon_n \in E} \sum_{X\in M} W(X)\ (F(X) - \sum_{i=1}^{n} c_i\ \phi_i\ (X))^2$$

The extreme (minimum or maximum) of the above sum occurs if and only if all its partial derivatives with respect to the individual c's go to zero. Since the maximum, for F(X) finite, is achieved for only infinitelylarge values of c's then we may see that any finite solution (values of c's) we get from the partial derivatives equated to zero will furnish the minimum distance rather than maximum.

Carrying out the operation of minimization we get:

$$\frac{\partial}{\partial c_i} \sum_{X \in M} [W(X)\ (F(X) - \sum_{j=1}^{n} c_j\ \phi_j\ (X))^2]$$

$$= 2 \sum_X [W(X)\ (F(X) - \sum_j c_j\ \phi_j\ (X))\ \frac{\partial}{\partial c_i}\ (- \sum_j c_j\ \phi_j\ (X))]$$

$$= -2 \sum_X [W(X)\ (F(X) - \sum_j c_j\ \phi_j\ (X))\ \phi_i\ (X)]$$

$$= -\ 2 \sum_X [W(X)\ (F(X)\ \phi_i\ (X) - \sum_j c_j\ \phi_j\ (X)\ \phi_i\ (X))]$$

$$= -\ 2 \sum_X W\ (X)\ F(X)\ \phi_i\ (X) + 2 \sum_X W(X) \sum_j c_j\ \phi_j\ (X)\ \phi_i\ (X) = 0 \qquad i=1,2,\ldots,n.$$

Here the 2's can be evidently discarded and we get

$$\sum_X W(X)\ F(X)\ \phi_i\ (X) = \sum_X W(X) \sum_j c_j\ \phi_j\ (X)\ \phi_i\ (X) \qquad i = 1,2,\ldots,\ n.$$

But the left hand side is nothing but $[F,\phi_i]$ and the right hand side can be rewritten as

$$\sum_X W(X) \sum_j c_j\ \phi_j\ (X)\ \phi_i\ (X) = \sum_j c_j \sum_X W(X)\ \phi_j\ (X)\ \phi_i(X) = \sum_j c_j\ \langle \phi_j, \phi_i \rangle\ .$$

Hence we end up with the system of linear algebraic equations for $c_1, c_2, \ldots, c_n$:

$$\sum_{j=1}^{n} \langle \phi_i, \phi_j \rangle\ c_j = \langle F, \phi_i \rangle \qquad i = 1, 2,\ldots,\ n.$$

This system of equations is known as the <u>normal equations</u>.

If we consider the mean-quadratic distance in G_m for M compact:

$$\rho(F-P_n) = \sqrt{\int_M W(X)\ (F(X) - P_n(X))^2\ dX}$$

we end up with exactly the same system of equations with the only difference that the scalar products $\langle \phi_i, \phi_j \rangle$, $\langle F, \phi_i \rangle$ will be defined as

$$\langle \phi_i, \phi_j \rangle = \int_M W(X) \cdot \phi_i(X) \cdot \phi_j(X)\ dX, \quad \langle F, \phi_i \rangle = \int_M W(X) \cdot F(X) \cdot \phi_i(X)\ dX.$$

The proof of this is left to the reader. Hence if we use only the general notation for the norm and the scalar product both developments remain exactly identical.

3.4 Remarks On The Normal Equations

We can, first of all, note that the determinant of the matrix of normal equations

$$A = \left[\langle \phi_i, \phi_j \rangle\right]$$

is nothing else but our old acquaintance from 3.2, the Gram's determinant:

$$G(\Phi) = \det\ (A) = \det\ \left(\langle \phi_i, \phi_j \rangle\right) .$$

Now, we can see the importance of the requirement that Φ be linearly independent on G_m. The linear independence of Φ insures that $G(\Phi) \neq 0$ and therefore the matrix A has an inverse A^{-1}. Hence the system of normal equations has the unique solution

$$\vec{c} = A^{-1}\ \langle F, \vec{\phi} \rangle .$$

We may note that if the system Φ is orthogonal $A = \text{diag}\left[\langle \phi_i, \phi_i \rangle\right] = \text{diag}\ (||\phi_i||^2)$ and the system of normal equations

degenerates into an n−tuple of equations containing only one coefficient each:

$$||\phi_i||^2 \, c_i = \langle F, \phi_i \rangle \qquad i = 1, 2, \ldots, n.$$

The solution then is trivial:

$$c_i = \langle F, \phi_i \rangle / ||\phi_i||^2 \qquad i = 1, 2, \ldots, n.$$

For an orthonormal system Φ, the normal equations degenerate still further and we get:

$$c_i = \langle F, \phi_i \rangle \quad i = 1, 2, \ldots, n.$$

3.5 Solution Of The Normal Equations

We may wish, for a certain class of problems, to preserve the inverse of the matrix of normal equations (note that it does not depend on F in any way) and compute the appropriate coefficients by multiplying the inverse by $\langle \vec{F}, \phi \rangle$ involving the given F. This may happen when a multitude of F's is given on a common M and we want to approximate them all using the same base and weight function.

This idea can be carried further. Let us denote by a_{ij} the elements of A^{-1}. Then we can write

$$c_i = \sum_{j=1}^{n} a_{ij} \langle F, \phi_j \rangle \quad i = 1, 2, \ldots, n.$$

Let us limit ourselves to the discrete M, for the moment, and we get

$$c_i = \sum_{j=1}^{n} a_{ij} \sum_{X \in M} W(X) \, F(X) \, \phi_j(X)$$

$$= \sum_{X \in M} W(X) \, F(X) \sum_{j=1}^{n} a_{ij} \, \phi_j(X) \qquad i = 1, 2, \ldots, n.$$

Denoting $\sum_{j=1}^{n} a_{ij}\,\phi_j(X)$ by $f_i(X)$ we can write

$$c_i = \sum_{X\in M} W(X)\,F(X)\,f_i(X) = \langle F, f_i\rangle \qquad i=1,2,\ldots,n.$$

Here f_i $(i=1,2,\ldots,n)$ are some new functions that are obviously derived from all the original functions ϕ_i $(i = 1,2,\ldots,n.)$.

The alternative approach would be to define

$$\tilde{f}_i(X) = W(X)\,f_i(X) \qquad i = 1,2,\ldots,n$$

and c_i then would degenerate further to

$$c_i = \sum_{X\in M} F(X)\,\tilde{f}_i(X) \qquad i = 1,2,\ldots,n\ .$$

Here f's or $\tilde{f}$'s can be stored (they do not depend on F, only on Φ and M) and the coefficients of the best fitting polynomial to any F can be computed from one of the above simple formulae.

It is left to the reader to prove that for compact M the following expressions hold:

$$c_i = \langle F, f_i\rangle \qquad i = 1,2,\ldots,n$$
$$f_i = \sum_{j=1}^{n} a_{ij}\,\phi_j(X)$$

and
$$c_i = \langle F, \tilde{f}_i\rangle, \quad W(X) = 1 \qquad i = 1,2,\ldots,n\ .$$
$$\tilde{f}_i = \sum_{j=1}^{n} a_{ij}\,W(X)\,\phi_j(X)$$

They are same as these for discrete M with the only difference that the scalar products are defined by integrals instead of summations.

3.6 Recapitulation of the Least Squares Approximation

Given 1) A base $\Phi \equiv \{ \phi_1, \phi_2, \ldots, \phi_n \}$, a set of n linearly independent functions from the functional space G_m,

2) The function to be approximated, F, defined on the set M,

$M \equiv \{ x_1, X_2, \ldots, X_m \}$ M DISCRETE

$M \equiv [a,b]$ M COMPACT .

3) A weight function W, defined and non-negative on M. Then the least squares approximation problem is to determine that vector of coefficients $(c_1, c_2, \ldots c_n)$ which minimizes the distance with weight function W, $\rho(F, P_n)$, defined as

$$\rho\ (F, P_n) \equiv \left[\sum_{x \epsilon M} W(x)\ [F(x) - P_n(x)]^2 \right]^{1/2} \qquad \text{M DISCRETE}$$

$$\rho\ (F, P_n) \equiv \left[\int_M W(x)\ [F(x) - P_n(x)]^2 dx \right]^{1/2} \qquad \text{M COMPACT}$$

where the approximating polynomial is

$$P_n = \sum^{n} c_i\ \phi_i \ .$$

The solution to this problem is the vector $(c_1, c_2, \ldots c_n)$ which satisfies the normal equations

$$\sum^{n} < \phi_i, \phi_j > c_i = < F, \phi_j > \qquad j = 1,2,\ldots,n$$

where the scalar product with weight function W, $< G,H >$, is defined as

$$<G,H> \equiv \sum_{x \epsilon M} W(x)\ G(x)\ H(x) \qquad \text{M DISCRETE}$$

$$<G,H> \equiv \int_M W(x)\ G(x)\ H(x)\ dx \qquad \text{M COMPACT} .$$

3.7 Orthogonal And Orthonormal Bases

Going back to the system of normal equations where we have left it in (3.4), we may now study it further from the point of view of the orthogonal and orthonormal bases. It is not difficult to see the computational advantages an orthogonal (orthonormal) system of basic functions has to offer. In addition to this, there is one more advantage in using an orthogonal (orthonormal) base. We can add further (orthogonal or orthonormal) basic functions to the ones for which the coefficients have already been computed and determine the new coefficients without having to change the established ones. Obviously, this is not the case with a general base Φ where an addition of new basic functions changes the whole matrix of normal equations and thus influences generally all the elements of the inverse of the original matrix.

Further, considering a general base Φ, the matrix of normal equations is non-singular but it still may be ill-conditioned. This very often happens, for instance, with generalized trigonometric systems

$$\Phi \equiv \{\cos\omega_1 X,\ \sin\omega_1 X,\ \cos\omega_2 X,\ \sin\omega_2 X,\ \ldots,\ \cos\omega_n X,\ \sin\omega_n X\}$$

where $\omega_1, \omega_2, \ldots, \omega_n$ are some real numbers. The use of an orthogonal base prevents this from happening. As we have seen in (3.4) an orthogonal base has always a diagonal matrix of normal equations. A diagonal matrix cannot be ill-conditioned (near-singular). It can only be singular but this is ruled out in our development due to the definition of an orthogonal system. Note that an <u>orthogonal</u> matrix, as usually defined (i.e. the inverse equals the transpose), is equivalent to an <u>orthonormal</u> matrix in the sense of these notes.

3.8 Orthogonalization

All the systems of orthogonal algebraic polynomials can be derived from the system of algebraic functions $\Phi \equiv \{1, X, X^2, \ldots, X^n\}$ by orthogonalization. Similarly, every general system of functions $\Phi \equiv \{\phi_1, \phi_2, \ldots, \phi_n\}$ can be transformed to an orthogonal system $\tilde{\Phi} \equiv \{P_1, P_2, \ldots, P_n\}$ on a certain M and with a certain W. One of the easiest ways to do this is to use the Schmidt's orthogonalization process. (also called the Gram - Schmidt process).

The Schmidt's process reads as follows:

i) Choose

$$P_1 \equiv \phi_1; \quad X \in M .$$

ii) Define

$$P_2 \equiv \phi_2 + \alpha_{2,1} P_1 ; \quad X \in M, \quad \alpha_{21} \in E .$$

Multiplying this equation by $W \cdot P_1$ and summing up all the equations for all the X's (for M compact we integrate the equation, multiplied by WP_1, with respect to X) we get

$$\langle P_2, P_1 \rangle = \langle \phi_2, P_1 \rangle + \alpha_{21} \langle P_1, P_1 \rangle .$$

Here $\langle P_2, P_1 \rangle$ has to be zero to make the system $\tilde{\Phi}$ orthogonal. Hence the unknown coefficient $\alpha_{2,1}$ can be obtained from

$$\alpha_{2,1} = -\frac{\langle \phi_2, P_1 \rangle}{\langle P_1, P_1 \rangle} .$$

iii) Define

$$P_3 \equiv \phi_3 + \alpha_{3,2} P_2 + \alpha_{3,1} P_1 ; \quad X \in M, \ \alpha_{3,2}, \alpha_{3,1} \in E$$

and get, by the same reasoning as above:

$$\langle P_3, P_2 \rangle = \langle \phi_3, P_2 \rangle + \alpha_{3,2} \langle P_2, P_2 \rangle + \alpha_{3,1} \langle P_1, P_2 \rangle$$

and similarly

$$\langle P_3, P_1 \rangle = \langle \phi_3, P_1 \rangle + \alpha_{3,2} \langle P_2, P_1 \rangle + \alpha_{3,1} \langle P_1, P_1 \rangle .$$

By virtue of the orthogonality $<P_1,P_2> = <P_2,P_1> = <P_3,P_2> = <P_3,P_1> = 0$ and we obtain

$$<\phi_3,P_2> + \alpha_{3,2}<P_2,P_2> = 0$$

$$<\phi_3,P_1> + \alpha_{3,1}<P_1,P_1> = 0 .$$

Hence

$$\alpha_{3,1} = -\frac{<\phi_3,P_1>}{<P_1,P_1>}$$

$$\alpha_{3,2} = -\frac{<\phi_3,P_2>}{<P_2,P_2>} .$$

We then progress the same way ending up with the equations for coefficients $\alpha_{n,i}$ $i=1,2,\ldots,n-1$. Evidently, the general formula for any $\alpha_{j,i}$ is

$$\alpha_{j,i} = -\frac{<\phi_j,P_i>}{<P_i,P_i>} = -\frac{<\phi_j,P_i>}{||P_i||^2} .$$

This method can be obviously used for both compact and discrete M.

3.9 Extension To Three Dimensions

3.9.1 Extension of definitions

The theory of least-squares approximation as explained so far can be very easily extended into three-dimensional space (as opposed to the two-dimensional we have been dealing with up to now). In three-dimensional space we shall be dealing with surfaces instead of curves. Otherwise we have to retain the two parallel streams dealing with either discrete or compact definition sets.

Denoting by β the two-dimensional area (either compact or discrete) we can define the norm:

$$||G|| = \sqrt{\sum_{X\in\beta} W(X)\, G(X)^2} \in E^+$$

for discrete β and

$$||G|| = \sqrt{\int_\beta W(X)\, G(X)^2 dX} \in E^+$$

for compact β . Here X is a pair of arguments, (x, y) say. Similarly, the scalar product may be defined as

$$\langle G,H\rangle = \begin{cases} \sum_{X\in\beta} W(X)\, G(X)\, H(X) & \beta \text{ discrete} \\ \int_\beta W(X)\, G(X)\, H(X)\, dX & \beta \text{ compact} . \end{cases}$$

G, H are, of course, functions of X defined on β , i.e. G, H $\in G_\beta$.

The condition of minimum least-squares distance leads again to the same system of normal equations. The linear independence of the system of prescribed functions Φ has to be hence also required.

We may notice that if the functions ϕ_i of the base Φ can be expressed as products of two functions of single variable, $\phi_i(X) = \psi_i(x) \cdot \chi_i(y)$ say, and if even the weight function W can be split into v(x) . u(y) we get:

$$\langle \phi_i,\phi_j\rangle = \langle \psi_i,\psi_j\rangle_v \cdot \langle \chi_i,\chi_j\rangle_u .$$

The proof of this is left to the reader.

3.9.2 Orthogonality in two-dimensional space

One may now ask the question whether relations like orthogonality also exist for functions defined on two-dimensional areas. There is no reason why they should not exist and as a matter of fact the orthogonality can be defined in exactly the same way as for functions defined on an interval:

$$\langle \phi_i,\phi_j\rangle = ||\phi_i||^2\, \delta_{ij} .$$

There is even a very simple relation between orthogonal systems in one and two dimensional spaces. Let $\Phi_1 \equiv \{\psi_1, \psi_2, \ldots, \psi_n\}$ be a system orthogonal on M_1 with u and $\Phi_2 \equiv \{ X_1, X_2, \ldots, X_n\}$ be a system orthogonal on M_2 with v. Then the product Φ of these two systems is orthogonal on the area $M_1 \times M_2 \equiv \beta$ with $W = u \cdot v$.

To show this, let us denote

$$\phi_i(X) = \psi_\ell (x) \cdot X_k (y) .$$

Then

$$\langle \phi_i, \phi_j \rangle = \langle \psi_\ell, \psi_s \rangle_u \langle X_k, X_r \rangle_v = \delta_{\ell s} \, ||\psi_\ell||^2 \, \delta_{kr} \, ||X_k||^2$$
$$= \delta_{ij} \, ||\phi_i||^2 .$$

The most widely used two-dimensional orthogonal system in geodesy are the spherical functions (spherical harmonics) orthogonal on a sphere (or in any rectangle 2π by 2π) with weight $W(X) = 1$. They originate as a product of associated Legendre's functions and the trigonometric system.

Another application is in approximating the topographic surface in an area. For this purpose any system can be used.

3.10 Problems

Solved Problems

1. Given the weight function $W(x) = 1$ and the definition set $M \equiv \{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$ find the numerical functions f_1 and f_2 which allow the linear approximation of any function F defined on M.

$P_2 = c_1 + c_2 x$ where $c_1 = \langle F, f_1 \rangle$, $c_2 = \langle F, f_2 \rangle$.

What property must the definition set M posess in order that this technique be applied?

Solution

The linear approximating polynomial is

$$P_2(x) = \sum_{i=1}^{2} c_i \phi_i(x) = c_1 \phi_1(x) + c_2 \phi_2(x) = c_1 + c_2 x$$

hence the base functions are $\phi_1(x) = 1$ and $\phi_2(x) = x$. The coefficients are the solutions to the normal equations

$$\sum_{j=1}^{2} \langle \phi_i, \phi_j \rangle c_j = \langle F, \phi_i \rangle \qquad i = 1,2.$$

Noting that in this case

$$\langle \phi_1, \phi_1 \rangle = \sum_{x \in M} W(x)\, \phi_1(x) \phi_1(x) = \sum_{x \in M} 1 = 11$$

$$\langle \phi_2, \phi_2 \rangle = \sum_{x \in M} W(x) \phi_2(x)\, \phi_2(x) = \sum_{x \in M} x^2 = 110$$

$$\langle \phi_1, \phi_2 \rangle = \langle \phi_2, \phi_1 \rangle = \sum_{x \in M} W(x) \phi_1(x) \phi_2(x) = \sum_{x \in M} x = 0$$

the normal equations become

$$\begin{bmatrix} 11 & 0 \\ 0 & 110 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \langle F, \phi_1 \rangle \\ \langle F, \phi_2 \rangle \end{bmatrix}$$

so that

$$c_1 = \langle F, \phi \rangle / 11$$

$$c_2 = \langle F, \phi_2 \rangle / 110$$

that is, the functional values of functions f_1 and f_2 are

$$f_1 = (1/11)\ (1,1,1,1,1,1,1,1,1,1,1)\ ,\quad f_1(x) = 1/11.$$

$$f_2 = (1/110)\ (-5,-4,-3,-2,-1,0,1,2,3,4,5),\quad f_2(x) = x/110.$$

In general, if M is <u>symmetric,</u> that is

$$\sum_{x\in M} x = 0$$

then

$$c_1 = \langle F,\phi_1\rangle/n \qquad \text{where } \phi_1(x) = 1$$

$$c_2 = \langle F,\phi_2\rangle/\Sigma x^2 \qquad \text{where } \phi_2(x) = x.$$

2. Derive a system of orthonormal algebraic polynomials from the system of algebraic functions $\Phi \equiv \{1,x,x^2,\ldots,x^n\}$, defined on $M \equiv \{-5,-4,-3,-2,-1,0,1,2,3,4,5\}$ and with weight function $W(x) = 1$. Choose your own value for n.

<u>Solution</u>

We select n = 2 and obtain an orthogonal system P by applying the Schmidt orthogonalization process to Φ, that is

$$P_1 = \phi_1 = 1$$

$$P_2 = \phi_2 + \alpha_{21}P_1$$

$$P_3 = \phi_3 + \alpha_{32}P_2 + \alpha_{31}P_1$$

where

$$\alpha_{ji} = -\frac{\langle \phi_j,\ P_i\rangle}{\langle P_i,\ P_i\rangle}\ .$$

Specifically

$$\alpha_{21} = -\frac{\langle \phi_2,P_1\rangle}{\langle P_1,P_1\rangle} = -\frac{\sum_{x\in M} W(x)\phi_2(x)P_1(x)}{\sum_{x\in M} W(x)P_1(x)P_1(x)} = 0$$

so that $P_2 = \phi_2$, and again

$$\alpha_{32} = - \frac{\langle \phi_3, P_2 \rangle}{\langle P_2, P_2 \rangle} = - \frac{\sum_{x \in M} W(x) \phi_3(x) P_2(x)}{\sum_{x \in M} W(x) P_2(x) P_2(x)} = 0$$

$$\alpha_{31} = - \frac{\langle \phi_3, P_1 \rangle}{\langle P_1, P_1 \rangle} = - \frac{\sum_{x \in M} W(x) \phi_3(x) P_1(x)}{\sum_{x \in M} W(x) P_1(x) P_1(x)} = - \frac{110}{11} = - 10$$

so that $P_3 = x^2 - 10$ and our orthogonal system is $P \equiv \{1, x, x^2 - 10\}$.

We obtain an orthonormal system $\tilde{P}$ by dividing each element of P by its norm, that is by

$$||P_i|| = \sqrt{\langle P_i, P_i \rangle}$$

Thus

$$\tilde{P}_1 = \frac{P_1}{||P_1||} = \frac{1}{\sqrt{11}}$$

$$\tilde{P}_2 = \frac{P_2}{||P_2||} = \frac{x}{\sqrt{110}}$$

$$\tilde{P}_3 = \frac{P_3}{||P_3||} = \frac{x^2 - 10}{\sqrt{858}}$$

and we have the orthonormal system $\tilde{P} \equiv \{1/\sqrt{11}, x/\sqrt{110}, (x^2-10)/\sqrt{858}\}$, that is

$$\langle \tilde{P}_i, \tilde{P}_j \rangle = \delta_{ij} .$$

Note the close relationship between these results and those in the previous problem.

Unsolved problems

3. Given F(x) defined for $x \in [-1, 1]$, approximate F in the least squares sense by

a) P_n for ϕ $\{1,x,x^2, \ldots, x^n\}$ with weight function $W(x) = 1$

b) P_n for ϕ $\{1,x,x^2, \ldots,x^n\}$ with weight function $W(x) = (1 - x^2)^{-1/2}$

c) P_n for ϕ $\{L_0,L_1, \ldots L_n\}$ (Legendre's polynomials)

d) P_n for ϕ $\{T_0,T_1, \ldots T_n\}$ (Tchebyshev's polynomials)

Compare a) with c) and b) with d). Choose your own F and n. For the individual orthogonal polynomials see the Appendix.

4. Write debug and document a general least squares approximation computer program with the following features. It should approximate and function F defined on a discrete set of arguments $M \equiv \{x_1,x_2, \ldots x_n\}$ where the x are not necessarily equidistant. The user should be free to choose his own base $\phi \equiv \{\phi_1,\phi_2, \ldots \phi_m\}$ and weight function W. Therefore the input must include the vectors x and F(x), the base, and the weight function. The output should comprise the best fitting coefficients $c_1,c_2,\ldots,c_m$ and the variance $p(F,P_m)/\sqrt{n-m}$. (For explanation of the variance see the next section). A check for linear independence of ϕ on M should be incorporated.

4. LEAST SQUARES ADJUSTMENT

In this chapter we discuss three least squares adjustments, called the parametric adjustment, the condition adjustment, and the combined adjustment. The purpose of this discussion is twofold. First we describe the relationship between the parametric adjustment and least squares approximation. Second we present geometrical interpretations of each of the three adjustments.

Linear mathematical models and uncorrelated observations are assumed throughout this chapter. A discussion of least squares adjustments using non-linear models and correlated observations can be found in Wells & Krakiwsky, 1971.

4.1 Equivalence of Parametric Adjustment and Least Squares Approximation

The parametric least squares adjustment differs only in intent and notation from the least squares approximation of a function F defined on a discrete domain M.

The intent of the least squares approximation is to find an approximating function P_n for a given function F. The intent of the least squares adjustment is to find the "least squares statistical estimates" (in our notation $\{c_i\}$) of unknown parameters (coefficients) which are related to the observed values (in our notation $\{F(x_j)\}$) by a linear (or linearized) mathematical model (or system of observation equations).

$$F(x_j) = \sum_{i=1}^{n} c_i \, \phi_i(x_j).$$

The connection between the two is simply demonstrated by restating the least squares approximation in matrix notation. The column vector

$$[F(x_1), F(x_2), \ldots F(x_m)]^T = {}_mF_1$$

is called in adjustment the vector of observed values (or misclosure vector in the case of linearized models). The matrix whose diagonal elements are the statistical weights attached to the observed values in F

$$\begin{bmatrix} W(x_1) & & & 0 \\ & W(x_2) & & \\ & & \ddots & \\ 0 & & & W(x_m) \end{bmatrix} = {}_mP_m$$

is called in adjustment the weight matrix, or the inverse of the covariance matrix of the observables. Note that in this formulation of the adjustment problem we assume the observations to be statistically independent (P has no off-diagonal elements). This is not always the case. The matrix whose column vectors are vectors of functional values for each of the base functions ϕ_i (Vandermonde's matrix)

$$\begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \ldots & \phi_n(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \ldots & \phi_n(x_2) \\ \vdots & \vdots & & \vdots \\ \phi_1(x_m) & \phi_2(x_m) & \ldots & \phi_n(x_m) \end{bmatrix} = {}_mA_n$$

is called in adjustments the design matrix. The vector of coefficients

$$[c_1, c_2, \ldots c_n]^T = {}_nC_1$$

is called in adjustment the vector of unknown parameters.

Note that the vector resulting from the matrix product

A C

is the vector whose elements are the functional values $P_n(x_j)$ of the approximating function P_n.

The column vector

$$V = F - A\,C$$

is called in adjustment the vector of residuals (or discrepancies).

We want to find the coefficient (or unknown) vector C which minimizes the quadratic form

$$V^T P V = \rho^2 (F, AC) = \rho^2 (F, P_n).$$

The solution is that value of C which satisfies the normal equations

$$(A^T P A)\, C = A^T P F.$$

In accordance with adjustment convention, the least squares estimate for the variance factor σ_o^2 is

$$\hat{\sigma}_o^2 = V^T P V/(m-n) = \sigma^2 (F, P_n)/(m-n) .$$

The estimate of the covariance matrix of the vector C is then

$$\hat{\Sigma}_c = \hat{\sigma}_o^2 (A^T PA)^{-1} .$$

If the base Φ is orthogonal with weight function W then the matrix

$$A^T P A = \begin{bmatrix} \langle \phi_1, \phi_1 \rangle & \langle \phi_1, \phi_2 \rangle & \cdots & \langle \phi_1, \phi_n \rangle \\ \langle \phi_2, \phi_1 \rangle & \langle \phi_2, \phi_2 \rangle & \cdots & \langle \phi_2, \phi_n \rangle \\ \langle \phi_n, \phi_1 \rangle & \langle \phi_n, \phi_2 \rangle & \cdots & \langle \phi_n, \phi_n \rangle \end{bmatrix}$$

is diagonal, and therefore so is its inverse. In this case all covariances between the coefficients c_i are zero, that is the coefficients are statistically independent. If the base is orthonormal with weight function W, then

$$A^T P A = (A^T P A)^{-1} = I$$

and all the individual variances of the coefficients c_i are equal to the variance factor $\hat{\sigma}_o^2$. Note that these results depend on the weight matrix P being diagonal (i.e. having all off-diagonal elements equal to zero).

4.2 Geometrical Interpretations of the Parametric Adjustment

A vector has both algebraic and geometric meaning. In this section we investigate the geometric meanings of the vectors involved in the parametric adjustment. We begin with the vector of observed values ${}_mF_1$, the weight matrix ${}_mP_m$, and the design matrix ${}_mA_n$. The vector of coefficients ${}_nc_1$, is obtained by solving the normal equations

$$(A^T P A) C = A^T P F .$$

The residual vector ${}_mV_1$, is obtained from

$$V = F - A C .$$

Note that the number of observations, m, must be greater than the number of coefficients, n. The augmented matrix

$$[A \mid F]$$

contains all the known data (except for the weights in P). The elements of this data matrix can be considered either to be organized into m-dimensional column vectors

$$[A \mid F] = [A_1 \mid A_2 \mid \ldots \mid A_n \mid F]$$

or to be organized into (n + 1)-dimensional row vectors, of which there are m.

Returning momentarily to the notation of the least squares approximation, we see that the column vectors of $[A \mid F]$ are the vectors of functional values

$$\vec{\phi}_i \equiv (\phi_i(x_1), \phi_i(x_2), \ldots \phi_i(x_m))$$

$$\vec{F} \equiv (F(x_1), F(x_2), \ldots F(x_m)).$$

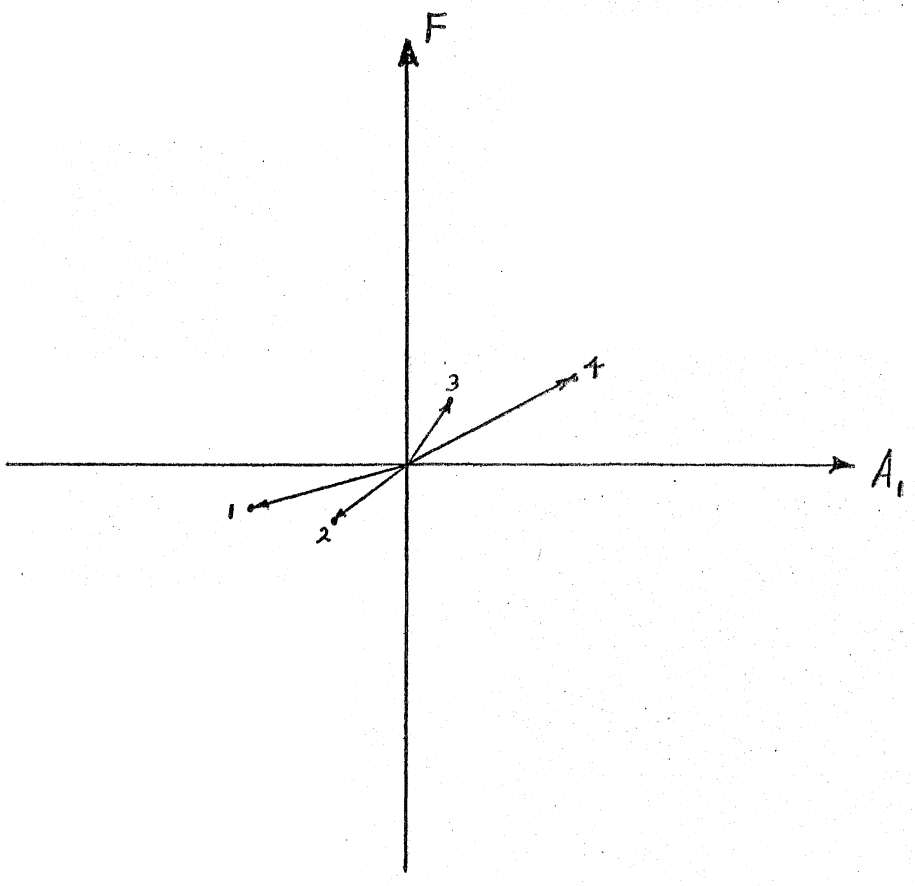
They can be considered as tables of values which serve to define the functions ϕ_i and F.

We will consider two geometries in which A, F and V can be interpreted as geometrical quantities. In the first geometry, the m row vectors of the data matrix are considered to be the position vectors of m data points, plotted in the (n + 1)-dimensional space for which the column vectors are a basis (that is, the column vectors span this space).

In the second geometry, the n + 1 column vectors of the data matrix are considered to be n+1 position vectors plotted in the m-dimensional space spanned by the row vectors. These are multi dimensional geometries, and can be visually illustrated only if the dimensionality is 3 or less.

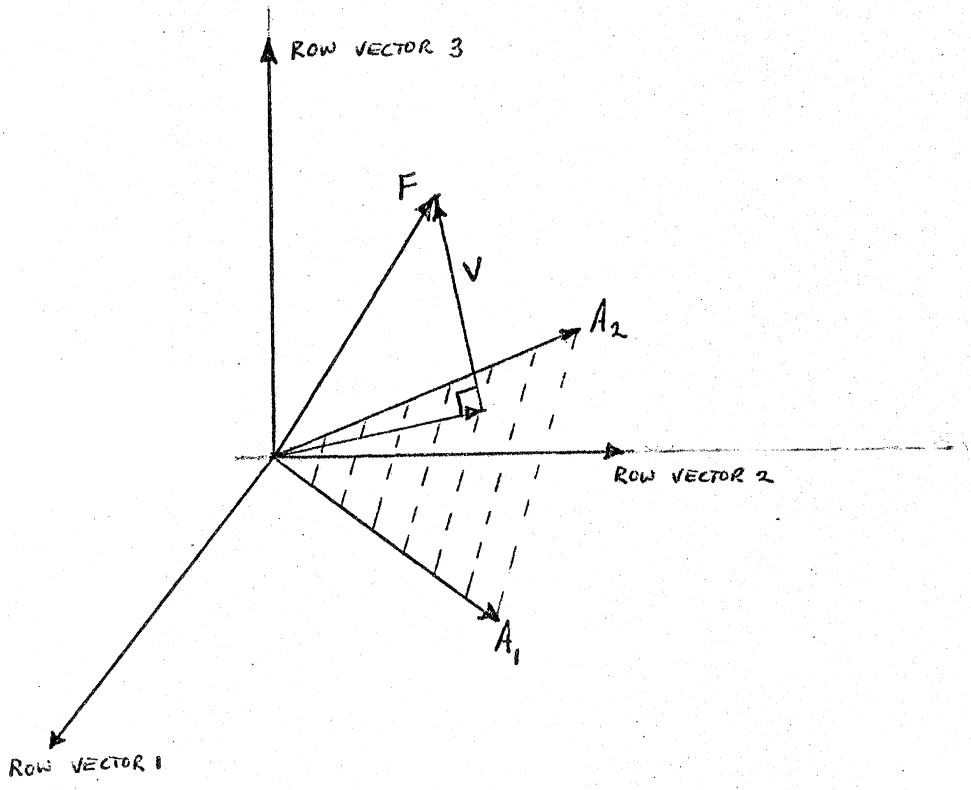
Illustrating the first geometry (often called the "scatter plot") for the case n=1 and m=4 (note n+1 space = 2-space), the data matrix is

$$\left[\begin{array}{c:c} \phi_1(x_1) & F(x_1) \\ \phi_1(x_2) & F(x_2) \\ \phi_1(x_3) & F(x_3) \\ \phi_1(x_4) & F(x_4) \end{array}\right]$$



Illustrating the second geometry for the case n = 2 and m = 3 (note m-space = 3-space): the data matrix is

$$\left[\begin{array}{c:c:c} \phi_1(x_1) & \phi_2(x_1) & F(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & F(x_2) \\ \phi_1(x_3) & \phi_2(x_3) & F(x_3) \end{array}\right].$$



Note that Geometrical concepts such as parallelism, orthogonality and length remain valid in spaces of any dimension even though we cannot visualize them as above. Some of the terms we use are m-dimensional hyperspace (a Euclidean space of dimension m), n-dimensional hyperplane (a Euclidean subspace of dimension n, of a Euclidean space of higher dimensionality), and m-dimensional hypersphere (a generalization of a sphere into m dimensions). We will consider these two geometries in detail, but will first consider the significance of the weight matrix P.

4.2.1 Geometrical Interpretation of the Weight Matrix P

We have seen that the statistical meaning of the P matrix is that it is the inverse of the covariance matrix of the observables. If P is diagonal, its elements are the reciprocals of the variances of the observables. Then the diagonal matrix whose elements are the square roots of the elements of P (which we will denote by $P^{1/2}$) has the reciprocals of the standard deviations of the observables as elements.

A multivariate statistical distribution is one in which the multivariate mean is zero and the multivariate variance is unity. It is convenient to transform the space in which the adjustment problem is stated, into a space in which the multivariate distribution of the observables has been transformed into the standard multivariate distribution, that is a space which statistically is hyperspherically symmetric. When introducing the two geometries above, we specified that the row (or column) vectors of the data matrix be considered as position (or radius) vectors in our first (or second) geometry. This is equivalent to transforming the multivariate mean to zero. To transform the multivariate variance to unity we must transform the weight matrix P into the unit matrix. We shall call this hyperspherically symmetric space "tilde space" (after the diacritic mark ~).

The matrix operation which accomplishes the statistical transformation from the problem-space to tilde-space is the premultiplication of the data matrix by $P^{1/2}$. This scales the row vectors of the data matrix by the reciprocals of the standard deviations of the corresponding elements of F. Thus

$$[\tilde{A} \vdots \tilde{F}] = P^{1/2}[A \vdots F] = [P^{1/2} A \vdots P^{1/2} F].$$

This transformation also scales the residual vector

$$\tilde{V} = \tilde{F} - A C = P^{1/2} V.$$

The adjustment problem is to find the coefficient vector C which minimizes

$$\tilde{V}^T \tilde{V} = V^T P V,$$

and the solution is that vector C which satisfies the normal equations

$$(\tilde{A}^T \tilde{A})\ C = \tilde{A}^T \tilde{F}.$$

In tilde-space the weight matrix has been transformed into the unit matrix

$$\tilde{P} = \mathbf{I}.$$

Note that if P is not a diagonal matrix no immediate geometrical interpretation or transformation to a statistically symmetric space can be made.

4.2.2. Geometrical Interpretation of the Approximant Vector and the Residual Vector.

In this section we will work in tilde-space. Let us take an arbitrary coefficient vector ${}_nC_1$. To each such vector there corresponds an approximant vector ${}_mG_1$, given by

$$G = \tilde{A}\ C = \sum_{i=1}^{n} c_i\ \tilde{A}_i$$

where c_i is the ith element of the coefficient vector C, and $\tilde{A}_i$ is the ith column vector of the design matrix ${}_m\tilde{A}_n$. To each approximant vector there corresponds a residual vector ${}_m\tilde{V}_1$, given by

$$\tilde{V} = \tilde{F} - G.$$

In the notation of the least squares approximation, G is the vector whose elements are the functional values of an approximating function P_n, that is

$$G \equiv (P_n(x_1),\ P_n(x_2), \ldots P_n(x_m)\ ).$$

The first geometry. We have already plotted the row vectors of $[\tilde{A}_i|\tilde{F}]$ as position vectors of points in the (n + 1)-dimensional space spanned by the column vectors of $[\tilde{A}_i|\tilde{F}]$. Now let us plot in this same space the row vectors of $[\tilde{A}_i|G]$. These new data points all lie on a surface in (n + 1)-space. Because G is a linear combination of the column vectors $\tilde{A}_i$, this surface is a subspace (hyperplane) of that spanned by the column vectors of $[\tilde{A}_i|\tilde{F}]$. There is one such hyperplane for each approximant G, or equivalently for each coefficient vector C.

The residual vector corresponding to an approximant G is the vector whose elements are the distances from the original set of data points (the row vectors of $[\tilde{A}\,|\,\tilde{F}]$) and the new data points (the row vectors of $[\tilde{A}\,|\,G]$). Each of these distances is measured perpendicularly to another hyperplane, that spanned by the column vectors of $\tilde{A}$.

Thus in the first geometry the parametric adjustment problem is to find that hyperplane (or equivalently that coefficient vector C) which "best" approximates the cluster of m original data points, in the sense that the sum of the squares of the distances from that hyperplane to each data point, measured orthogonally to a second hyperplane (that spanned by the column vectors of the given matrix $\tilde{A}$), be a minimum.

The second geometry. The vector ${}_m\tilde{F}_1$ and the n column vectors of ${}_m\tilde{A}_n$ have been plotted in the m-dimensional space spanned by the row vectors of $[\tilde{A}\,|\,\tilde{F}]$. The n column vectors of $\tilde{A}$ span an n-dimensional subspace (hyperplane) of this space (which we will call the $\tilde{A}$ hyperplane). Since an approximant ${}_mG_1$ is a linear combination of the column vectors $\tilde{A}_i$, then G lies in the $\tilde{A}$ hyperplane. The residual vector ${}_m\tilde{V}_1$ then is the distance from a point in this hyperplane to the point whose position vector is $\tilde{F}$. Obviously the residual vector is of minimum length (for a given $\tilde{F}$ and given $\tilde{A}$ hyperplane) if and only if it is the orthogonal distance from the $\tilde{A}$ hyperplane to $\tilde{F}$ (that is if it is perpendicular to the $\tilde{A}$ hyperplane).

Thus in the second geometry the parametric adjustment problem is to decompose the given vector $\tilde{F}$ into two other vectors; an approximant vector G and a residual vector $\tilde{V}$, that is

$$\tilde{F} = G + \tilde{V}$$

such that a) G lies in the $\tilde{A}$ hyperplane, that is

$$G = \tilde{A}\,C$$

b) $\tilde{V}$ is orthogonal to the $\tilde{A}$ hyperplane, and thus to every vector in it, and in particular to the column vectors of A, that is

$$\tilde{A}^T\,\tilde{V} = 0.$$

It is simple to show that these three equations are equivalent to the normal equations. Premultiplying the first by $\tilde{A}^T$ and substituting from the other two we have

$$\tilde{A}^T \tilde{F} = \tilde{A}^T G + \tilde{A}^T \tilde{V} = \tilde{A}^T \tilde{A} C + 0$$

which are the normal equations in our "tilde-space". This geometry thus illustrates the meaning of the work "normal" used in "normal equations". It refers to the orthogonality (or normality) between the residual vector and the approximant vector (or the hyperplane).

In the next sections we will extend the application of this second geometry to interpret the condition and combined adjustments.

4.3 The Condition Adjustment

In our second geometry we have seen that in tilde-space the column vectors $(\tilde{A}_1, \tilde{A}_2, \ldots \tilde{A}_n)$ of the design matrix $\tilde{A}$ form a basis for (span) an n-dimensional hyperplane (the $\tilde{A}$ hyperplane) in the m-dimensional space spanned by the row vectors of the data matrix $[\tilde{A} \mid \tilde{F}]$.

Therefore there must exist (m-n) linearly independent vectors in m-space, $(\tilde{B}_1, \tilde{B}_2 \ldots \tilde{B}_{m-n})$, each of which is orthogonal to each of the column vectors $\tilde{A}_i$. That is

$$\tilde{B}_j^T \tilde{A}_i = 0 \qquad \text{for } i=1,2,\ldots n \text{ and } j = 1,2,\ldots,m-n.$$

If we form the matrix

$$\tilde{B}^T = [\tilde{B}_1, \tilde{B}_2, \ldots \tilde{B}_{m-n}]$$

Then

$$B A = 0.$$

If this result is to be valid in both tilde-space and our problem-space then

$$\tilde{B} \tilde{A} = B A = 0 .$$

But $\tilde{A} = P^{1/2} A$. Hence $\tilde{B} = B P^{-1/2}$, that is the transformation of B to tilde-space is accomplished by scaling the column vectors of B by the standard deviations of the corresponding observables. Continuing in tilde-space, we see that the row vectors of $\tilde{B}$ span an (m-n)-dimensional hyperplane (the $\tilde{B}$ hyperplane), and that the $\tilde{A}$ and $\tilde{B}$ hyperplanes are orthogonal.

Premultiplying the equation defining the residual vector $\tilde{V}$, namely

$$\tilde{V} = \tilde{F} - \tilde{A}\,C$$

by the matrix $\tilde{B}$, we obtain

$$\tilde{B}\,\tilde{V} = \tilde{B}\,\tilde{F} - \tilde{B}\,\tilde{A}\,C = \tilde{B}\,\tilde{F}.$$

Geometrically this means that the projection of $\tilde{V}$ onto each of the row vectors of $\tilde{B}$ is equal to the projection of $\tilde{F}$ onto the same row vectors. However, we have seen that for the length of $\tilde{V}$ to be minimized, $\tilde{V}$ must be orthogonal to $\tilde{A}$ hyperplane. Therefore it must lie in the $\tilde{B}$ hyperplane, that is $\tilde{V}$ must be a linear combination of the row vectors of $\tilde{B}$, or

$$\tilde{V} = \tilde{B}^T K$$

where the vector of coefficients K is called the vector of correlates. Substituting this in the previous equation we have

$$\tilde{B}\,\tilde{B}^T K = \tilde{B}\,\tilde{F},$$

the normal equations for the condition adjustment.

Thus the condition adjustment problem is to find that vector $\tilde{V}$ (or equivalently that vector K) such that $\tilde{V}$ lies in the $\tilde{B}$ hyperplane and has the same projection as the given vector $\tilde{F}$ on each of the row rectors of $\tilde{B}$ which span the $\tilde{B}$ hyperplane.

Note that the condition and parametric adjustments are duals. Any adjustment problem amenable to solution using one, can also be solved using the other.

4.4 The Combined Adjustment

4.4.1 The Combined Adjustment resulting from Implicit Mathematical Models

The parametric and condition adjustments are used for explicit mathematical models, and the combined adjustment for implicit mathematical models. Let us define what is meant by explicit and implicit. If the vector of observed values F is expressable as a direct function of the vector of unknown coefficients C, that is

$$F = F(C),$$

then the relationship (mathematical model) is explicit. However, if F can not be explicitly expressed as a function of C, that is if the relationship must be given as

$$G(F,C) = 0$$

then the relationship (mathematical model) is implicit.

We have seen in the case of an explicit linear relationship, the mathematical model is

$$F = A\ C + V,$$

where A is the design matrix, and V is the residual vector. All adjustment problems involving such an explicit mathematical model can be handled by both the parametric and condition adjustments. (Note that the least squares approximation, as described in these notes, always involves an explicit mathematical model).

In the case of an implicit linear relationship, F itself will not be explicitly related to C, but linear combinations of the elements of F will be. That is, there exists a (design) matrix B such that the misclosure vector F*

$$F^* = B\ F$$

is explicitly related to C, that is

$$F^* = A\ C + V^* .$$

Let m be the number of original observed values, n be the number of coefficients, and r be the number of linear combinations of the observed values which are related explicitly to C (i.e. r is the number of equations). Note that there must be at least one observed value for each equation, that is

$$m \geq r$$

and there must be more equations than unknown coeffients, that is

$$r > n .$$

This would be a parametric adjustment problem if we wished to minimize the length of the residual vector V*, whose elements are the residuals corresponding to the elements of the misclosure vector F*. However we wish instead to minimize the length of the residual vector V, whose elements are the residuals of the original vector of observed values F, that is

$$V^* = B\ V.$$

Thus we are unable to use the parametric or condition adjustments without some modification.

There are two possibilities. The first possibility is to apply the covariance law to $V^* = B\ V$ to obtain the covariance matrix of V* as

$$\Sigma_{V^*} = B\ \Sigma_V\ B^T$$

where Σ_V is the covariance matrix of V. Since the weight matrix we have so far considered is

$$P = \Sigma_V^{-1}$$

then

$$P^* = \bar{\Sigma}^1 = (B\ P^{-1}\ B^T)^{-1}$$

and we proceed to apply the parametric adjustment to

$$F^* = AC + V^*$$

by choosing that vector C which minimizes the quadratic form

$$V^{*T}\ P^*\ V^* .$$

The solution is that vector C which satisfies the normal equations

$$A^T\ P^*\ A\ C = A^T\ P^*\ F^* .$$

The second possibility is to use the combined adjustment, for which the mathematical model is

$$F^* = A\ C + B\ V$$

where

$$F^* = B\ F$$

and the known quantities are the two design matrices ${}_rA_n$ and ${}_rB_m$ and the misclosure vector ${}_rF_1^*$, so that the data matrix in this case is

$$[A\,|\,B\,|\,F^*] .$$

The unknown quantities are the vector of coefficients ${}_nC_1$ and the residual vector ${}_mV_1$. The combined adjustment problem is to find that vector C which minimizes the length of the vector V, or equivalently the quadratic form

$$V^T P V$$

where ${}_mP_m$ is the weight matrix of the original vector of observed values F.

The solution is that vector C which satisfies the normal equations

$$A^T (B P^{-1} B^T)^{-1} A C = A^T (B P^{-1} B^T)^{-1} F*$$

(which are the same normal equations as in the first possibility above).

If the weight matrix P of the observed values F is diagonal, then the statistical transformation to "tilde-space" is accomplished by the matrix operations

$$\tilde{B} = B P^{-1/2}$$

$$\tilde{V} = P^{1/2} V.$$

Note that

$$\tilde{A} = A$$

$$\tilde{F}* = F*$$

$$\tilde{V}* = V*$$

and in tilde-space

$$\tilde{P} = I.$$

4.4.2 Geometrical Comparison between Parametric and Combined Adjustments

The difference between the parametric and combined adjustments can be explained geometrically rather simply. We will use the second geometry already described, and we will work in tilde-space.

In the parametric adjustment we are given the data matrix

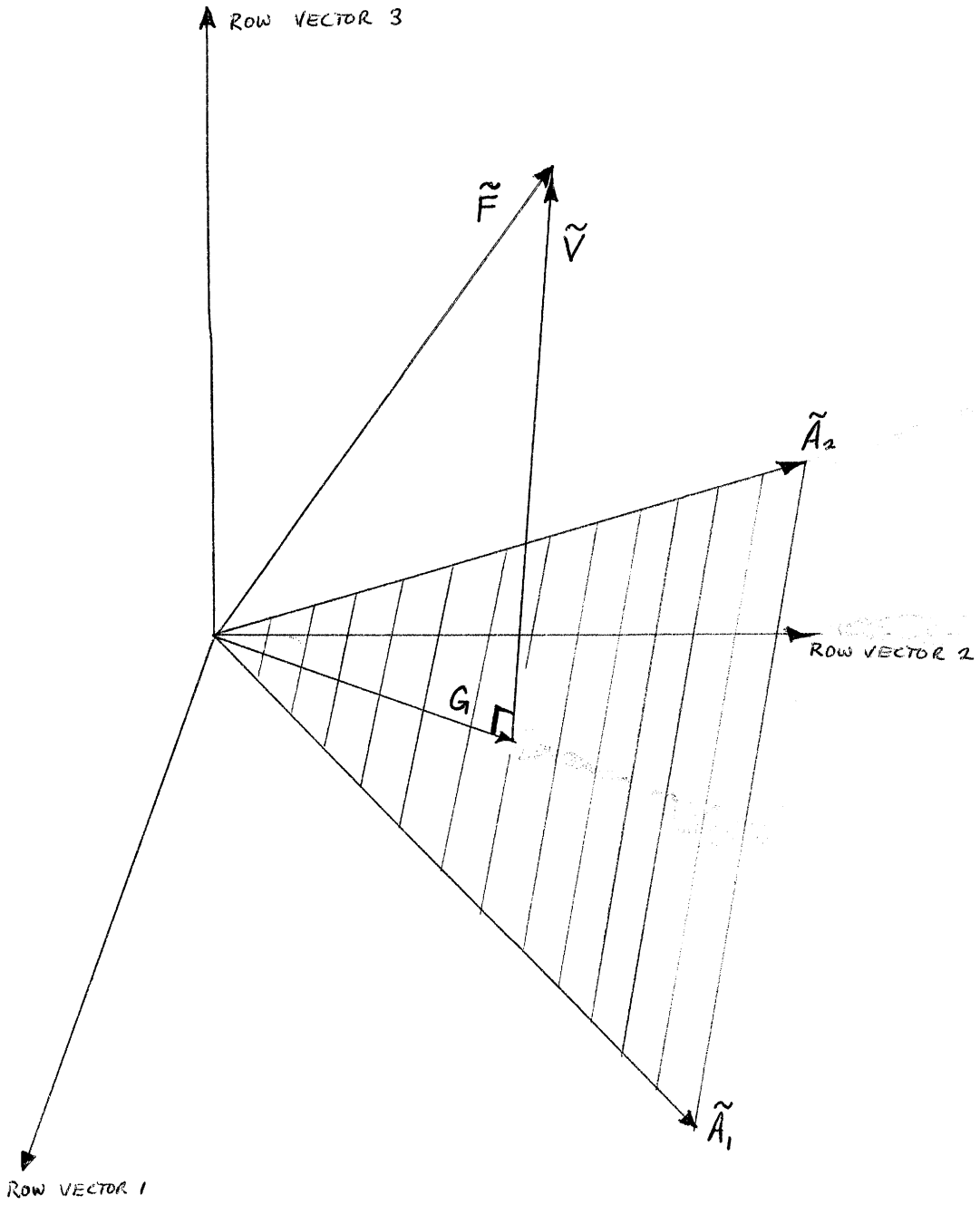
$$[\tilde{A} \mid \tilde{F}]$$

and the adjustment problem is to decompose the given vector $\tilde{F}$ into two vectors

$$G = \tilde{A} C \text{ and } \tilde{V}$$

$$\tilde{F} = \tilde{A} C + \tilde{V}$$

such that G lies in the given hyperplane (the $\tilde{A}$ hyperplane) and V is orthogonal to the $\tilde{A}$ hyperplane. For example


ROW VECTOR 3
$\tilde{F}$
$\tilde{V}$
$\tilde{A}_2$
ROW VECTOR 2
G
$\tilde{A}_1$
ROW VECTOR 1

In the combined adjustment we are given the data matrix

$$[A \mid \tilde{B} \mid F^*]$$

and applying our second geometry, we wish to plot the (n + m + 1) column vectors of the data matrix as the position vectors of points in the r-dimensional space spanned by the row vectors of the data matrix. Again, since $n < r$, the column vectors of ${}_rA_n$ define an n-dimensional hyperplane (the A hyperplane). However there are $m \geq r$ column vectors of ${}_r\tilde{B}_m$ so that only r of them can be linearly independent, and these column vectors will in general span all of r-space itself. The combined adjustment problem is to decompose the given vector F* into two vectors $G = A\ C$ and V*, that is

$$F^* = A\ C + V^*$$

such that G lies in the given A hyperplane as before, but V* is <u>not</u> orthogonal to the A hyperplane. Instead there is a <u>vector of correlates</u> K which is orthogonal to the A hyperplane, that is

$$A^T K = 0$$

and the vector V* is a linear combination of the column vectors of $\tilde{B}$, where the coefficients of this linear combination are the elements of a vector $\tilde{V}$, that is

$$V^* = \tilde{B}\ \tilde{V} = \Sigma\ v_i\ \tilde{B}_i$$

and the elements of $\tilde{V}$ are the projections of K onto each of the column vectors of $\tilde{B}$, that is

$$\tilde{V} = \tilde{B}^T K.$$

The normal equations are obtained by premultiplying the first of the above equations by $A^T (\tilde{B}\ \tilde{B}^T)^{-1}$ and substituting from the other three equations

$$A^T(\tilde{B}\ \tilde{B}^T)^{-1}F^* = A^T(\tilde{B}\ \tilde{B}^T)^{-1}\ A\ C + A^T(\tilde{B}\ \tilde{B}^T)^{-1}\ V^*$$

where

$$V^* = \tilde{B}\ \tilde{V} = \tilde{B}\ \tilde{B}^T K$$

so that

$$A^T(\tilde{B}\ \tilde{B}^T)^{-1}\ V^* = A^T(\tilde{B}\ \tilde{B}^T)^{-1}\ \tilde{B}\ \tilde{B}^T K = A^T K = 0\ .$$

5. CONNECTION BETWEEN THE LEAST -SQUARES APPROXIMATION AND INTERPOLATION

Suppose, in a least-squares approximation problem the number of basic functions equal the number of points of the discrete M and $W(X) \equiv 1$. Then the best-fitting least-squares polynomial is identical with the interpolating polynomial.

For the interpolating polynomial P_n, we get the system of n equations for the coefficients:

$$P_n(X_j) = \sum_{i=1}^{n} c_i \phi_i(X_j) = F(X_j) \qquad j=1,2,\ldots,n.$$

To show that the same coefficients can be obtained from normal equations, let us multiply each equation of the above system by $\phi_\ell(X_j)$, $\ell = 1,2,..,n$. We get

$$\sum_{i=1}^{n} c_i \phi_i(X_j)\, \phi_\ell(X_j) = F(X_j)\, \phi_\ell(X_j) \qquad j,\ell = 1,2,\ldots,n.$$

When we sum these equations up with respect to j we obtain:

$$\sum_{j=1}^{n} \sum_{i=1}^{n} c_i \phi_i(X_j) \phi_\ell(X_j) = \sum_{j=1}^{n} F(X_j) \phi_\ell(X_j) \qquad \ell = 1,2,..,n$$

which can be rewritten as

$$\sum_{X \in M} \sum_{i=1}^{n} c_i \phi_i(X) \phi_\ell(X) = \sum_{X \in M} F(X) \phi_\ell(X) \qquad M \equiv \{X_1, X_2, \ldots, X_n\},\ \ell = 1,2,\ldots,n.$$

Interchanging the summations in the left hand term, we get

$$\sum_{i=1}^{n} c_i \sum_{X \in M} \phi_i(X)\phi_\ell(X) = \sum_{X \in M} F(X)\phi_\ell(X) \qquad \ell = 1,2,\ldots,n$$

where the summations over X can be interpreted as scalar products for $W(X) \equiv 1$. We hence end up with

$$\sum_{i=1}^{n} \langle \phi_i, \phi_j \rangle c_i = \langle F, \phi_j \rangle \qquad j = 1, 2, \ldots, n$$

which is the system of normal equations for the least-squares best-fitting P_n (for $W(X) \equiv 1$). Hence both polynomials are identical.

5.1 Problems

1. Prove by direct computation (choose your own F) that the interpolating algebraic polynomial of second degree (using three interpolation nodes) is identical to the best fitting least squares algebraic polynomial of second degree for F given on three points ($M \equiv \{ x_1, x_2, x_3 \}$).

6. APPLICATIONS OF THE LEAST SQUARES APPROXIMATION

6.1 Fourier Series

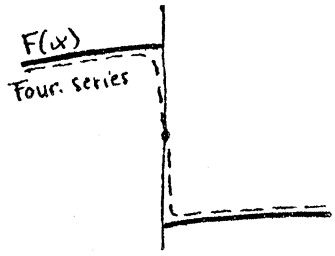
6.1.1 Generalized Fourier Series

It is conceivable that when working with compact M we may increase the number of functions in the base beyond all limits, since a compact set contains infinitely many points. Then the best approximating polynomial in the least-squares sense becomes an infinite series. If the base is, in addition, orthogonal then the series is known as generalised Fourier series.

It can be shown that if the base is complete, i.e. if there exists no other function apart from $f(x) = 0$ that is orthogonal to all of the base functions, than the generalised Fourier series equals to the given function F in the "mean sense". This is usually written as

$$F(x) \overset{\mu}{=} \lim_{n\to\infty} P_n(x) \ .$$

The equality in the "mean sense" becomes the identity when F is continuous on M . When F is not continuous but differentable at x it is approximated as shown on the diagram:



Here, as well as in the least-squares approximation, the integrability of F on M is a prerequisite. Note that the best approximating polynomials (in the least-squares sense) for an orthogonal basis can be regarded also as truncated Fourier series.

The overall behaviour of the coefficients c_i of a generalized Fourier series, known as Fourier coefficients $c_i = \langle F, \phi_i \rangle / ||\phi_i||^2$, is governed by Bessel's inequality

$$\sum_{i=1}^{\infty} (c_i \, ||\phi_i||)^2 \leq \langle F, F \rangle .$$

The equality is valid for continuous F. This can be shown as follows:

$$F = \sum_{i=1}^{\infty} c_i \, \phi_i = \sum_{i=1}^{\infty} \frac{\langle F, \phi_i \rangle}{||\phi_i||^2} \, \phi_i .$$

This equation can be multiplied by W F and integrated with respect to X over M. We obtain

$$\int_M W(X) \, F^2(X) dX = \int_M W(X) \, F(X) \sum_{i=1}^{\infty} \frac{\langle F, \phi_i \rangle}{||\phi_i||^2} \, \phi_i(X) \, dX$$

or

$$\langle F, F \rangle = \sum_{i=1}^{\infty} \frac{\langle F, \phi_i \rangle}{||\phi_i||^2} \int_M W(X) \, F(X) \, \phi_i(X) \, dX = \sum_{i=1}^{\infty} \frac{\langle F, \phi_i \rangle^2}{||\phi_i||^2} =$$

$$= \sum_{i=1}^{\infty} c_i^2 \, ||\phi_i||^2 .$$

The Bessel's inequality ensures that the Fourier series converges.

Note that exactly the same treatment can be adopted for two-dimensional orthogonal systems. The resulting infinite series in two-dimensions is known as Fourier double-series. The infinite series of spherical harmonics is one possible example.

There is evidently a close relationship between the Fourier coefficients and the coefficients α_{ji} used for orthogonalization (see 3.8). It is left to the reader to interpret the orthogonalization coefficients in terms of Fourier coefficients.

6.1.2 Trigonometric Fourier Series

Perhaps, the most important of all the Fourier series is the one based on the trigonometric system of functions. It became subsequently known as trigonometric Fourier series. In fact, many authors when writing about Fourier series mean really trigonometric Fourier series.

Since it is so important we shall state here the particular form of this Trigonometric Fourier series, although it should be evident from our earlier explanations. It reads, for an integrable function F defined on $M \equiv [-\ell, \ell]$:

$$F(X) \overset{\mu}{=} \sum_{i=0}^{\infty} \left(a_i \cos \frac{\Pi}{\ell} iX + b_i \sin \frac{\Pi}{\ell} iX\right)$$

where

$$a_o = \frac{1}{2\ell} \int_{-\ell}^{\ell} F(X)\, dX$$

$$a_i = \frac{1}{\ell} \int_{-\ell}^{\ell} F(X) \cos \frac{\Pi}{\ell} iX dX \qquad i = 1, 2, \ldots .$$

$$b_i = \frac{1}{\ell} \int_{-\ell}^{\ell} F(X) \sin \frac{\Pi}{\ell} iX dX$$

Note that the Trigonometric Fourier series of an odd function F defined on $[-\ell, \ell]$ contains only the sine terms and that of an even function only the cosine terms. These are usually called sine or cosine series respectively. This is a consequence of a more general feature of all the Fourier series based on the commonly used orthogonal systems. We can see, that the commonly used systems, defined on symmetrical intervals

and for even weights, can all be split into odd and even functions. It is known that a product of odd and even functions is an odd function. A finite integral over a symmetrical interval of an odd function equals zero. Hence the coefficients by the odd functions become zero for F even and vice versa. This can be expressed in a form of a principle that even (odd) function cannot be very well represented by odd (even) functions or their linear combinations on a symmetrical interval.

In connection with trigonometric Fourier series one should quote the Riemann's theorem. It reads: for an absolutely integrable F (i.e. $\int_a^b |F(X)|\, dX$ exists) the following equality holds

$$\lim_{\omega\to\infty} \int_a^b F(X) \sin \omega X dX = 0$$

even for $a \to -\infty$ and/or $b \to \infty$. It can also be written as

$$\lim_{i\to\infty} b_i = 0$$

and we can see that the series of b coefficients converges to zero.

It is not difficult to foresee that similar theorems should be valid for the coefficients of other orthogonal systems too, since the convergence of the series of the coefficients is ensured. However, the proofs of these are somewhat involved and not readily accessible in the literature.

6.1.3. Trigonometric Fourier series in complex form

The trigonometric Fourier series is very often formulated using complex numbers. The reason for it is that the complex form is much simpler and clearer.

To arrive at the complex form let us first express the trigonometric functions we shall be dealing with in complex form. Using de Moivre's theorem we can write:

$$\cos \frac{\pi}{\ell} nx = \cos nt = \frac{1}{2} (e^{int} + e^{-int})$$

$$\sin \frac{\pi}{\ell} nx = \sin nt = \frac{1}{2i} (e^{int} - e^{-int}) ,$$

where we use the index n instead of i in 6.1.2. and i equals now $\sqrt{-1}$ Hence one term of the trig. Fourier series, known as a <u>trigonometric term</u>, can be written as:

$$a_n \cos nt + b_n \sin nt = \frac{a_n}{2} (e^{int} + e^{-int}) + \frac{b_n}{2i} (e^{int} - e^{-int}) =$$

$$= \frac{a_n}{2} (e^{int} + e^{-int}) - i \frac{b_n}{2} (e^{int} - e^{-int})$$

$$= \frac{a_n - ib_n}{2} e^{int} + \frac{a_n + ib_n}{2} e^{-int} =$$

$$= c_n e^{int} + c_{-n} e^{-int} .$$

Note that the two coefficients c_n, c_{-n} are complex conjugates. The subscript -n was chosen to allow the following interpretation. Substituting the last result back into the original trig. Fourier series we obtain

$$f(X) \stackrel{\mu}{=} \sum_{n=0}^{\infty} (c_n e^{int} + c_{-n} e^{-int}) = \sum_{n=-\infty}^{\infty} c_n e^{int} = \sum_{n=-\infty}^{\infty} c_n e^{in\frac{\pi}{\ell}X} .$$

For the coefficients c_n we get:

$$c_n = \frac{1}{2} (a_n - ib_n) = \frac{1}{2} [\frac{1}{\ell} \int_{-\ell}^{\ell} f(X) \cos \frac{\pi}{\ell} nX dX - i \frac{1}{\ell} \int_{-\ell}^{\ell} f(X) \sin \frac{\pi}{\ell} nX dX]$$

$$= \frac{1}{2\ell} [\int_{-\ell}^{\ell} f(X) (\cos \frac{\pi}{\ell} nX - i \sin \frac{\pi}{\ell} nX) dX]$$

$$= \frac{1}{2\ell} \int_{-\ell}^{\ell} f(X) e^{-in \frac{\pi}{\ell} X} .$$

The coefficient c_{-n} can then be easily obtained from c_n being its complex conjugate.

Note that the complex series is real. It can be shown, denoting the n-th term of the complex series by $f_n(x)$, that $f_n(x) + f_{-n}(x)$ is real for every $n = 0, 1, 2, \ldots$ The proof is left to the reader.

6.2 Harmonic Analysis

One special problem of least-squares approximation, harmonic analysis, is worth mentioning here separately. We speak about harmonic analysis when the given function F, periodic on $M \equiv [a, b]$ (i.e. $F(x) = F(x + b - a)$) is known at $2n + 1$ points $x_i = a + \frac{b-a}{2n} i$ $(i = 0, 1, 2, \ldots, 2n)$ and we want to approximate it using the best-fitting, in the least-squares sense, trigonometric polynomial

$$P_{2k+1}(x) = \sum_{j=0}^{k} (a_j \cos j\xi + b_j \sin j\xi), \quad \xi = \xi(x).$$

The system $\Phi \equiv \{1, \cos \xi(x), \sin \xi(x), \cos 2\xi(x), \ldots \sin k\xi(x)\}$ has to be orthogonal on M.

First of all let us realize that because of the periodicity of F we have

$$F(x_o) = F(x_{2n})$$

so that we have to use only 2n functional values, say $F(x_i)$ $i = 1, 2, \ldots, 2n$. Hence we can accommodate only up to 2n functions in the system Φ and $2k + 1 \leq 2n$. This condition can be rewritten as

$$k \leq n - 1/2 \quad .$$

The second question is what kind of transformation $\xi = \xi(x)$

do we have to use to make the system Φ orthogonal on

$$M \equiv \{x_1, x_2, \ldots, x_{2n}\} ?$$

We know that the system $\Phi(\xi) = \{1, \cos \xi, \sin \xi, \ldots, \sin k\xi\}$ is orthogonal for $\xi_i = -\pi + \frac{\pi}{n} i$, $i = 1, 2, \ldots 2n$ (see 3.7). Hence the transformation

$$\frac{x - a}{b - a} = \frac{\xi + \pi}{2\pi}$$

ensures that the system $\Phi(\xi(x)) = \Phi(x)$ will be orthogonal for $x_i = a + \frac{b - a}{2n} i$, $i = 1, 2, \ldots, 2n$. Therefore

$$\xi = \frac{2\pi}{b-a} x - \pi \left(\frac{b}{b-a}\right) = K_1 x + K_2$$

is the transformation we are looking for and the trigonometric system (base) to be used in the problem is

$$\Phi \equiv \{1, \cos (K_1x+K_2), \sin (K_1x+K_2), \cos 2(K_1x+K_2), \ldots, \sin k(K_1x+K_2)\}.$$

Since the base is orthogonal on M we get a diagonal matrix of normal equations and the coefficients of P_{2k+1} become:

$$a_o = \langle F, 1\rangle / |||1|||^2 = \frac{1}{2n} \sum_{i=1}^{2n} F(x_i)$$

$$a_j = \frac{\langle F, \cos j(K_1x+K_2)\rangle}{||\cos j(K_1x + K_2)||^2} = \frac{1}{n} \sum_{i=1}^{2n} F(x_i) \cos j(K_1x_i + K_2)$$

$$b_j = \frac{\langle F, \sin j(K_1x+K_2)\rangle}{||\sin j(K_1x+K_2)||^2} = \frac{1}{n} \sum_{i=1}^{2n} F(x_i) \sin j(K_1x_i + K_2). \qquad j = 1, 2, \ldots, k.$$

When evaluating the scalar products we can exploit the property of the trigonometric functions:

$$\sin j\xi_i = -\sin j\xi_{2n-i}$$

$$\cos j\xi_i = \cos j\xi_{2n-i} \qquad j = 1, 2, \ldots, k, \quad i = 1, 2, \ldots, 2n-1.$$

We can write:

$$\sum_{i=1}^{2n} F(\xi) \cos j\xi_i = \frac{1}{2} \sum_{i=1}^{n-1} (F(\xi_i) + F(\xi_{2n-i})) \cos j\xi_i + F(\xi_n) - F(\xi_{2n}) \qquad j = 1, 2, \ldots, k.$$

$$\sum_{i=1}^{2n} F(\xi_i) \sin j\xi_i = \frac{1}{2} \sum_{i=1}^{n-1} (F(\xi_i) - F(\xi_{2n-i})) \sin j\xi_i .$$

The derivation of the above formula is left to the reader. We shall just remark that this computational trick reduces the computing time considerably and is often used whenever we deal with evaluation of scalar products involving odd and even functions defined on a symmetrical M. Note that the problem of harmonic analysis can be also regarded as the problem of computing the truncated trig. Fourier series with the integrals being numerically evaluated using the rectangular formula. The harmonic analysis has a wide field of applications whenever the periodic functions are investigated.

6.3 Fourier Integral

If a function F, given on $M \equiv (-\infty, \infty)$ is on M absolutely integrable then it can also be represented in terms of the Fourier integral

$$F(x) \overset{\mu}{=} \int_0^\infty (a(\omega) \cos\omega x + b(\omega) \sin\omega x)\, d\omega$$

where

$$a(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} F(x) \cos\omega x\, dx$$

$$b(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} F(x) \sin\omega x\, dx .$$

The Fourier integral is obviously a close relative of the Fourier trig. series. It can be derived from the trigonometric Fourier series when the finite interval $[-\ell, \ell]$, for which the trig. Fourier series is defined, is extended beyond all the limits. We can see that for the infinite M, the trig. Fourier series ceases to be defined and has to be replaced by the Fourier integral.

The only two differences between the trigonometric series and the integral are:

1) the coefficients a, b in the Fourier integral are continuous functions of the frequency ω as compared to the coefficients a_n, b_n of the series which are discrete functions of the frequency $\frac{\pi}{\ell} n$;

2) the summation over the discrete frequencies $\frac{\pi}{\ell} n$ in the trigonometric series is here replaced by the integration over the continuous frequency ω.

Fourier integral represents the given function F in the same way, i.e. in the mean sense, as the Fourier series does. The same theorems about representation of odd and even functions by cosine and sine terms only are valid even for the integral.

It is left to the reader to show that the Fourier integral can be written, using complex numbers, as

$$F(x) \overset{\mu}{=} \int_{-\infty}^{\infty} c(\omega) e^{i\omega x} \, d\omega$$

where

$$c(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(x) e^{-i\omega x} dx.$$

6.4 Spectral Analysis

6.4.1 Problem of spectral analysis

In dealing with various time series f(t) in practice we are very often faced with following problem: "given a time series f(t) defined on M (discrete) we assume that f can be expressed as

$$f(t) = \sum_{i=1}^{m} (a_i \cos \omega_i t + b_i \sin \omega_i t),$$

where ω_i, a_i, b_i are some real numbers. The task is to evaluate ω_i, a_i, b_i, $i = 1, 2, \ldots, m$."

If we formulate the problem as a problem of the best least-squares approximation, we would end up with a system of 3m normal equations that would not be linear. This is because the generalized trigonometric polynomial above contains the frequencies ω_i as paramaters within the functions. The solution of such non-linear system of equations may or may not exist. Even if it did, the numerical computation would be very awkward and most troublesome.

On the other hand, if the frequencies ω_i were known, the coefficients a_i, b_i could be determined quite easily using the least-squares technique. They would be furnished by a system of 2m linear algebraic equations, that usually has a solution.

The problem of finding the unknown frequencies ω_i constitutes the proper problem of spectral analysis. Hence the spectral analysis can be considered as the first step in the complete decomposition of the given function f (Usually a time series, hence the use of it) into the individual trigonometric terms. This type of analysis is of a great importance and finds applications in electrotechnique, oceanography, meteorology, geophysics, biology, medicine, statistics and many other fields.

Precise solution of this problem is seldom possible. In the next section we shall show one of the most widely used techniques of spectral analysis.

6.4.2 Fourier spectral analysis

To begin with, let us write the trigonometric Fourier series of the time function f(t) defined on a finite compact $M \equiv [-\ell, \ell]$:

$$f(t) \overset{M}{=} \sum_{n=0}^{\infty} (a_n \cos \omega_n t + b_n \sin \omega_n t),$$

where the frequency ω_n is given by $\frac{\pi}{\ell} n$ (see 6.1.2). This can be rewritten as

$$f(t) \overset{M}{=} \sum_{n=0}^{\infty} R_n (\cos \nu_n \cos \omega_n t + \sin \nu_n \sin \omega_n t)$$

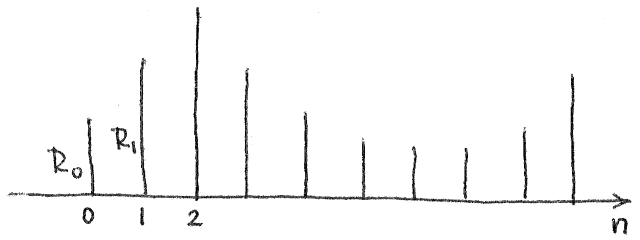
$$= \sum_{n=0}^{\infty} R_n \cos (\omega_n t - \nu_n)$$

where

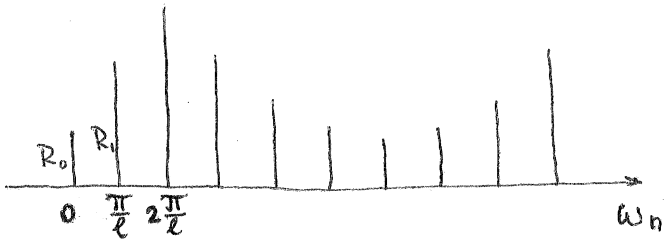
$$R_n \cos \nu_n = a_n, \qquad R_n \sin \nu_n = b_n \text{ and therefore}$$

$$R_n = \sqrt{(a_n^2 + b_n^2)}, \qquad \nu_n = \operatorname{arctg} \frac{b_n}{a_n} \quad .$$

Each amplitude R_n is obviously a non-negative real number that describes the magnitude of the cosinusoidal wave $\cos (\omega_n t - \nu_n)$. These magnitudes can be plotted either against the subscript n:



or against the frequencies $\omega_n = \frac{\pi}{\ell} n$



This diagram is known as the line spectrum of f and represents the discrete transformation of f from time space into the frequency space

$$f(t) \rightarrow R(\omega)$$

(actually $R(\omega_i)$ in this discrete case).

Such a diagram is evidently able to provide some information about the contribution of the individual integer frequencies.

The Fourier integral can be interpreted in the same way, to furnish a continuous spectrogramme, i.e. for any real value of n in $\omega_n = \frac{\pi}{\ell} n$,

Not just for integral values of n. $R(\omega)$ given by

$$R(\omega) = \sqrt{(a^2(\omega) + b^2(\omega))}$$

is a continuous function of ω known as the Fourier transform of f. Its graphical representation is called Fourier periodogramme and used extensively in all kinds of experimental sciences.

If f is known on a finite M only, and this is usually the case with experimental functions, then we define its extension $\tilde{f}$

$$\tilde{f}(t) = \begin{cases} f(t) & t \in M \\ 0 & t \notin M \end{cases}$$

and regard the periodogramme of $\tilde{f}$ as an approximation of that of f. The integrals involved in evaluating $a(\omega)$, $b(\omega)$ are generally evaluated also only approximately since M is seldom compact.

$R(\omega)$ is then analyzed for peaks or "characteristic peaks". The number of these peaks decide the number of terms in the expansion for f(t) in section 6.4.1. The location of the peaks (ω values) become the ω_i in this expansion.

6.5 Problems

1. What does the trigonometric Fourier series of the function $F(x) = |\sin x|$ on $x \in M \equiv [-\pi, \pi]$ look like?

2. Express the function

$$F(x) = \begin{cases} 1 & \text{if } |x| < 1 \\ 1/2 & |x| = 1 \\ 0 & |x| > 1 \end{cases}$$

in terms of its Fourier integral.

3. What is the line spectrum of

$$f(t) = 5 + 3\cos(3t - 5.8) - 2\sin(5t + 3.1) \ ?$$

4. Write a computer program to compute the values $a(\omega)$, $b(\omega)$, $\mu(\omega)$, $R(\omega)$ for any function given on an equidistant discrete set M.

5. Using $\phi(X) = R(x)\,R(y)$, where R is the trigonometric system, design the algorithm for a computer program that would compute the best fitting double trigonometric polynomial to any topographic surface given on a rectangular, equidistant grid. Can you use any other system? What about mixed algebraic polynomials?

6. Write, debug and document a program that would approximate a function F given on a grid $5^g \times 5^g$ on the surface of a sphere by a series of spherical harmonics. Consider P_{nm} to be orthogonal on the grid and leave the choice of the highest order of spherical harmonics up to the user. Include a check for the proper selection of the highest order. Print out the estimates of variances of the individual coefficients.

APPENDIX: SOME ORTHOGONAL SYSTEMS OF FUNCTIONS

Following are the most widely used systems of functions, orthogonal on compact sets:

i) Tchebyshev's of 1-st kind, which we have met in 2 , orthogonal on $M \equiv [-1,1]$ with $W(X) = (1 - X^2)^{-1/2}$.

ii) Tchebyshev's of 2-nd kind:

$$U_K(X) = \frac{1}{1+k} \frac{d}{dx} T_{K+1}(X)$$

$$= \frac{\sin((1+k) \arccos X)}{\sqrt{(1 - X^2)}} ,$$

orthogonal on $M \equiv [-1,1]$ with $W(X) = (1 - X^2)^{1/2}$.

iii) Legendre's

$$L_o(X) = 1, \quad L_k(X) = \frac{1}{2^k k!} \frac{d^k}{dX^k} (X^2 - 1)^k ,$$

orthogonal on $M \equiv [-1,1]$ with $W(X) = 1$.

iv) Legendre's associated functions

$$P_{nm}(X) = (1 - X^2)^{\frac{m}{2}} \frac{d^m}{dX^m} L_n(X)$$

are orthogonal with respect to n (only for the same value of m) on $M \equiv [-1,1]$ with $W(X) = 1$. These are excessively used in geodesy within the "spherical harmonics".

v) Trigonometric system

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

is orthogonal on $M \equiv [-\pi,\pi]$ with $W(X) = 1$. It can be shown that this system is orthogonal even on discrete equidistant set $M \equiv [-\pi,\pi)$ or $(-\pi,\pi]$ (note that in this case the interval is open from one side).

Other systems for which we are not going to list the formulae are:

vi) Laguerre's orthogonal on $M \equiv [0,\infty)$ with $W(X) = e^{-X}$.

vii) Generalized Laguerre's orthogonal on $M \equiv [0,\infty)$ with weight $W(X) = X^{\alpha} e^{-X}$, where $\alpha > -1$ is a parameter of the system.

viii) Hermite's orthogonal on $M \equiv (-\infty,\infty)$ with $W(X) = e^{-X^2}$.

ix) Bessel's orthogonal on $M \equiv [0,c]$ with $W(X) = X$, where c is a parameter of the system. They are of the 1st, 2nd and 3rd kinds.

x) Jacobi's orthogonal on $M \equiv [-1,1]$ with $W(X) = (1-X)^{\alpha} (1+X)^{\beta}$ where $\alpha,\beta > -1$ are parameters of the system.

All these systems can be also regarded as series of eigenfunctions of some particular differential equation of 2nd order (of the Sturm-Luiville's type). Each of them has got its own "generating function", i.e., a function that developed into power series has coefficients equal to the individual functions of the system. For the individual functions in the different systems of algebraic polynomials, there exist recursive formulae, known sometimes as Rodriguez's, expressing the n-th element as a linear combination of (n-1)st and (n-2)nd elements.

If a system of functions is orthogonal for $X \in M \equiv [a,b]$, then a system of functions of the argument

$$\xi = \alpha + \frac{\beta-\alpha}{b-a} (X-a) = k_1 + k_2 X$$

is orthogonal on $M \equiv [\alpha,\beta]$. This equation follows simply from proportionality equation

$$\frac{X - a}{b - a} = \frac{\xi - \alpha}{\beta - \alpha}$$

as the reader can easily prove. Note that if a system is orthogonal on an infinite interval (Laguerre's, Hermite's) there does not exist any linear transformation for the argument that would make the system orthogonal on a finite interval ($b-a \to \infty$ and $(\beta-\alpha)/(b-a) \to 0$).

We may also note that λ-times an orthogonal system is again an orthogonal system as long as λ is not a function of X. λ can be even different for different functions of the system.

Many orthogonal systems of functions can be expressed as special cases of a generalized system of functions called the Generalized Hypergeometric Series, which is written

$$ {}_pF_q\ (\alpha_1,\alpha_2,\ldots,\alpha_p;\ \rho_1,\rho_2,\ldots\rho_q;\ x) $$

$$ = \sum_{n=0}^{\infty} \frac{(\alpha_1)_n\ (\alpha_2)_n \ldots (\alpha_p)_n}{(\rho_1)_n\ (\rho_2)_n \ldots (\rho_q)_n} \frac{x^n}{n\ !} $$

where Pockhammer's Symbol $(\alpha)_r$ is defined in terms of gamma functions

$$ (\alpha)_r = \frac{\Gamma\ (\alpha+r)}{\Gamma\ (\alpha)} $$

and $\alpha_1,\alpha_2,\ldots\alpha_p$ and $\rho_1,\rho_2,\ldots\rho_q$ are constants.

For example

$$ T_n(x) = {}_2F_1\ \left(-n,n;1/2;\ \frac{1-x}{2}\right); \qquad \text{Tchebyshev of 1st kind} $$

$$ U_n(x) = (n+1)\ {}_2F_1\ \left(-n,n+2;3/2;\ \frac{1-x}{2}\right) \qquad \text{Tchebyshev of 2nd kind} $$

$$ L'_n(x) = {}_2F_1\ \left(-n,n+1;1;\frac{1-x}{2}\right) \qquad \text{Legendre.} $$

Following are tables giving the lower degree polynomials in some of these systems of orthogonal functions.

TCHEBYSHEV 1-ST KIND

k	$T_k^*(x)$	k	$T_k^*(x)$
0	1	5	$16x^5 - 20x^3 + 5x$
1	x	6	$32x^6 - 48x^4 + 18x^2 - 1$
2	$2x^2 - 1$	7	$64x^7 - 112x^2 + 56x^3 - 7x$
3	$4x^3 - 3x$	8	$128x^8 - 256x^6 + 160x^4 - 32x^2 + 1$
4	$8x^4 - 8x^2 + 1$	9	$256x^9 - 576x^7 + 432x^5 - 120x^3 + 9x$
$T_n^*(x) = 2^{n-1}T_n(x).$			

TCHEBYSHEV 2-ND KIND

k	$U_k(x)$	k	$U_k(x)$
0	1	5	$32x^5 - 24x^3 + 6x$
1	$2x$	6	$64x^6 - 80x^4 + 24x - 1$
2	$4x^2 - 1$	7	$128x^7 - 192x^5 + 80x^3 - 8x$
3	$8x^3 - 4x$	8	$256x^8 - 448x^6 + 240x^4 - 40x^2 + 1$
4	$16x^4 - 12x^2 + 1$	9	$512x^9 - 1026x^7 + 672x^5 - 160x^3 + 10x$

LAGUERRE

k	$Q_k(x)$
0	$+ 1$
1	$- x + 1$
2	$x^2 - 4x + 1$
3	$-x^3 + 9x^2 - 18x + 6$
4	$x^4 - 16x^3 + 72x^2 - 96x + 24$
5	$-x^5 + 25x^4 - 200x^3 + 600x^2 - 600x + 120$

HERMITE

k	$H_k(x)$
0	1
1	$2x$
2	$4x^2 - 2$
3	$8x^3 - 12x$
4	$16x^4 - 48x^2 + 120x$
5	$32x^5 - 160x^3 + 120x$

LEGENDRE

k	$L_k(x)$	k	$L_k(x)$
0	1	5	$(1/8)\ (63x^5 - 70x^3 + 15x)$
1	x	6	$(1/16)\ (231x^5 - 315x^4 + 105x^2 - 5)$
2	$(1/2)\ (3x^2 - 1)$	7	$(1/16)\ (429x^7 - 693x^5 + 315x^3 - 35x)$
3	$(1/2)\ (5x^3 - 3x)$	8	$(1/128)\ (6435x^8 - 12012x^6 + 6930x^4 - 1260x^2 + 35)$
4	$(1/8)\ (35x^4 - 30x^2 + 3)$	9	$(1/128)\ (12155x^9 - 25730x^7 + 18018x^5 - 4620x^3 + 315x)$

SUGGESTIONS FOR FURTHER READING

a) Approximation Theory

Cheney, E.W. (1966). "Introduction to Approximation Theory", McGraw-Hill.

Berezin, I.S. and Zhidkov, N.P. (1965). "Computing Methods" 2 Vols. Addison-Wesley.

b) Set theory

Lipschutz, S. (1964). "Set Theory and Related Topics" Schaum's Outline

c) Functional Analysis

Rudin, W. (1964). "Principles of Mathematical Analysis" McGraw-Hill.

Yosida, K. (1965). "Functional Analysis" Springer.

d) Orthogonal Functions

Abramowitz, M. and Stegun, I.A. (1965). "Handbook of Mathematical Functions" Dover.

Sneddon, I.N. (1961). "Special Functions of Mathematical Physics and Chemistry" Oliver and Boyd.

Courant, R. and Hilbert, D. (1953). "Methods of Mathematical Physics" 2 Vols, Wiley (see especially Vol. 1, Chapter 7).

e) Least Squares Adjustment

Thompson, E.H. (1969). "Introduction to the Algebra of Matrices with some Applications" University of Toronto Press (see especially Chapter 10).

Hamilton, W.C. (1964). "Statistics in Physical Science" Ronald.

Wells, D.E. and Krakiwsky, E.J. (1971). "The Method of Least Squares" U.N.B. S.E. Lecture Notes #18.

Vaníček, P. (1971). "Introduction to Adjustment Calculus" U.N.B. S.E. Lecture Notes #20.

f) Applications of Least Squares Approximation

Lanczos, C. (1957). "Applied Analysis" Prentice-Hall.

Zygmund, A. (1955). "Trigonometrical Series" Dover.

Blackman, R.B. and Tukey, J.W. (1959). "Measurement of Power Spectra" Dover.

g) n-dimensional geometry

Sommerville, D.M.Y. (1929). "An Introduction to the Geometry of N Dimensions" Dover edition 1958.