1969

Error direction dependence and best straight line approximations

Approved

by

First Reader [redacted]

Professor of Mathematics

Second Reader [redacted]

Professor of Mathematics

OUTLINE

INTRODUCTION

Let f(x) be an arbitrary given continuous real valued function on [0,1]. Let $G=\{g(x)\}$ be a given set of admissible approximations, and let $||.||$ be an appropriate norm. A typical approximation problem concerns finding a g^* belonging to G such that

$$||f-g^*|| \leq ||f-g||$$

for all $g \varepsilon G$.

Implicit in the above formulation is the assumption that the direction perpendicular to the x axis is the most appropriate direction in which to measure errors. There are important cases when this assumption is justified: e.g., in linear regression problems where it is assumed that inaccuracies are present in the ordinate but not the abcissa. There are, however, many practical cases where this assumption doesn't hold: e.g., when both variables are subject to inaccuracies. Within the context of a linear regression problem, if we were to reverse the roles of the experimental and controlled variables, the regression line would, in general, be different from the original regression line. Reversing the roles of the variables is equivalent to measuring error in the direction parallel to the abcissa. Thus we see that best approximations can depend very strongly on the direction of measuring the error.

CHAPTER I

PRELIMINARIES

1.1 THE PROBLEM

We will restrict our attention to approximating defined and continuous real valued functions on the interval [0.1]. Our approximating functions will be straight lines: ax+b. The direction of measurement, the angle β , will be that angle formed by rotating a line perpendicular to the x axis through an angle β, the counterclockwise sense taken as positive. From the definition it follows that the slope of lines with direction of measurement β is $\tan(\beta+\frac{\pi}{2})$.

1.2 THE ERROR FUNCTION $e_\beta(x)$

Given a function f(x) on [0,1] a straight line ax+b and a direction of measurement β we wish to find the error from the function to the approximating straight line for all directions β. If (x,f(x)) is a point on the curve, let s be the abcissa of the straight line ax+b which corresponds to the point (x,f(x)) for direction of measurement .

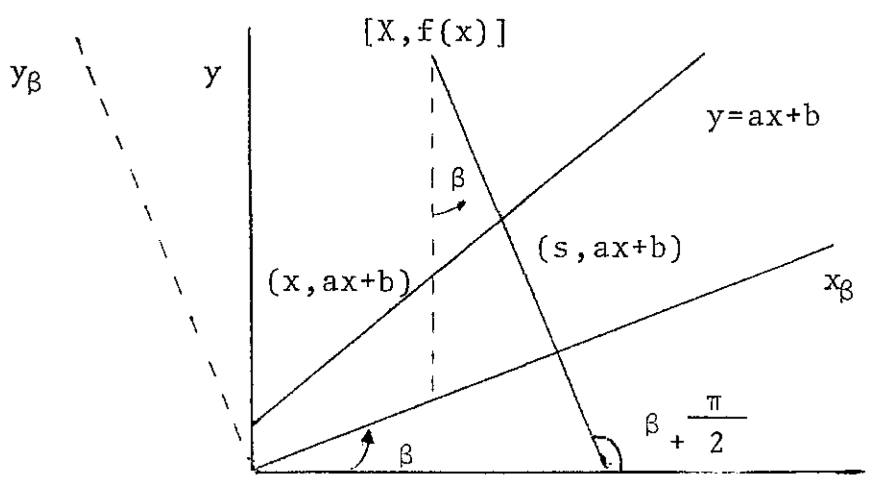

FIG. 1

taking differences along direction of measurement β. The general equations relating points in the original coordinate system to the β rotated coordinate system is

$$x_\beta(x)=x\cos\beta+f(x)\sin\beta$$
$$y_\beta(x)=-x\sin\beta+f(x)\cos\beta \qquad (3)$$

$x_\beta(x)$ is the abcissa in the new frame of reference. A straight line is thus $cx_\beta(x)+d$. If we are given the straight line as+b in the original coordinate system, the rotated version of ax+b must be equal to $cx_\beta(x)+d$. The equations relating the slopes and y intercepts of the same straight line in the two coordinate systems is:

$$c=(a\cos\beta -\sin\beta)/(a\sin\beta+\cos\beta)$$
$$d=b/(a\sin\beta+\cos\beta) \qquad (4)$$
$$a=(\sin\beta+c\cos\beta)/(\cos\beta-c\sin\beta)$$
$$b=d/(\cos\beta -c\sin\beta)$$

Within the new frame of reference, the error between f(x) and the straight line ax+b is

$$e_\beta(x)=y_\beta(x)-cx_\beta(x)-d=(f(x)-ax-b)/a\sin\beta+\cos\beta) \qquad (5)$$

the second equality following from (4) and the definitions (3) where $x_\beta(x)$ is the abcissa of the point (x,f(x)) in the rotated coordinate system.

The region enclosed by $e_\beta(x)$ and the x_β axis is, in general, open at the end points. We can close the region by including perpendicular lines from the end points to the x_β axis. Hence, the area enclosed by $e_\beta(x)$ is the region R bounded by the closed curve C consisting of the curve $e_\beta(x)$, the perpendicular distance from one end point to the x_β axis, the x_β axis from one end point perpendicular to the other, and a perpendicular from the x_β axis to the second end point of $e_\beta(x)$. We will denote these four pieces of C by C_1, C_2, C_3, C_4, all of which are rectifiable Jordan arcs. Since $f(x)$ is continuous on [0,1], it is bounded and hence $e_\beta(x)$ is bounded. Thus the region enclosed by C is a closed bounded region of E_2.

C is not, in general, a Jordan curve, since a straight line may intersect a Jordan arc infinitely often. For most functions of interest, however, this is not the case. Hence, in the sequel, we shall assume that f satisfies the condition that no straight line intersects the function infinitely often. Thus by the indicated construction of C and the restrictions of f we can satisfy the hypothesis of Green's theorem. The area is given by (7) for C and we have

$$A=\int_C -y\,dx =\int_{C_1} -y\,dx+\int_{C_2} -y\,dx+\int_{C_3} -y\,dx+\int_{C_4} -y\,dx \quad (8)$$

At C_2 and C_4, dx is zero. Along C_3 the ordinate is zero. Therefore we have the result

CHAPTER II

L_∞ NORM

2.1 DEFINITIONS AND CHARACTERIZATION OF BEST APPROXIMATIONS

The L_∞ or Chebychev norm of a continuous real valued function f is

$$||f||_\infty = \max_{x\,[0,1]} |f(x)| \quad (10)$$

For polynomial and therefore straight line approximations the best approximation exists, is unique and is characterized by the fact that the maximum value of the error in absolute value is attained at at least n+2 points, where n is the degree of the polynomial, on the interval of approximation. [Handscomb, 1966, Chap. 7]. For our purposes, the characterization theorem implies that a best straight line approximation is one whose error function attains its max at, at least 3 points, in the interval [0.1].

2.2 SOLUTION

Let a*x+b be a best approximation to f for $\beta=0$. Then for arbitrary β, the error function $e_\beta(x)$ given by (2) for a*x + bx is

$$|e_\beta(x)| = |(f(x)-a^*x-b^*)|/|(a^*\sin\beta+\cos\beta)| \quad (11)$$

which has the characterization property for allβ . Thus we have proven the following:

Theorem 1: The best straight line approximation to a continuous real valued function f(x) on a closed bounded interval is independent of the direction of measuring

CHAPTER III

L_1 NORM

3.1 DEFINITIONS

The L_1 norm is defined as

$$||f||_1=\int_0^1|f(x)|dx \qquad (12)$$

The L_1 error for e (x) from (9) is

$$||e_\beta(x)||_1 =\int_0^1|e_\beta(x)|\, x'_\beta(x)dx \qquad (13)$$

3.2 SOME PROPERTIES OF THE INTEGRAL OF $e_\beta(x)$

Area is by definition invariant except for sign to the coordinate system in which it is measured. It may then appear that (9) must be invariant for all β. This is not so because the closed curve C is in general different for each β and not strictly the rigid rotation of the error in the original coordinate system. The change in areas under the curve $e_\beta(x)$ for different β results from the fact that areas are added or deleted at the ends of the curve for different β.

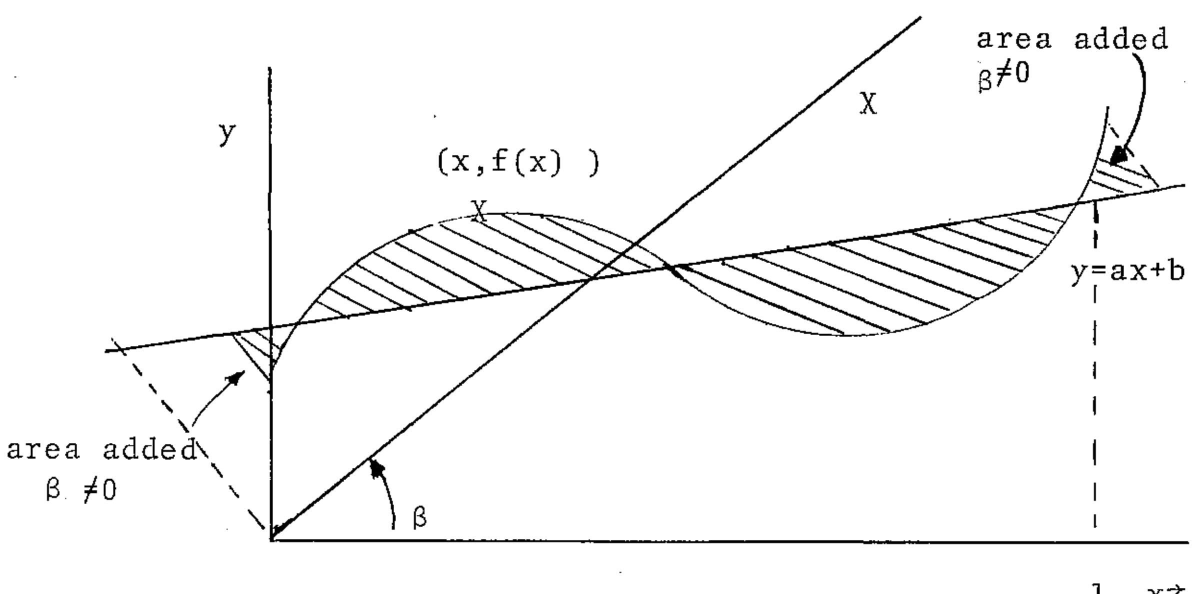

FIG. 2

approximation. A polygonal approximation is linear over subintervals of [0,1] and agrees with f at the ends of these subintervals. The ends of the subintervals are called nodes, and we assume that the points 0 and 1 are included among the nodes. For the purposes of the discussion here we will include in the set of nodes all of the points of intersection of the approximating lines in each subinterval. Let us call this expanded set of nodes the set U where U= $0=u_0<u<\ldots u_n=1$. Note that U is necessarily finite since we have restricted our definition of area to functions which are not intersected an infinite number of times by any straight line. The area over each subinterval is independent of β for all β, so that the area is independent of β. If we form the L_1 error over [0,1] then it is the sum of the L_1 errors over each subinterval. But by construction of the subintervals, the error function is either wholly positive or negative in the subinterval and the L_1 error in each case is thus exactly plus or minus the area of each subinterval. But each area, is independent of β, thus (13) must be also. Hence we have proven the following

Theorem 3: Polygonal approximation in L_1 is independent of β, the direction of measuring the error.

CHAPTER IV

DISCRETE LEAST SQUARE APPROXIMATION

4.1 INTRODUCTION

We are given a function f defined on a finite point set

$$X = \{x_i \mid i=1,2,\ldots,N\}$$

The error of the approximation at x_i is

$$e(x_i) = f(x_i)-ax_i-b$$

and we wish to minimize

$$F(a,b) = \sum_{i=1}^{N} (f(x_i)-ax-b)^2 \qquad (16)$$

Equation (16) is an inner product and the square root of (16) defines a norm. As we have set up the problem, the solution to (16) is the regression line, a subject of major importance in statistics.

Since the function F is convex in a and b [Rice, 1964, p. 31, Cheney, 1966, p.25] then setting the partials of F equal to zero defines the conditions for a minimum. Thus the a* and b* which minimize (16) must satisfy the conditions

$$\sum y_i x_i = a^* \sum x_i^2 + b^* \sum x_i \qquad (17)$$
$$\sum y_i = a^* \sum x_i + b^* N$$

which leads to the solutions

$$b^* = \bar{y} - a^*\bar{x}$$

$$a^* = \frac{N \sum y_i x_i - \sum x_i \sum y_i}{N \sum x_i^2 - (\sum x_i)^2}$$

over all a,b,β . By using the form of the error (5) we can put (26) in the more tractable form

$$F(c,d,\beta) = \sum (y_{i_\beta} - cx_{i_\beta} - d)^2 \qquad (27)$$

where

$$x_{i_\beta} = x_i \cos\beta + y_i \sin\beta$$

$$y_{i_\beta} = -x_i \sin\beta + y_i \cos\beta \qquad (28)$$

From (27) it is clear that F is a convex function in c and d. The analysis of (27) thus proceeds in the same manner as (16) in 4.1. We have the conditions for minimum

$$\sum y_{i_\beta} x_{i_\beta} = c^* \sum x_{i_\beta}^2 + d^* \sum x_{i_\beta} \qquad (29)$$

$$\sum y_{i_\beta} = c^* \sum x_{i_\beta} + d^* N$$

If we define

$$\bar{x}_\beta = \sum x_{i_\beta}/N$$

$$\bar{y}_\beta = \sum y_{i_\beta}/N \qquad (30)$$

and we assume that the data has been normalized for $\beta=0$, then $\bar{x}_\beta$ and $\bar{y}_\beta$ is zero for all β; i.e., the normalized origin is independent of β. We therefore assume the normalization condition (20) holds which again simplifies the solution for a minimum which is

$$d^* = 0$$

$$c^* = \sum x_{i_\beta} y_{i_\beta} / \sum x_{i_\beta}^2 \qquad (31)$$

and bounded.

4.4 ORTHOGONAL REGRESSION LINE AND PRINCIPAL AXIS OF INERTIA

The orthogonal regression line is that approximating straight line which minimizes the square perpendicular distance from the data points to the line of approximation. [Linnik,1961] We can state the orthogonal regression line problem within the context of the least squares problem where the approximating straight line is given in terms of the original coordinate system. Let β be a given direction of error measurement. The orthogonal regression line problem then restricts its attention to straight line approximations whose slope is tan β. Thus the orthogonal regression error in terms of (26) is

$$F(a,b,\beta) = F(\tan\beta, b,\beta) \qquad (38)$$

and the orthogonal regression line is that β^*, b^* such that

$$F(\tan\beta^*, b^*,\beta^*) \leq F(\tan\beta, b,\beta) \qquad (39)$$

for all β, b. For the least squares problem, the best approximation is given by a^*, b^*, β^* such that

$$F(a^*, b^*,\beta^*) \leq F(a,b,\beta) \qquad (40)$$

for all a,b,β.

For any given frame of reference β, it is true that the best approximation may not have $\tan\beta = a$; nevertheless, we can prove that for β^*, the best approximation has $\tan\beta^* = a^*$. Hence

Theorem 4: $F(a^*,b^*,\beta^*)=F(\tan\beta^*, b^*,\beta^*)$

$$\tan 2\beta = \frac{2K_{xy}}{\sigma_x^2 - \sigma_y^2} \tag{41}$$

There are two solutions $\beta \varepsilon [-\pi/2, \pi/2]$ corresponding to the absolute minβ^* and the absolute max. The unnormalized condition for β^* is

$$\tan 2\beta = \frac{2\Sigma(x_i - x)(y_i - y)}{\Sigma(x_i - \overline{x}) - \Sigma(y_i - \overline{y})} \tag{42}$$

Since $$a^* = \tan\beta^* \tag{43}$$

then $$b^* = \overline{y} - \overline{x}\tan\beta^* \tag{44}$$

since the line with slope $\tan\beta^*$ must go through the point $\overline{x}, \overline{y}$. Thus (42), (43) and (44) are the solution to the discrete least square approximation problem for all directions of measurement β, for unnormalized data.

4.6 REMARKS ON THE ORTHOGONAL REGRESSION LINE

From a curve fitter's point of view, the solution derived in 4.5 satisfies the fundamental requirement of minimizing square residuals without regard to any particular coordinate system for which the data happens to be given. From a statistician's point of view,however, there are some disturbing problems. It is shown in [Roos,1938], [Jones,1938] that the orthogonal regression line is not invariant with changes in scale in y or x. If, e.g., we change the scale in x to 2 x, the orthogonal regression line is not in general a linear change in the original

CHAPTER V

L_2 NORM

5.1 DEFINITIONS AND INTRODUCTION

The L_2 norm is defined as

$$||f||_2 = \left(\int_0^1 f(x)^2 dx\right)^{1/2} \tag{45}$$

and the best approximation in the L_2 norm by a straight line is to find a*, b* which minimize

$$C(a,b) = \int_0^1 (f(x)-ax-b)^2 dx \tag{46}$$

C(a,b) is a convex function in a and b and exactly the same analysis as in 4.1 applies for the solution a*,b*. The normalization conditions are

$$\int_0^1 x dx = 0 \qquad \int_0^1 f(x) dx = 0 \tag{46}$$

The normalized solutions are

$$b^* = 0$$
$$a^* = \int_0^1 x f(x) dx / \int_0^1 x \; dx = 3 \int_0^1 x f(x) dx \tag{47}$$

and if we define

$$\sigma_x^2 = \int_0^1 x^2 dx = 1/3$$
$$\sigma_y^2 = \int_0^1 f(x)^2 dx \tag{48}$$
$$K_{xy} = \int_0^1 x f(x) dx$$

then we can write the error of the best approximating straight line as

$$C(a^*,b^*) = \frac{\sigma_x^2 \quad \sigma_y^2 \quad -K_{xy}^2}{\sigma_x^2} \tag{49}$$

5.2 THE SQUARE L_2 ERROR FOR DIRECTION OF MEASUREMENT β

The square L_2 error for direction of measurement β is the integral of the square of the error function where it is

5.3 RESTRICTIONS ON THE RANGE OF β

The usefulness of a study on the dependence of best approximations on the direction of error measurement is to point out the existence of independence properties or to dismiss direction effects by finding the direction with least error without altering the essential characteristics of the curve being approximated. Consistent with these goals is to restrict our attention to the directions of measurement for which the curve is still a function within the context of the rotated coordinate system.

Definition: Let P be the set of all β such that the curve $y_\beta(x)$ is a function in the x_β, y_β coordinate system and $x_\beta'(x)$ is non-negative.

The characterization of the set P is not necessarily simple and is dependent on the particular function being approximated for such essential properties as open, half open, closed. For example the function x gives rise to the closed set P=[-27°, 90°]. The straight line y=c gives rise to P=(-90°, 90°).

A given direction of measurement β gives rise to a curve which isn't a function if at least two points on the curve intersect a line with slope tan β . Since $y_\beta(x)$ is continuous with continuous derivatives it is a smooth curve and therefore there must exist x_o in [0,1] such that the tangent of the curve in the x_β, y_β coordinate system by the mean value theorem of differential calculus

$$x_\beta(x) - x_\beta(x_o) = x'_\beta(\zeta)(x-x)_o \zeta\varepsilon(x_o,x)C[0,1] \quad (53)$$

and we can infer that $x_\beta(x)$ is a monotone increasing function of x.

Definition: Let P' be the set of all β such that

$$x'_\beta(x) > 0 \qquad x\varepsilon[0,1]$$

By definition P' is an open set and we set $P' = (\beta_1, \beta_2)$. Clearly, P' is contained in P and since $\beta = 0$ is in P', P and P' is non-empty. If

$$M = \max f'(x) \qquad m = \min f'(x)$$

then

$$P' = (\tan^{-1} M - \frac{\pi}{2}, \tan^{-1} m + \frac{\pi}{2}).$$

Our attention now focuses on the angles β_1 and β_2 for which the set of x such that $x'_\beta(x) = 0$ is non-empty.

Definition: Let Q_1, Q_2 be the sets such that

$$Q_1 = \{x \mid x'_{\beta_1}(x) = 0, x\varepsilon[0,1]\}$$

$$Q_2 = \{x \mid x'_{\beta_2}(x) = 0, x\varepsilon[0,1]\}$$

Proposition 6: If Q_i contains a convex interval (more than one point) then β_i is not contained in P.

Proof: Let $[x_1, x_2]$ be contained in Q_i, then, by (53) $x_\beta(x)$ is not monotone.

Proposition 7: if Q_i contains a single point then β_i is contained in P.

Proof: Let x_o be the single point and consider the intervals $[0,x_o]$ and $[x_o,1]$. Let x and y be two points in [0,1], $x<y$. If x and y both in first or second interval then $x_\beta(x)$

If p is even, the methods of theorem 9 directly apply and L_p error is direction dependent. If p is odd, we can use the argument of theorem 3 for the permanence of sign in subintervals of [0,1]. Hence, the L_p error is plus or minus the value of the integrals in each subinterval, all of which are direction dependent. Thus we have proven the following

Theorem 11: The L_p error (56), $2 \leq p < \infty$, is direction dependent.

5.5 BEST APPROXIMATIONS FOR β BELONGING TO P

For β belonging to P, by the remarks of 5.2, the analysis of 5.1 is valid and the conditions for a minimum in c and d are

$$\int_0^1 y_\beta(x)x_\beta(x)x_\beta'(x)dx = c\int_0^1 x^2{}_\beta(x)x_\beta'(x)dx + d\int_0^1 x_\beta(x)x_\beta'(x)dx$$

$$\int_0^1 y_\beta(x)x_\beta'(x)dx = c\int_0^1 x_\beta(x)x_\beta'(x)dx + d\int_0^1 x_\beta'(x)dx \quad (57)$$

If we define

$$\overline{x}_\beta = \frac{\int_0^1 x_\beta(x)x_\beta' dx}{\int_0^1 x'_\beta(x)dx} \qquad \overline{y}_\beta = \frac{\int_0^1 y_\beta(x)x_\beta'(x)dx}{\int_0^1 x_\beta'(x)dx} \quad (58)$$

and let

$$u_\beta(x) = x_\beta(x) - \overline{x}_\beta \qquad v_\beta(x) = y_\beta(x) - \overline{y}_\beta \quad (59)$$

then

$$\int_0^1 u_\beta(x)x_\beta'(x)dx = 0 \qquad \int_0^1 v_\beta(x)x_\beta'(x)dx = 0 \quad (60)$$

or which, in terms of the normalized variables is the condition

$$\int_0^1 x_\beta(x)x_\beta'(x)dx = 0 \qquad \int_0^1 y_\beta(x)x_\beta'(x)dx = 0 \quad (61)$$

This assumption (61) simplifies the conditions for a minimum (57) and the solution for a minimum with the normalized variable assumption is case (58) is easily seen since

$$\bar{x}_\beta = \frac{\int_0^1 x_\beta(x)x'_\beta(x)\,dx}{\int_0^1 x'_\beta(x)\,dx} = 1/2x_\beta(x)\Big|_0^1 = 1/2(\cos\beta+(f(1)-f(0))\sin\beta) \qquad (68)$$

which is directly dependent on β .

5.7 PROPERTIES OF C(β)

Since f(x) is continuous on [0,1], f(x) is bounded and therefore for all real c and d, the error function (5) and therefore C(c,d,β) is bounded. Hence C(β) is bounded when c* and d* are real.

In the case of d*, there is nothing to prove, since it is always zero. The numerator of c* is bounded, therefore c* is infinite only when $\sigma^2_{x_\beta}$ is zero. This is equivalent to determining when

$$x^3_\beta(x)\Big|_0^1 = (\cos\beta+f(1)\sin\beta)^3 - (f(0)\sin\beta)^3 = 0 \qquad (69)$$

Condition (69) is in the form

$$u^3-v^3 = (u-v)(u^2+uv+v^2)=0 \qquad (70)$$

The second factor has only imaginary, non-trivial, solutions. Thus the only real solutions of interest of (69) are

$$x_\beta(1)-x_\beta(0)=\cos\beta+f(1)\sin\beta-f(0)\sin\beta= 0 \qquad (71)$$

or

$$\tan\beta= 1/(f(0)-f(1)) \qquad (72)$$

If β' is a solution of (71) or (72), thenβ' does not belong to P by definition of P. If f(x) is a straight line, then P is open and β' is an end point of the P interval. If f(x) is not a straight line, the range of $x_\beta(x)$ is a compact interval, not a point. This means it must achieve

We turn our attention to finding the solutions of $C'(\beta)=0$, since any relative mins in P must satisfy this condition. If we let N and D denote the numerator and denominator of (75), then

$$C'(\beta)=0 \quad \text{implies } DN'-ND'=0 \tag{76}$$

If we define

$$\delta_x^2={}_0\int^1 x^2 f'(x)dx$$
$$\delta_y^2={}_0\int^1 f(x)\ f'(x)dx \tag{76}$$
$$\delta_{xy}={}_0\int^1 xf(x)f'(x)dx$$

then we can write N, using (48), as

$$N=(\sigma_x^2\sigma_y^2-K_{xy}^2)\cos^2\beta+ (\delta_x^2\delta_y^2-\delta_{xy}^2)\sin^2\beta \tag{77}$$
$$(\sigma_x^2\delta_y^2+\sigma_y^2\delta_x^2 - 2K_{xy}\delta_{xy})\sin\beta\cos\beta$$

and D as

$$3D=\cos^3\beta+ 3y_1\cos^2\beta\sin\beta+ 3y_1^2 \cos\beta\sin^2\beta+ (y_1^3-y_0^3)\sin^3\beta \tag{78}$$

N' is a second degree equation in cos and sin, D' is a third degree equation in cos and sin, therefore the derivative will be a fifth degree equation in cos and sin. By dividing throughout by $\cos^5$, $C'(\beta)=0$ leads to a fifth degree equation in tan. The function tan is periodic with period π, but since $C(\beta)$ is periodic with period π except for sign this means that our fifth degree equation in tan will lead to all critical points which characterize the behavior of $C(\beta)$. Thus we will have five (at most) solutions in an interval of length π or ten solutions from $[0,2\pi]$.

$$\overline{x} = \frac{\int_{s_0}^{s_1} x\,ds}{s_1 - s_0} \qquad \overline{y} = \frac{\int_{s_0}^{s_1} y\,ds}{s_1 - s_0} \quad (85)$$

The two solutions of (83) correspond to the max and min of moments of inertia about all axes. Since theorem 4 applies, the solution (83) is equivalent to minimizing the error

$$G(c,d,\beta) = \int_{s_0}^{s_1} e_\beta(x)\,ds \quad (86)$$

The fundamental difference between the minimization of the continuous least square error for directions β and (86) is that the arc length is direction independent whereas the interval of integration $d(x_\beta(x))$ is direction dependent. Alternately, we can view (86) as the continuous least square error for a constant interval of integration.

5.10 EXAMPLES

For the class of functions, $f(x)=x^n$, some simplifications are immediate; i.e., $y_1 = 1$, $y_0=0$.

Thus (80) reduces to

$$(b-c/3)\tan^5_\beta + [(5b-2a)/3]\tan^4\beta + (b-a+c/3)\tan^3\beta + (b-a-c/3)$$
$$\tan^2\beta[(2b-5a)/3]\tan\beta + (c/3-a)=0$$

Further we can evaluate the quantities

$$\sigma x^2 = 1/3 \quad \sigma_y^2 = 1/((2n+1) \quad K_{xy} = 1/(n+2) \quad \delta_x^2 = n/(n+2)$$
$$\delta_y^2 = 1/3 \quad \delta_{xy} = n/(2n+1)$$

from which we derive the values of

$$a=(n-1)^2/[3(n+2)^2(2n+1)] \qquad b=n(n-1)^2/[3(n+2)(2n+1)^2]$$
$$c=2(n-1)^2/[9(n+2)(2n+1)]$$

Using these results we can put (80) in the following form

$$(5n-2)(n+2)\tan^5\beta + 3(5n^2+6n-2)\tan^4\beta + (13n^2+10n-5)\tan^3$$
$$(5n^2-10n-13)\tan^2\beta + 3(2n^2-6n-5)\tan\beta + (2n+1)(2n-5)=0$$

BIBLIOGRAPHY

Apostol, T.M. [1957]: Mathematical Analysis. Reading, Mass: Addison-Wesley.

Cheney, E.W. [1966]: Approximation Theory. New York: McGraw-Hill.

Courant, R. and Fritz, J. [1965]: Calculus and Analysis. New York: Interscience.

Davis, P.J. [1963]: Interpolation and Approximation. New York: Blaisdell.

Handscomb, D.C. [1966]: Methods of Numerical Approximation. Oxford: Pergamon.

Jones, H.E. [1937]: "Some Geometrical Considerations in the General Theory of Fitting Lines and Planes", Metron, Vol. 13, No. 1, pp. 21-30.

Linnik, Y.U. [1961]: Method of Least Squares and Principles of the Theory of Observations. Trans. by Elandt, R.C., New York: Pergamon.

Rice, J.R. [1964]: The Approximation of Functions. Reading, Mass: Addison-Wesley.

Roos, C.F., [1937]: "A General Invariant Criterion of Fit for Lines and Planes where all Variates are Subject to Error", Metron, Vol. 13, No. 1, pp. 1-20.

Taylor, A.E. [1955]: Advanced Calculus. Boston: Ginn.

4) L_2 approximation (continuous least squares) is direction dependent. A direction of measurement β, is equivalent to a rotation of the coordinate system by an angle β, the counterclockwise direction taken as positive. If the function f is still a function in the rotated frame of reference, then the L_2 error is well defined. If P is the set of angles β where the L_2 error is well defined and positive, and if f has no straight line segments, then P is closed. The solution for β^* leads to a fifth degree equation in tan β. When P is closed β^* exists and is one of the solutions of the fifth degree equation or is one of the end points of P.