

## A Review Article On Mathematical Aspects Of Nonlinear Models

<sup>1</sup>B. Mahaboob, <sup>2</sup>P. Venkateswararao, <sup>3</sup>P. S. Prema Kumar, <sup>4</sup>S. Venu Madhava Sarma, <sup>5</sup>S. Raji Reddy, <sup>6</sup>Y. Hari Krishna,

<sup>1,4</sup>Department of Mathematics, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, AP, India

<sup>2</sup>KLU Business School, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, AP, India

<sup>3</sup>Department of Mechanical Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, AP, India

<sup>5</sup>Department of M&H, MGIT, Gandipet, Hyderabad, Telangana-76

<sup>6</sup>Department of Mathematics, ANURAG Engineering College, Ananthagiri(v), Kodad, Suryapet, Telangana-508206

**Article History:** Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

### Abstract

The main objective of this review article is to propose some mathematical aspects of nonlinear models. In mathematics, nonlinear modelling is empirical or semi-empirical modelling which takes at least some nonlinearities into account. Nonlinear modelling in practice therefore means modelling of phenomena in which independent variables affecting the system can show complex and synergetic nonlinear effects. Contrary to traditional modelling methods, such as linear regression and basic statistical methods, nonlinear modelling can be utilized efficiently in a vast number of situations where traditional modelling is impractical or impossible. This review article mainly explores on mathematical preliminaries of nonlinear models, solution of algebraic and transcendental equations and solution of system of nonlinear equations. In addition to these Taylor polynomial and finite difference operators, least - squares polynomial approximation and the roots of the equations are also discussed here.

**Keywords:** Nonlinear Model, Algebraic Equation, Mathematical Model, Transcendental Equation, Difference Operator

### 1. INTRODUCTION

Mathematical models involving nonlinearity are becoming increasingly popular in recent years. In practice nonlinear mathematical models have a wide variety of applications. A mathematical model is said to be nonlinear model, if the derivatives of the model with respect to the model parameters depend on one or more parameters. This aspect is essential to distinguish nonlinear from curve linear model. Generally, the direct interpretation of the process under study can be described by the model parameters.

Several theorems in mathematical analysis are stated as preliminaries for nonlinear models. Certain numerical analysis techniques to obtain solutions of algebraic and transcendental equations have been discussed to use in iterative estimation procedures for estimating parameters of nonlinear models.

### 2. MATHEMATICAL PRELIMINARIES FOR NONLINEAR MODELS:

The following mathematical preliminaries will be useful in the present study:

#### Intermediate Value Theorems:

- (i) If  $f(x)$  is continuous in  $a \leq x \leq b$  and if  $f(a)$  and  $f(b)$  are of opposite signs then  $f(\xi) = 0$  for atleast one number  $\xi$  such that  $a < \xi < b$
- (ii) Let  $f(x)$  be continuous in  $[a, b]$  and  $k$  be any number between  $f(a)$  and  $f(b)$ . Then there exists atleast one number  $\xi$  in  $(a, b)$  such that  $f(\xi) = k$ .

#### Rolle's Theorem:

- If
- (i)  $f(x)$  is continuous in  $a \leq x \leq b$
  - (ii)  $f'(x)$  exists in  $a < x < b$  and
  - (iii)  $f(a) = f(b)$

then there exists atleast on value of  $x$ , say  $\xi$  such that  $f'(\xi) = 0$ ,  $a < \xi < b$

#### Generalized Rolle's Theorem:

Let  $f(x)$  be a function which is  $n$  times differentiable on  $[a, b]$ . If  $f(x)$  vanishes at  $(n+1)$  distinct points  $x_0, x_1, \dots, x_n$  in  $(a, b)$  then there exists a number  $\xi$  in  $(a, b)$  such that  $f^{(n)}(\xi) = 0$

#### Mean-value theorem for derivatives:

If (i)  $f(x)$  is continuous in  $[a, b]$

and (ii)  $f'(x)$  exists in  $(a, b)$

then there exists atleast on value of  $x$  say  $\xi$  between  $a$  and  $b$  such that

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}, \quad a < \xi < b$$

By assuming  $b = a + h$  one can find

$$f(a + h) = f(a) + hf'(a + \theta h), \quad 0 < \theta < 1$$

**Taylor's series for a function of one variable:**

If  $f(x)$  is continuous and possesses continuous derivatives of order  $n$  in an interval that includes  $x = a$  then in that interval

$$f(x) = f(a) + (x - a)f'(a) + \frac{(x - a)^2}{2} f''(a) + \dots + \frac{(x - a)^{n-1}}{(n-1)!} f^{(n-1)}(a) + R_n(x)$$

where  $R_n(x)$ , the remainder term can be expressed in the form

$$R_n(x) = \frac{(x - a)^n}{n!} f^{(n)}(\xi), \quad a < \xi < x$$

**Maclaurin's Expansion:**

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2} f''(0) + \dots + \frac{x^n}{n!} f^{(n)}(0) + \dots \dots \dots (1)$$

**Taylor's series for a function of several variables:**

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2) = f(x_1, x_2) + \frac{\partial f}{\partial x_1} \Delta x_1 + \frac{\partial f}{\partial x_2} \Delta x_2 + \dots \dots \dots (2)$$

$$\frac{1}{2} \left[ \frac{\partial^2 f}{\partial x_1^2} (\Delta x_1)^2 + 2 \frac{\partial^2 f}{\partial x_1 \partial x_2} \Delta x_1 \Delta x_2 + \frac{\partial^2 f}{\partial x_2^2} (\Delta x_2)^2 \right] + \dots$$

**Taylor's series for a function of several variables:**

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n) = f(x_1, x_2, \dots, x_n)$$

$$+ \frac{\partial f}{\partial x_1} \Delta x_1 + \frac{\partial f}{\partial x_2} \Delta x_2 + \dots + \frac{\partial f}{\partial x_n} \Delta x_n +$$

$$\frac{1}{2} \left[ \frac{\partial^2 f}{\partial x_1^2} (\Delta x_1)^2 + \dots + \frac{\partial^2 f}{\partial x_n^2} (\Delta x_n)^2 + 2 \frac{\partial^2 f}{\partial x_1 \partial x_2} \Delta x_1 \Delta x_2 + \dots + 2 \frac{\partial^2 f}{\partial x_{n-1} \partial x_n} \Delta x_{n-1} \Delta x_n \right] + \dots \dots \dots (3)$$

**3. SOLUTION OF ALGEBRAIC AND TRANSCENDENTAL EQUATIONS:**

In research work, a frequently occurring problem is to find the roots of the equation of the form

$$f(x) = 0 \dots \dots \dots (4)$$

If  $f(x)$  is quadratic, cubic or a biquadratic expression then algebraic formulae are available for expressing the roots in terms of the coefficients. On the other hand when  $f(x)$  is a polynomial of higher degree or an expression involving transcendental functions, algebraic methods are not available and approximate methods are to be performed.

Here some numerical methods for the solution of (4) will be discussed where  $f(x)$  is algebraic or transcendental or a combination of both.

**Bisection Method:**

This method is based on Intermediate Value Theorem which states that if a function  $f(x)$  is continuous in  $[a, b]$  and  $f(a)f(b)<0$  then there exists atleast one root for  $f(x)=0$  between  $a$  and  $b$ . For definiteness let  $f(a)$  be positive and  $f(b)$  be negative. Then the root lies between  $a$  and  $b$  and let its approximate value be given by  $x_0 = \frac{a+b}{2}$ . If  $f(x_0) = 0$  then obviously  $x_0$  is a root of (3.3.1). If not then the root lies either between  $x_0$  and  $b$  or between  $x_0$  and  $a$  depending on the sign of  $f(x_0)$ . Let this new interval be  $[a_1, b_1]$  whose length is  $\frac{1}{2}(b-a)$ . As above this is bisected at  $x_1$  and the new interval possesses length  $\frac{1}{4}(b-a)$ . This process is repeated until the latest interval which contains the root is as small as desired say  $\epsilon$ . At each step the length of the interval is reduced by a factor of one-half and at the end of the  $n^{\text{th}}$  step the new interval will be  $[a_n, b_n]$  of length  $\frac{1}{2^n}(b-a)$ . Now

$$\begin{aligned} \frac{b-a}{2^n} &\leq \epsilon \\ \text{Multiplying by } \frac{2^n}{\epsilon} & \\ \frac{b-a}{\epsilon} &\leq 2^n \\ \log \frac{b-a}{\epsilon} & \\ \frac{\log \frac{b-a}{\epsilon}}{\log 2} &\leq n \dots\dots\dots(5) \end{aligned}$$

This inequation gives the number of iterations required to get an accuracy  $\epsilon$ .

**The Iteration Method:**

This method solves  $x = F(x)$  by the recursion  $x_n = F(x_{n-1})$  and converges to a root if  $|F'(x)| \leq L < 1$ . The error  $e_n = r - x_n$  where  $r$  is the exact root has the property

$$e_n \approx F'(r)e_{n-1} \dots\dots\dots(6)$$

so that each iteration reduces the error by a factor near  $F'(r)$ . If  $F'(r)$  is near 1 this is slow convergence.

**Regula Falsi Method:**

Let two points  $x_0$  and  $x_1$  be such that  $f(x_0)$  and  $f(x_1)$  are of opposite signs and  $f$  is continuous on  $[x_0, x_1]$ . Since the graph of  $y = f(x)$  crosses the  $x$ -axis between these two points, a root must lie between these points. Now the equation of the chord joining the two points  $(x_0, f(x_0))$  and  $(x_1, f(x_1))$  is

$$(x - x_0)(f(x_0) - f(x_1)) = (y - f(x_0))(x_0 - x_1) \dots\dots\dots(7)$$

The method consists in replacing the part of the curve between the points  $(x_0, f(x_0))$  and  $(x_1, f(x_1))$  by means of the chord joining these points and taking the point of intersection of the chord with the  $x$ -axis as an approximation to the root. The point of intersection in the present case is given by putting  $y=0$  in (3.3.4).

Hence the second approximation to the root of  $f(x)=0$  is given by

$$x_2 = x_0 - \frac{f(x_0)}{f(x_1) - f(x_0)}(x_1 - x_0) \dots\dots\dots(8)$$

If now  $f(x_2)$  and  $f(x_0)$  are of opposite signs then the root lies between  $x_0$  and  $x_2$  and replace  $x_1$  by  $x_2$  in (3.3.5) to get the next approximation. Otherwise replace  $x_0$  by  $x_2$  and generate the next approximation. The procedure is repeated till the root is obtained to the desired accuracy.

**Newton - Raphson Method:**

Let  $x_0$  be an approximation to the desired root of  $f(x) = 0$  and let  $x_1 = x_0 + h$  be the correct root so that  $f(x_1) = 0$ . Expanding  $f(x_0 + h)$  by Taylor's series one can get

$$f(x_0) + hf'(x_0) + \frac{h^2}{2} f''(x_0) + \dots = 0 \dots\dots\dots(9)$$

Neglecting the derivatives from second order onwards, one may find

$$f(x_0) + hf'(x_0) = 0 \dots\dots\dots(10)$$

This gives  $h = -\frac{f(x_0)}{f'(x_0)} \dots\dots\dots(11)$

A better approximation than  $x_0$  is therefore given by  $x_1$  where

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \dots\dots\dots(12)$$

Successive approximations are given by  $x_2, x_3, \dots, x_{n+1}$

where  $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (n=0,1,2,\dots) \dots\dots\dots(13)$

This is known as Newton – Raphson formula.

If  $f'(x)$  is complicated the previous iterative method may be preferable, but Newton's method converges such more rapidly and usually gets the nod. The error  $e_n$  here satisfies

$$e_n \approx \frac{-f''(r)}{2f'(r)} e_{n-1}^2 \dots\dots\dots(14)$$

This is known as quadratic convergence, each error roughly proportional to the square of the previous error. The number of correct digits almost doubles with each iteration. The square root iteration

$$x_n = \frac{1}{2} \left( x_{n-1} + \frac{Q}{x_{n-1}} \right) \dots\dots\dots(15)$$

is a special case of Newton's method corresponding to  $f(x) = x^2 - Q$ . It converges quadratically to the positive square root of  $Q$ , for  $Q > 0$ .

The more general root finding formula

$$x_n = x_{n-1} - \frac{x_{n-1}^p - Q}{p x_{n-1}^{p-1}} \dots\dots\dots(16)$$

is a special case of Newton's method. It produces a  $p^{\text{th}}$  root of  $Q$ .

**Generalized Newton's Method:**

If  $\xi$  is a multiple root of order  $p$  of  $f(x) = 0$  then the iteration formula corresponding to (3.3.10) is written as

$$x_{n+1} = x_n - p \frac{f(x_n)}{f'(x_n)} \dots\dots\dots(17)$$

This means that  $\frac{1}{p} f'(x_n)$  is the slope of the straight line passing through  $(x_n, y_n)$  and intersecting the x-axis at the point  $(x_{n+1}, 0)$ .

**Muller’s Method:**

In this method  $f(x)$  is approximated by a second degree curve in the nbd of a root. The roots of the quadratic are then assumed to be the approximations to the roots of the equation  $f(x) = 0$ . The method is iterative, converges almost quadratically and can be used to obtain non real complex roots

Let  $x_{i-2}, x_{i-1}, x_i$  be three different approximations to a root of  $f(x)=0$  and let  $y_{i-2}, y_{i-1}$ , and  $y_i$  be the corresponding values of  $y=f(x)$ .

Let  $P(x) = A(x - x_i)^2 + B(x - x_i) + y_i$  .....(18)

be the parabola passing through the points  $(x_{i-2}, y_{i-2}), (x_{i-1}, y_{i-1})$  and  $(x_i, y_i)$ . Then

$y_{i-1} = A(x_{i-1} - x_i)^2 + B(x_{i-1} - x_i) + y_i$  .....(19)

and  $y_{i-2} = A(x_{i-2} - x_i)^2 + B(x_{i-2} - x_i) + y_i$  .....(20)

From equations (19) and (20), one may get

$y_{i-1} - y_i = A(x_{i-1} - x_i)^2 + B(x_{i-1} - x_i)$  .....(21)

$y_{i-2} - y_i = A(x_{i-2} - x_i)^2 + B(x_{i-2} - x_i)$  .....(22)

Solutions of equations (21) and (22) give

$A = (x_{i-2} - x_i)(y_{i-1} - y_i) - (x_{i-1} - x_i)(y_{i-2} - y_i)$  .....(23)

$B = \frac{(x_{i-2} - x_i)^2 (y_{i-1} - y_i) - (x_{i-1} - x_i)^2 (y_{i-2} - y_i)}{(x_{i-2} - x_{i-1})(x_{i-1} - x_i)(x_{i-2} - x_i)}$  .....(24)

With the above values of A and B the quadratic equation  $A(x - x_i)^2 + B(x - x_i) + y_i = 0$  gives the next approximation

$x_{i+1} = x_i + \frac{-B \pm \sqrt{B^2 - 4Ay_i}}{2A}$  .....(25)

**Ramanujan’s Method:**

Srinivasa Ramanujan described an iterative method which can be used to compute the smallest root of the equation

$f(x) = 0$  .....(26)

where  $f(x) = 1 - (a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \dots)$  .....(27)

For smaller values of x, one can write

$[1 - (a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \dots)]^{-1} = b_1 + b_2x + b_3x^2 + \dots$  .....(28)

Expanding LHS by binomial theorem for negative integer gives

$1 + (a_1x + a_2x^2 + a_3x^3 + \dots) + (a_1x + a_2x^2 + a_3x^3 + \dots)^2 + \dots = b_1 + b_2x + b_3x^2 + \dots$  .(29)

Comparing the coefficients of like powers of x, one can get

$b_1 = 1$   
 $b_2 = a_1 = a_1(1) = a_1b_1$   
 $b_3 = a_1^2 + a_2 = a_1a_1 + a_2(1) = a_1b_2 + a_2b_1$   
 .....  
 $b_n = a_1b_{n-1} + a_2b_{n-2} + \dots + a_{n-1}b_1$   
 $n = 2, 3, \dots$

Ramanujan states that the successive convergents namely  $\frac{b_n}{b_{n+1}}$  approach a root of the equation

$f(x) = 0$ .

**Graeffe’s Root-Squaring method:**

Let  $P_n(x)$  be a polynomial of degree n. Graeffe’s method consists in transforming  $P_n(x)$  into another polynomial say  $Q_n(z)$  of the same degree but whose roots are the squares of the roots of the original polynomial. The process is repeated so that the roots of the new polynomial are distributed more sparsely. This is possible provided that the roots of the original polynomial are all real and distinct. The roots are finally computed directly from the coefficients.

**Lin-Bairstow’s method:**

This is often used in estimating the factors of polynomials. Let the polynomial be given by

$$f(x) = A_3x^3 + A_2x^2 + A_1x + A_0 = 0 \dots\dots\dots(31)$$

Let  $x^2 + Rx + S$  be a quadratic factor of  $f(x)$  and let an approximate factor be  $x^2 + rx + s$ . Usually first approximations to r and s can be obtained from the last three terms of the given polynomial. Thus

$$r = \frac{A_1}{A_2} \quad \text{and} \quad s = \frac{A_0}{A_2} \dots\dots\dots(32)$$

Let

$$\begin{aligned} f(x) &= (x^2 + rx + s)(B_2x + B_1) + Cx + D \\ &= B_2x^3 + (B_2r + B_1)x^2 + (C + B_1r + sB_2)x + (B_1s + D) \dots\dots\dots(33) \end{aligned}$$

where the constants  $B_1, B_2, C$  and  $D$  are to be determined. Equating the coefficients of like powers of  $x$  in equations (3.3.28), (3.3.30) one may get

$$\left. \begin{aligned} B_2 &= A_3 \\ B_1 &= A_2 - rB_2 \\ C &= A_1 - rB_1 - sB_2 \\ D &= A_0 - sB_1 \end{aligned} \right\} \dots\dots\dots(34)$$

From (3.3.31) it is clear that the coefficients  $B_1, B_2$  of the factored polynomial and also the coefficients  $C$  and  $D$  are functions of  $r$  and  $s$ . Since  $x^2 + Rx + S$  is a factor of the given polynomial it follows that

$$\left. \begin{aligned} C(R, S) &= 0 \\ \text{and} \quad D(R, S) &= 0 \end{aligned} \right\} \dots\dots\dots(35)$$

Letting  $R = r + \Delta r$  and  $S = s + \Delta s$  equations (3.3.32) can be expanded by Taylor’s series

$$\left. \begin{aligned} C(R, S) &= C(r, s) + \Delta r \cdot \frac{\partial C}{\partial r} + \Delta s \cdot \frac{\partial C}{\partial s} = 0 \\ D(R, S) &= D(r, s) + \Delta r \cdot \frac{\partial D}{\partial r} + \Delta s \cdot \frac{\partial D}{\partial s} = 0 \end{aligned} \right\} \dots\dots\dots(37)$$

where the derivatives are to be calculated at  $r$  and  $s$ . Equations (3.3.34) can then be solved for  $\Delta r$  and  $\Delta s$ . Using of these values in (3.3.33) will give the next approximation to  $R$  and  $S$ . The process can be repeated until successive values of  $R$  and  $S$  show no significant change.

**Quotient – Difference method:**

Let the given cubic equation be

$$f(x) = a_0x^3 + a_1x^2 + a_2x + a_3 = 0 \dots\dots\dots(38)$$

and let  $x_1, x_2$  and  $x_3$  be its roots such that  $0 < |x_1| < |x_2| < |x_3|$

Then  $f(x)$  can be expressed as given below

$$\begin{aligned} (f(x))^{-1} &= \sum_{r=1}^3 \frac{p_r}{x - x_r} = \sum_{r=1}^3 \frac{p_r}{-x_r \left(1 - \frac{x}{x_r}\right)} = \sum_{r=1}^3 \frac{-p_r}{x_r} \left(1 - \frac{x}{x_r}\right)^{-1} \\ &= \sum_{r=1}^3 \frac{-p_r}{x_r} \left(1 + \frac{x}{x_r} + \frac{x^2}{x_r^2} + \dots\right) = \sum_{r=1}^{\infty} -p_r \left(\frac{1}{x_r} + \frac{x}{x_r^2} + \frac{x^2}{x_r^3} + \dots\right) \\ &= \sum_{i=0}^{\infty} \alpha_i x_i \dots\dots\dots(39) \end{aligned}$$

where  $\alpha_i = -\sum_{r=1}^3 \frac{p_r}{x_r^{i+1}} \dots\dots\dots(40)$

Quotients and differences are defined by the relations

$$q_1^{(i)} = \frac{\alpha_i}{\alpha_{i-1}} \dots\dots\dots(41)$$

and  $\Delta_1^{(i)} = q_1^{(i+1)} - q_1^{(i)} \dots\dots\dots(42)$

Substituting (3.3.37) in (3.3.38), one can get

$$\begin{aligned} q_1^{(i)} &= \frac{\frac{p_1}{x_1^{i+1}} + \frac{p_2}{x_2^{i+1}} + \frac{p_3}{x_3^{i+1}}}{\frac{p_1}{x_1^i} + \frac{p_2}{x_2^i} + \frac{p_3}{x_3^i}} \\ &= \frac{1}{x_1} \cdot \frac{p_1 + p_2 \left(\frac{x_1}{x_2}\right)^{i+1} + p_3 \left(\frac{x_1}{x_3}\right)^{i+1}}{p_1 + p_2 \left(\frac{x_1}{x_2}\right)^i + p_3 \left(\frac{x_1}{x_3}\right)^i} \end{aligned}$$

Since  $\left|\frac{x_1}{x_2}\right| < 1$  and  $\left|\frac{x_1}{x_3}\right| < 1$ , one can have

$$\lim_{i \rightarrow \infty} q_1^{(i)} = \frac{1}{x_1} \dots\dots\dots(43)$$

Thus the quotients  $q_1^{(i)}$  approach, in the limit, the reciprocal of a root of the given cubic equation (38).

In a similar way it can be shown that

$$\lim_{i \rightarrow \infty} \frac{\frac{1}{x_1} - q_1^{(i)}}{\left(\frac{x_1}{x_2}\right)^i} = \frac{1}{x_1} \frac{p_2}{p_1} \left(1 - \frac{x_1}{x_2}\right) \dots\dots\dots(44)$$

and  $\lim_{i \rightarrow \infty} \frac{\frac{1}{x_1} - q_1^{(i+1)}}{\left(\frac{x_1}{x_2}\right)^i} = \frac{1}{x_2} \frac{p_2}{p_1} \left(1 - \frac{x_1}{x_2}\right) \dots\dots\dots(45)$

Subtracting (45) from (44) and using (42) one can get

$$\lim_{i \rightarrow \infty} \frac{\Delta_1^{(i)}}{\left(\frac{x_1}{x_2}\right)^i} = \left(\frac{1}{x_1} - \frac{1}{x_2}\right) \frac{p_2}{p_1} \left(1 - \frac{x_1}{x_2}\right) \dots\dots\dots(46)$$

Similarly

$$\lim_{i \rightarrow \infty} \frac{\Delta_1^{(i+1)}}{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^{i+1}} = \left( \frac{1}{x_1} - \frac{1}{x_2} \right) \frac{p_2}{p_1} \left( 1 - \frac{x_1}{x_2} \right) \dots\dots\dots(47)$$

Form (46) and (47) one can get

$$\frac{\Delta_1^{(i)}}{\Delta_1^{(i+1)}} = \frac{x_2}{x_1} = \frac{1}{\frac{x_1}{x_2}} \dots\dots\dots(48)$$

Put  $\frac{1}{x_2} = \lim_{i \rightarrow \infty} q_2^{(i)} \dots\dots\dots(49)$

and using (43), eq(48) gives

$$\Delta_1^{(i)} \cdot q_2^{(i)} = \Delta_1^{(i+1)} \cdot q_1^{(i+1)} \dots\dots\dots(50)$$

Following (42) define

$$\Delta_2^{(i)} - \Delta_1^{(i+1)} = q_2^{(i+1)} - q_2^{(i)} \dots\dots\dots(51)$$

In the similar way with (48) it can be shown that

$$\lim_{i \rightarrow \infty} \frac{\Delta_2^{(i)}}{\Delta_2^{(i+1)}} = \frac{1}{\frac{x_2}{x_3}} \dots\dots\dots(52)$$

The relations (50), (51) can be easily generalized as

$$\Delta_k^{(i)} \cdot q_{k+1}^{(i)} = \Delta_k^{(i+1)} \cdot q_k^{(i+1)} \dots\dots\dots(53)$$

and  $\Delta_k^{(i)} + q_k^{(i)} = \Delta_{k-1}^{(i+1)} + q_k^{(i+1)} \dots\dots\dots(54)$

where  $\Delta_0^{i+1} = 0 \dots\dots\dots(55)$

**4. SOLUTION OF SYSTEMS OF NONLINEAR EQUATIONS:**

**(i) The method of iteration:**

Let the equations be given by

$$\left. \begin{aligned} f(x, y) &= 0 \\ g(x, y) &= 0 \end{aligned} \right\} \dots\dots\dots(56)$$

whose real roots are required within a specified accuracy. Put the equations in the form

$$x = F(x, y) \dots\dots\dots(57)$$

and  $y = G(x, y)$

where the functions F and G satisfy the equations

$$\left. \begin{aligned} \left| \frac{\partial F}{\partial x} \right| + \left| \frac{\partial F}{\partial y} \right| &< 1 \\ \text{and } \left| \frac{\partial G}{\partial x} \right| + \left| \frac{\partial G}{\partial y} \right| &< 1 \end{aligned} \right\} \dots\dots\dots(58)$$

in the neighbourhood of the root.

Let  $(x_0, y_0)$  be the initial approximation to a root  $(\alpha, \beta)$  of the system (56). Then construct the successive approximations according to the formulae



$$\left. \begin{aligned} x_1 &= F(x_0, y_0), & y_1 &= G(x_0, y_0) \\ x_2 &= F(x_1, y_1), & y_2 &= G(x_1, y_1) \\ x_3 &= F(x_2, y_2), & y_3 &= G(x_2, y_2) \\ &\dots\dots\dots \\ &\dots\dots\dots \\ x_{n+1} &= F(x_n, y_n), & y_{n+1} &= G(x_n, y_n) \end{aligned} \right\} \dots\dots\dots(59)$$

For faster convergence recently computed values of  $x_i$  may be used in the evaluation of  $y_i$  in (59). If the iteration process (59) converges then one can get

$$\left. \begin{aligned} \alpha &= F(\alpha, \beta) \\ \text{and } \beta &= G(\alpha, \beta) \end{aligned} \right\} \dots\dots\dots(60)$$

in the limit. Thus  $\alpha$  and  $\beta$  are the roots of the system (57)

**Theorem:**

Let  $x = \alpha$  and  $y = \beta$  be one pair of roots of the system  $x = F(x, y)$  and  $y = G(x, y)$  in the closed neighbourhood R.

If (a) the functions F and G and their first partial derivatives are continuous in R

$$(b) \left. \begin{aligned} \left| \frac{\partial F}{\partial x} \right| + \left| \frac{\partial F}{\partial y} \right| < 1 \\ \text{and } \left| \frac{\partial G}{\partial x} \right| + \left| \frac{\partial G}{\partial y} \right| < 1 \end{aligned} \right\} \dots\dots\dots(61)$$

for all x,y in R and

(c) the initial approximation  $(x_0, y_0)$  is chosen in R then the sequence of approximations given by (59) converges to the roots  $x = \alpha$  and  $y = \beta$  of the system (56).

**Newton-Raphson method:**

Let  $(x_0, y_0)$  be an initial approximation to the root of the system

$$\left. \begin{aligned} f(x, y) &= 0 \\ g(x, y) &= 0 \end{aligned} \right\} \dots\dots\dots(62)$$

If  $(x_0 + h, y_0 + k)$  is the root of the system then

$$\left. \begin{aligned} f(x_0 + h, y_0 + k) &= 0 \\ g(x_0 + h, y_0 + k) &= 0 \end{aligned} \right\} \dots\dots\dots(63)$$

Let f and g be sufficiently differentiable. Expanding (63) by Taylor's series gives

$$\left. \begin{aligned} f_0 + h \frac{\partial f}{\partial x_0} + k \frac{\partial f}{\partial y_0} + \dots = 0 \\ \text{and } g_0 + h \frac{\partial g}{\partial x_0} + k \frac{\partial g}{\partial y_0} + \dots = 0 \end{aligned} \right\} \dots\dots\dots(64)$$

where  $\frac{\partial f}{\partial x_0} = \left( \frac{\partial f}{\partial x} \right)_{x=x_0}$ ,  $f_0 = f(x_0, y_0)$  etc.

Neglecting the second and higher order terms one can obtain the system of linear equations

$$\left. \begin{aligned} h \frac{\partial f}{\partial x_0} + k \frac{\partial f}{\partial y_0} &= -f_0 \\ \text{and } h \frac{\partial g}{\partial x_0} + k \frac{\partial g}{\partial y_0} &= -g_0 \end{aligned} \right\} \dots\dots\dots(65)$$

If the Jacobian

$$J(f, g) = \begin{vmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{vmatrix} \dots\dots\dots(66)$$

does not vanish then the linear equations (65) possess a unique solution given by

$$\left. \begin{aligned} h &= \frac{\begin{vmatrix} -f & \frac{\partial f}{\partial y} \\ -g & \frac{\partial g}{\partial y} \end{vmatrix}}{J(f, g)} \\ \text{and } k &= \frac{\begin{vmatrix} \frac{\partial f}{\partial x} & -f \\ \frac{\partial g}{\partial x} & -g \end{vmatrix}}{J(f, g)} \end{aligned} \right\} \dots\dots\dots(67)$$

The new approximations are then given by

$$\left. \begin{aligned} x_1 &= x_0 + h \\ y_1 &= y_0 + k \end{aligned} \right\} \dots\dots\dots(68)$$

The process is to be repeated till one can obtain the roots to the desired accuracy. This method, if possible, possesses quadratic convergence. The following gives the conditions which are sufficient for convergence.

**Theorem:**

Let  $(x_0, y_0)$  be an approximation to a root  $(\xi, \eta)$  of the system (56) in the closed neighbourhood R containing  $(\xi, \eta)$ . If (a) f,g and all their first and second derivatives are continuous and bounded in R and (b)  $J(f, g) \neq 0$  in R then the sequence of approximations given by

$$\left. \begin{aligned} x_{i+1} &= x_i - \frac{\begin{vmatrix} f & g \\ \frac{\partial f}{\partial y} & \frac{\partial g}{\partial y} \end{vmatrix}}{J(f, g)} \\ \text{and } y_{i+1} &= y_i - \frac{\begin{vmatrix} g & f \\ \frac{\partial g}{\partial x} & \frac{\partial f}{\partial x} \end{vmatrix}}{J(f, g)} \end{aligned} \right\} \dots\dots\dots(69)$$

converge to the root  $(\xi, \eta)$  of the system (56)

**5. TAYLOR POLYNOMIAL AND FINITE DIFFERENCE OPERATORS:**

The Taylor polynomial is the ultimate in osculation. For a single argument  $x_0$  the values of the polynomial and its first  $n$  derivatives are required to match those of a given function  $y(x)$ . That is

$$p^{(i)}(x_0) = y^{(i)}(x_0) \text{ for } i = 0, 1, 2, \dots, n \dots\dots\dots(70)$$

The Taylor formula settles the existence issue directly by exhibiting such a polynomial in the form

$$p(x) = \sum_{i=0}^n \frac{y^{(i)}(x_0)}{i!} (x - x_0)^i \dots\dots\dots(71)$$

The error of the Taylor polynomial when viewed as an approximation to  $y(x)$  can be expressed by the integral formula

$$y(x) - p(x) = \frac{1}{n!} \int_{x_0}^x y^{(n+1)}(x_0) (x - x_0)^n dx_0 \dots\dots\dots(72)$$

Lagrange's error formula may be deduced by applying a mean value theorem to the integral formula. It is

$$y(x) - p(x) = \frac{y^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1} \dots\dots\dots(73)$$

and clearly resembles the error formulae of collocation and osculation.

If the derivatives of  $y(x)$  are bounded independently of  $n$  then either error formula serves to estimate the degree  $n$  required to reduce  $|y(x) - p(x)|$  below a prescribed tolerance over a given interval of arguments  $x$ .

Analytic functions have the property that for  $n$  tending to infinity the above error of approximation has limit zero for all arguments  $x$  in a given interval. Such functions are then represented by the Taylor series

$$y(x) = \sum_{i=0}^{\infty} \frac{y^{(i)}(x_0)}{i!} (x - x_0)^i \dots\dots\dots(74)$$

The binomial series is an especially important case of the Taylor series. For  $-1 < x < 1$  one can have

$$(1 + x)^p = \sum_{i=0}^{\infty} \binom{p}{i} x^i \dots\dots\dots(75)$$

**Differentiation operator D:**

The differentiation operator  $D$  is defined by

$$D = h \frac{d}{dx} \dots\dots\dots(76)$$

The exponential operator may then be defined by

$$e^{kD} = \sum_{i=0}^{\infty} \frac{k^i D^i}{i!} \dots\dots\dots(77)$$

and the Taylor series in operator form becomes

$$y(x_k) = e^{kD} y_0(x_0) \dots\dots\dots(78)$$

The relationship between  $D$  and  $\Delta$  may be expressed in either of the forms

$$\Delta + 1 = e^D \qquad D = \Delta - \frac{1}{2}\Delta^2 + \frac{1}{3}\Delta^3 \dots\dots\dots(79)$$

both of which involve "infinite series" operator.

The Euler transformation is another useful relationship between two infinite series operators. It may be written as

$$(1 + E)^{-1} = \frac{1}{2} \left( 1 - \frac{1}{2}\Delta + \frac{1}{4}\Delta^2 - \frac{1}{8}\Delta^3 + \dots \right) \dots\dots\dots(80)$$

by using the binomial series.

The Bernoulli numbers  $B_i$  are defined by

$$\frac{x}{e^x - 1} = \sum_{i=0}^{\infty} \frac{1}{|i|} B_i x^i \dots\dots\dots(81)$$

Actually expanding the LHS into its Taylor series one can find  $B_0 = 1, B_1 = \frac{-1}{2}, B_2 = \frac{1}{6}$ , and so on

These numbers occur in various operator equations. For example the indefinite summation operator  $\Delta^{-1}$  is defined by

$$\Delta F_k = y_k \quad F_k = \Delta^{-1} y_k$$

and is related to D by

$$\Delta^{-1} = D^{-1} \sum_{i=0}^{\infty} \frac{1}{|i|} B_i D^i \dots\dots\dots(82)$$

where  $B_i$  are Bernoulli numbers. The operator  $D^{-1}$  is the familiar indefinite integral operator.

The Euler-Maclaurin formula may be deduced from the previous relationship

$$\sum_{i=0}^{n-1} y_i = \int_0^n y_k dk - \frac{1}{2}(y_n - y_0) + \frac{h}{12}(y'_n - y'_0) + \dots\dots\dots(83)$$

and is often used for the evaluation of either sums or integrals. The powers of D may be expressed in terms of the central difference operator  $\delta$  by using Taylor series. Some examples are the following

$$D = \mu \left( \delta - \frac{1^2}{|3|} \delta^3 + \frac{1^2 \cdot 2^2}{|5|} \delta^5 - \frac{1^2 \cdot 2^2 \cdot 3^2}{|7|} \delta^7 + \dots \right) \dots\dots\dots(84)$$

$$D^2 = \delta^2 - \frac{1}{12} \delta^4 + \frac{1}{90} \delta^6 - \frac{1}{560} \delta^8 + \frac{1}{3150} \delta^{10} \dots\dots\dots(85)$$

**6. LEAST - SQUARES POLYNOMIAL APPROXIMATION:**

**The Least-Squares Principle:**

The basic idea of choosing a polynomial approximation  $p(x)$  to a given function  $y(x)$  in a way which minimizes the squares of the errors (in some sense) was developed by Gauss. There are several variations depending on the set of arguments involved and the error measure to be used.

First of all when the data are discrete one may minimize the sum

$$S = \sum_{i=0}^N (y_i - a_0 - a_1 x_i - \dots - a_m x_i^m)^2 \dots\dots\dots(86)$$

for the given data  $x_i, y_i$  and  $m < N$ . The condition  $m < N$  makes it unlikely that the polynomial

$$p(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_m x^m \dots\dots\dots(87)$$

can collocate at all N data points. So S probably cannot be made zero. The idea of Gauss is to make as small as one can. Standard techniques of Calculus then lead to the normal equations which determine the coefficients  $a_j$ . These equations are

$$\left. \begin{aligned} s_0 a_0 + s_1 a_1 + \dots + s_m a_m &= t_0 \\ s_1 a_0 + s_2 a_1 + \dots + s_{m+1} a_m &= t_1 \\ \dots\dots\dots \\ s_m a_0 + s_{m+1} a_1 + \dots + s_{2m} a_m &= t_m \end{aligned} \right\} \dots\dots\dots(88)$$

where  $s_k = \sum_{i=0}^N x_i^k, t_k = \sum_{i=0}^N y_i x_i^k$ . This system of linear equations does determine the  $a_i$  uniquely and the resulting  $a_j$  do actually produce the minimum possible value of S. For the case of a linear polynomial

$$p(x) = Mx + B \dots\dots\dots(89)$$

The normal equations are easily solved and yield

$$M = \frac{s_0 t_1 - s_1 t_0}{s_0 s_2 - s_1^2} \dots\dots\dots(90)$$

$$B = \frac{s_2 t_0 - s_1 t_1}{s_0 s_2 - s_1^2} \dots\dots\dots(91)$$

In order to provide a unifying treatment of the various least – squares methods to be presented, including this first method just described a general problem of minimization in a vector space is considered. The solution is easily found by an algebraic argument using the idea of orthogonal projection. Naturally the general problem produces  $p(x)$  and normal equations. It will be reinterpreted to solve other variations of the least – squares principle as one proceed. In most cases a duplicate argument for the special case in hand will also be provided.

Except for very low degree polynomials the above system of normal equations proves to be ill – conditioned. This means that although it does define the coefficients  $a_j$  uniquely in practice it may prove to be impossible to extricate these  $a_j$ . Standard methods for solving linear systems may either produce no solution at all or else badly magnify data errors. As a result orthogonal polynomials are introduced. (This amounts to choosing an orthogonal basis for the abstract vector space). For the case of discrete data these are polynomials  $P_{m,N}(t)$  of degree  $m=0,1,2,\dots$  with property

$$\sum_{t=0}^N P_{m,N}(t) P_{n,N}(t) = 0 \dots\dots\dots(92)$$

This is the orthogonal property. The explicit representation

$$P_{m,N}(t) = \sum_{i=0}^m (-1)^i \binom{m}{i} \binom{m+i}{i} \frac{t^{(i)}}{N^{(i)}} \dots\dots\dots(93)$$

will be obtained in which binomial coefficients and factorial polynomials are prominent.

An alternate form of our least – squares polynomial now becomes convenient namely

$$p(t) = \sum_{k=0}^m a_k P_{k,N}(t) \dots\dots\dots(94)$$

with new coefficients  $a_k$ . The equations determining these  $a_k$  prove to be extremely easy to solve. In fact

$$a_k = \frac{\sum_{t=0}^N y_t P_{k,N}(t)}{\sum_{t=0}^N P_{k,N}^2(t)} \dots\dots\dots(95)$$

These  $a_k$  do minimize the error sum  $S$ , the minimum being

$$S_{\min} = \sum_{t=0}^N y_t^2 - \sum_{k=0}^m W_k a_k^2 \dots\dots\dots(96)$$

where  $W_k$  is the denominator sum in the expression for  $a_k$ .

**Applications**

There are two major applications of least – squares polynomials for discrete data

**i) Data Smoothing:** By accepting the polynomial

$$p(x) = a_0 + a_1 x + \dots\dots\dots + a_m x^m \dots\dots\dots(97)$$

in place of the given  $y(x)$ , one can obtain a smooth line, parabola or other curve in place of the original, probably irregular data function. What degree  $p(x)$  should have depends on the circumstances. Frequently a five-point least – squares parabola is used corresponding to points  $(x_i, y_i)$  with  $i= k-2, k-1, \dots\dots, k+2$ . It leads to the smoothing formula

$$y(x_k) = p(x_k) = y_k - \frac{3}{35} \delta^4 y_k \dots\dots\dots(98)$$

This formula blends together the five values  $y_{k-2}, \dots\dots, y_{k+2}$  to provide a new estimate to the unknown exact value  $y(x_k)$ . Near the ends of a finite data supply, minor modifications are required.

The root mean square error of a set of approximations  $A_i$  to corresponding true values  $T_i$  is defined as

$$RMS\ error = \left[ \sum_{i=0}^N \frac{(T_i - A_i)^2}{N} \right]^{\frac{1}{2}} \dots\dots\dots(99)$$

In various test cases where the  $T_i$  are known we shall use this measure to estimate the effectiveness of least – squares smoothing.

**(ii) Approximate differentiation:**

Fitting a collocation polynomial to irregular data leads to very poor estimates of derivatives. Even small errors in the data are magnified to troublesome size. But a least – squares polynomial does not collocate. It passes between the data values and provides smoothing. This smoother function usually brings better estimates of derivatives, namely, the values of  $p'(x)$ . The five point parabola just mentioned leads to the formula

$$y'(x_k) \approx p'(x_k) = \frac{1}{10^n} (-2y_{k-2} - y_{k-1} + y_{k+1} + 2y_{k+2}) \dots\dots\dots(100)$$

Near the ends of a finite data supply this also requires modification. This formula usually produces results much superior to those obtained by differentiating collocation polynomials. However reapplying it to the  $p'(x_k)$  values in an effort to estimate  $y''(x_k)$  again leads to questionable accuracy.

**Continuous Data:**

For continuous data  $y(x)$  one may minimize the integral.

$$I = \int_{-1}^1 [y(x) - a_0P_0(x) - \dots\dots\dots - a_mP_m(x)]^2 dx \dots\dots\dots(101)$$

The  $P_j(x)$  being Legendre polynomials. This means that one can have chosen to represent our least – squares polynomial  $p(x)$  from the start in terms of orthogonal polynomials in the form

$$p(x) = a_0P_0(x) + \dots\dots\dots + a_mP_m(x) \dots\dots\dots(102)$$

The coefficients prove to be

$$a_k = \frac{2k+1}{2} \int_{-1}^1 y(x)P_k(x)dx \dots\dots\dots(103)$$

For convenience in using the Legendre polynomials the interval over which the data  $y(x)$  are given in first normalized to  $(-1, 1)$ . Occasionally it is more convenient to use the interval  $(0,1)$ . In this case the Legendre polynomials must also be subjected to a change of argument. The new polynomials are called shifted Legendre polynomials.

Some type of discretization is usually necessary when  $y(x)$  is of complicated structure. Either the integrals which give the coefficients must be computed by approximation methods or the continuous agreement set must be discretized at the outset and a sum minimized rather than an integral. Plainly there are several alternate approaches and the computer must decide which to use for a particular problem. Smoothing and approximate differentiation of the given continuous data function  $y(x)$  are again the foremost applications of least squares polynomial  $p(x)$ . One can simply accept  $p(x)$  and  $p'(x)$  as substitutes for the more irregular  $y(x)$  and  $y'(x)$  .

A generalization of the least – squares principle involves minimizing the integral

$$I = \int_a^b w(x)[y(x) - a_0Q_0(x) - \dots\dots\dots - a_mQ_m(x)]^2 dx \dots\dots\dots(104)$$

where  $w(x)$  is a negative weight function. The  $Q_k(x)$  are orthogonal polynomials in the generalized sense

$$\int_a^b w(x)Q_j(x)Q_k(x)dx = 0 \dots\dots\dots(105)$$

for  $j \neq k$  . The details parallel those for the case  $w(x)=1$  already mentioned, the coefficients  $a_k$  being given by

$$a_k = \frac{\int_a^b w(x) y(x) Q_k(x) dx}{\int_a^b w(x) Q_k^2(x) dx} \dots\dots\dots(106)$$

The minimum value of I can be expressed as

$$I_{\min} = \int_a^b w(x) y^2(x) dx - \sum_{k=0}^m W_k a_k^2 \dots\dots\dots(107)$$

where  $W_k$  is the denominator integral in the expression for  $a_k$ . This leads to the Bessel's inequality.

$$\sum_{k=0}^m W_k a_k^2 \leq \int_a^b w(x) y^2(x) dx \dots\dots\dots(108)$$

and to the fact that for  $m$  tending to infinity the series  $\sum_{k=0}^{\infty} W_k a_k^2$  is convergent. If the orthogonal family involved has a property known as completeness and if  $y(x)$  is sufficiently smooth then the series actually converges to the integral which appears  $I_{\min}$ . This means that the error of approximation tends to zero as the degree of  $p(x)$  is increased.

**Chebyshev Polynomials:**

Approximation using Chebyshev polynomials is the important special case  $w(x) = \frac{1}{\sqrt{1-x^2}}$  of the generalized least squares method, the interval of integration being normalized to  $(-1,1)$ . In this case the orthogonal polynomials  $Q_k(x)$  are the Chebyshev polynomials

$$T_k(x) = \cos(K \text{ arc cos } x)$$

The first few prove to be

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1, \quad T_3(x) = 4x^3 - 3x$$

Properties of the Chebyshev polynomials include

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \dots\dots\dots(109)$$

$$\int_{-1}^1 \frac{T_m(x)T_n(x)}{\sqrt{1-x^2}} dx = \begin{cases} 0 & \text{if } m \neq n \\ \frac{\pi}{2} & \text{if } m = n \neq 0 \\ \pi & \text{if } m = n = 0 \end{cases} \dots\dots(110)$$

$$T_n(x) = 0 \quad \text{for } x = \text{Cos} \left[ \frac{(2i+1)\pi}{2n} \right], \quad i = 0, 1, 2, \dots, n-1$$

$$T_n(x) = (-1)^i \quad \text{for } x = \text{Cos} \left[ \frac{i\pi}{n} \right], \quad i = 0, 1, 2, \dots, n$$

An especially attractive property is the equal – error property which refers to the oscillation of the Chebyshev polynomials between extreme values of  $\pm 1$  reaching these extremes at  $n+1$  arguments inside the interval  $(-1,1)$ . As a consequence of this property the error  $y(x) - p(x)$  is frequently found to oscillate between maxima and minima of approximately  $\pm E$ . Such an almost – equal – error is desirable since it implies that our approximation has almost uniform accuracy across the entire interval.

The powers of  $x$  may be expressed in terms of Chebyshev polynomials by simple manipulations. For example

$$1 = T_0 \quad x = T_1 \quad x^2 = \frac{1}{2}(T_0 + T_2) \quad x^3 = \frac{1}{4}(3T_1 + T_3)$$

This has suggested a process known as economization of polynomials by which each power of  $x$  in a polynomial is replaced by the corresponding combination of Chebyshev polynomials. It is often found that a number of the higher – degree Chebyshev polynomials may then be dropped, the terms retained then constituting a least – squares approximation to the original polynomial of sufficient accuracy for many purposes. The result obtained will have the almost – equal–errors property. This process of economization may be used as an approximate substitute for direct evaluation of the coefficient integrals of an approximation by Chebyshev polynomials. The unpleasant weight factor  $w(x)$  makes these integrals formidable for most  $y(x)$ .

Another variation of the least – squares principle is to minimize the sum

$$\sum_{i=0}^{N-1} [y(x_i) - a_0 T_0(x_i) - \dots - a_m T_m(x_i)]^2 \dots\dots\dots(111)$$

the arguments being  $x_i = \cos \left[ \frac{(2i+1)\pi}{2N} \right]$ . These arguments may be recognized as the zeros of  $T_N(x)$ . The coefficients are easily determined using a second orthogonality property of the Chebyshev polynomials

$$\sum_{i=0}^{N-1} T_m(x_i) T_n(x_i) = \begin{cases} 0 & \text{if } m \neq n \\ \frac{N}{2} & \text{if } m = n \neq 0 \\ N & \text{if } m = n = 0 \end{cases} \dots\dots\dots(112)$$

and prove to be  $a_0 = \frac{1}{N} \sum_{i=0}^{N-1} y(x_i)$        $a_k = \frac{2}{N} \sum_{i=0}^{N-1} y(x_i) T_k(x_i)$

The approximating polynomial is then of course

$$p(x) = a_0 T_0(x) + \dots + a_m T_m(x)$$

This polynomial also has an almost – equal – error

**The  $L_2$  Norm:**

The underlying theme of the above discussion is to minimize the norm

$$\|y - p\|_2$$

where  $y$  represents the given data and  $p$  the approximating polynomial

**7. ROOTS OF EQUATIONS:**

The problem treated here is the ancient problem of finding roots of equations or of systems of equations. The long list of available methods shows the long history of this problem and its continuing importance. Which method to use depends upon whether one needs all the roots of a particular equation or only a few whether the roots are real or complex, simple or multiple , whether one has a ready first approximation or not and so on.

**(i) Interpolation Methods:**

These methods use two or more approximations usually some too small and some too large to obtain improved approximations to a root by use of collocation polynomials. The most ancient of these is based on linear interpolation between two previous approximations. It is called regula falsi and solves  $f(x) = 0$  by the iteration

$$x_n = x_{n-1} - \frac{(x_{n-1} - x_{n-2})f(x_{n-1})}{f(x_{n-1}) - f(x_{n-2})} \dots\dots\dots(113)$$

The rate of convergence is between those of the previous two methods. A method based on quadratic interpolation between three successive approximations  $x_0, x_1, x_2$  uses the formula

$$x_3 = x_2 - \frac{2C}{B \pm \sqrt{B^2 - 4AC}} \dots\dots\dots(114)$$



**(ii). Bernoulli’s Method:**

This method produces the dominant root of a real polynomial equation

$$a_0x^n + a_1x^{n-1} + \dots + a_n = 0 \dots\dots\dots(115)$$

provided a single dominant root exists by computing a solution sequence of the difference equation

$$a_0x_k + a_1x_{k-1} + \dots + a_nx_{k-n} = 0 \dots\dots\dots(116)$$

and taking  $\text{Lim} \left( \frac{x_{k+1}}{x_k} \right)$ . The initial values  $x_{-n+1} = \dots = x_{-1} = 0, x_0 = 1$  are usually used. If a complex conjugate pair of roots is dominant then the solution sequence is still computed but the formulas

$$r^2 \approx \frac{x_k^2 - x_{k+1}x_{k-1}}{x_{k-1}^2 - x_kx_{k-2}} \dots\dots\dots(117)$$

$$-2r\cos\phi \approx \frac{x_{k+1}x_{k-2} - x_{k-1}x_k}{x_{k-1}^2 - x_kx_{k-2}} \dots\dots\dots(118)$$

serve to determine the roots as  $r_1, r_2 \approx r(\cos\phi \pm i\sin\phi)$ .

**(iii) Optimization Methods:**

These are based upon the idea that the system  $F=0$  or  $f_i = 0$  for  $i=1,2,\dots,n$  is solved whenever the function

$$S = f_1^2 + f_2^2 + \dots + f_n^2 \dots\dots\dots(119)$$

is minimized since the minimum clearly occurs when all the  $f_i$  are zero. Direct methods for seeking this minimum or descent methods have been developed. For example the two – dimensional problem (with a familiar change of notation)

$$f(x, y) = 0 \quad g(x, y) = 0$$

is equivalent to minimizing this sum

$$S(x, y) = f^2 + g^2 \dots\dots\dots(120)$$

Beginning at an initial approximation  $(x_0, y_0)$  we select the next approximation in the form

$$x_1 = x_0 - tS_{x_0} \quad y_1 = y_0 - tS_{y_0}$$

where  $S_{x_0}$  and  $S_{y_0}$  are the components of the gradient vector of  $S$  at  $(x_0, y_0)$ . Thus progress is in the direction of steepest descent and the algorithm is known as the steepest descent algorithm. The number  $t$  may be chosen to minimize  $S$  in this direction though alternatives have been proposed. Similar steps then follow. The method is often used to provide initial approximations to the Newton Method.

The above equivalence is of course often exploited in the opposite way. To optimize a function  $f(x_1, x_2, \dots, x_n)$  one looks for places where the gradient of  $f$  is zero.

$$\text{grad}(f) = (f_1, f_2, \dots, f_n) = (0, 0, \dots, 0) \dots\dots\dots(121)$$

Here  $f_i$  denotes the partial derivative of  $f$  relative to  $x_i$ . The optimization is then attempted through the solution of the system of  $n$  nonlinear equations.

**(iv) Bairstow’s Method:**

This method produces complex roots of a real polynomial equation  $p(x) = 0$  by applying the Newton method to a related system. More specifically division of  $p(x)$  by a quadratic polynomial suggests the identity

$$p(x) = (x^2 - ux - v)q(x) + r(x) \dots\dots\dots(122)$$

where  $r(x)$  is a linear remainder

$$r(x) = b_{n-1}(u, v)(x - u) + b_n(u, v) \dots\dots\dots(123)$$

The quadratic divisor will be a factor of  $p(x)$  if one can choose  $u$  and  $v$  so that

$$b_{n-1}(u, v) = 0 \quad b_n(u, v) = 0$$

This is the system to which Newton's method is now applied. Once  $u$  and  $v$  are known a complex pair of roots may be found by solving

$$x^2 - ux - v = 0 \dots\dots\dots(124)$$

**Conclusion**

Nonlinear modelling can be utilized in situations where the phenomena are not well understood or expressed in mathematical terms. Thus nonlinear modelling can be an efficient way to model new and complex situations where relationships of different variables are not known. In the above conversation the main concepts in nonlinear modeling namely Taylor polynomial, finite difference operators, least - squares polynomial approximation and the roots of the equations are discussed.

**References**

1. J. Peter Praveen, R. Donthi, S.V.Prasad, B.Mahaboob, B.Venkateswarlu B., A glance on the estimation of Cobb-Douglas production functional model, AIP Conference Proceedings, 2177 (1), 020067. 2019
2. J.Peter Praveen, B. Mahaboob, R.Donthi, S.V. Prasad, B.Venkateswarlu, On stochastic linear regression model selection, AIP Conference Proceedings, 2177 (1), 020068, 2019.
3. B.Mahaboob, J.Peter Praveen, R.Donthi, S.V.Prasad, B.Venkateswarlu B., Criteria for selection of stochastic linear model selection, AIP Conference Proceedings, 2177 (1), 020041,2019.
4. R.Donthi, S.V.Prasad, B.Mahaboob, J.Peter Praveen, B.Venkateswarlu. Estimation methods of nonlinear regression models, AIP Conference Proceedings, 2177 (1), 020081, 2019.
5. R.Donthi, J.Peter Praveen, S.V.Prasad, B.Mahaboob, B.Venkateswarlu, Numerical techniques of nonlinear regression model estimation, AIP Conference Proceedings, 2177 (1), 020082. 2019.
6. B.Mahaboob, S.V.Prasad, J.Peter Praveen, R.Donthi, B.Venkateswarlu, On misspecification tests for stochastic linear regression model, AIP Conference Proceedings, 2177 (1), 020039.2019.
7. B.Mahaboob, Ajmath K.A., Venkateswarlu B., Narayana C., Praveen J.P., On Cobb-Douglas Production function model, AIP Conference Proceedings, 2177 (1), 020040. 2019.
8. B.Mahaboob, K.A.Azmith, B.Venkateswarlu, C.Narayana, B.M.Rao An evaluation in generalized LSE of linearized stochastic statistical model with non-spherical errors, AIP Conference Proceedings, 2177(1), 020038. 2019.
9. J. Peter Praveen, B. Nageswara Rao, Y.Haranadh, B. V. Appa Rao, C. Narayana, B. Mahaboob, Existence Of Solution Of The Nonlinear Differential Equation In The Modeling Of Eardrum By Using Homotopy Perturbation Method” *Advances in Mathematics Scientific journal* Vol.9 (7), pp.1- 9, 2020.
10. B. Mahaboob, J. Peter Praveen, B. V. Appa Rao, Y. Haranadh, C. Narayana, and G. Balaji Prakash, A Study On Multiple Linear Regression Using Matrix Calculus” , *Advances in Mathematics Scientific journal*, Vol.9(7), pp.1-10 ,2020.
11. J. Peter Praveen, B. Nageswara Rao, B. Mahaboob, C. Rajaiah, Y. Harnath, And C. Narayana, “Series Decomposition Method For Asymmetric Nonlinear Oscillations”, Vol.0,No.10,pp.8069-8076. (2020)
12. B. Mahaboob, K. Sreenivasulu, M. Rajaiah, Y. Harnath, C. Narayana, and J. Peter Praveen, “Jordan Generalized Derivations In Gamma Nearrings” *Advances in Mathematics Scientific journal*, Vol.9,no.10, (2020)
13. J.Peter Praveen, B.Nageswara Rao, B.Mahaboob, B.V.Appa Rao, “An Application of Laplace Transform, Inverse Laplace Transform and Pade’s Approximant in the Periodic Solution of Duffing Equation of Motion” , *International Journal of Emerging Trends in Engineering Research* ,Volume 8. No. 9, September 2020
14. J. Peter Praveen, B. Nageswara Rao, B.Mahaboob, M.Rajaiah, Y.Harnath, C.Narayana, “On The Simulation of Beck Column through a Simple Xiong-Wang-Tabarrok Experimental Model of Centripetally Loaded Column” *International Journal of Emerging Trends in Engineering Research*, Volume 8. No. 9, September 2020
15. B. Mahaboob, B. Venkateswarlu, G.S.G.N. Anjaneyulu, C. Narayana, “ A Discourse on Applications of Lie Groups and Lie Algebras” *International Journal Of Scientific & Technology Research* Volume 9, Issue 02, February 2020,pp.6226-6231.
16. B.Venkateswarlu, B. Mahaboob, K.A. Ajmath, C. Narayana, “Application of DEA in Super Efficiency Estimation” , *International Journal Of Scientific & Technology Research*, Research Volume 9, Issue 02, February 2020 pp.4496-4499.
17. B. Venkateswarlu, B. Mahaboob, C. Subbarami Reddy, C. Narayana, “Evaluating Different Types Of Efficiency Stability Regions And Their Infeasibility In DEA”, *International Journal of Scientific Technology and Research*, Volume 9, Issue 02, February 2020, Pages:3944-3949
18. B.Venkateswarlu, B. Mahaboob, C. Subbaramireddy, C. Narayana, “Multi-Criteria Optimization Techniques in DEA: Methods & Applications”, *International Journal of Scientific Technology and Research*, Volume 9, Issue 02, February 2020, Pages:509-515.
19. B. Mahaboob, B.Venkateswarlu, K.A. Azmath, C. Narayana, J. Peter Praveen, “On OLS Estimation of Stochastic Linear Regression Model, *International Journal of Engineering and Advanced Technology*, Volume-8 Issue-6, pp.1953-1955

20. B. Mahaboob, B.Venkateswarlu, C. Narayana, M. Sivaiah,” Bias In The Maximum Likelihood Estimation Of Parameters Of Nonlinear Regression Models”, International Journal of Scientific & Technology Research, Volume 8, Issue 11, November 2019 pgs:1252\_1255
21. B. Venkateswarlu, B. Mahaboob, C. Narayana, C. Subbarami Reddy, “Evaluation Of Slack Based Efficiency Of A Decision Making Unit” , International Journal Of Scientific & Technology Research, , Volume 8, Issue 11, November 2019 .Pgs:1178-1182
22. B.Venkateswarlu, B. Mahaboob, K.A. Azmath, C. Narayana, C. Muralidaran, “An Application of Linear Programming in the Estimation of Technical Efficiency of DMU”, International Journal of Engineering and Advanced Technology , Volume-8 Issue-6, August, 2019,pp.1956-1959.
23. B.Venkateswarlu B. Mahaboob, C. Subbarami Reddy, C. Narayana, “New Results Production Theory By Applying Goal Programming”, International Journal Of Scientific & Technology Research, Volume 8, Issue 12, December 2019 Pgs:1918-1923
24. B. Mahaboob, B.Venkateswarlu, C. Narayana, J. Ravi sankar and P. Balasiddamuni, “A Treatise on Ordinary Least Squares Estimation of Parameters of Linear model” International Journal of Engineering and Technology, vol. 7(4.10)(2018) 518-522
25. B. Mahaboob, B.Venkateswarlu, C. Narayana, J. Ravi sankar and P. Balasiddamuni, “A Monograph on Nonlinear Regression Models”, International Journal of Engineering and Technology,vol. 7(4.10)(2018) 543-546
26. B.Venkateswarlu, B. Mahaboob,J. Ravi sankar, C. Narayana and B. Madhusudhana Rao, “An Application of Goal Programming in Data Envelopment Analysis” International Journal of Engineering and Technology,vol. vol. 7(4.10)(2018) 523-525
27. C. Narayana , B. Mahaboob , B.Venkateswarlu , J. Ravi sankar, “A Memoir on Model Selection Criterion between Two Nested and Non-Nested Stochastic Nonlinear Regression Models”, International Journal of Engineering and Technology, vol. 7(4.10)(2018) 529-531.
28. C. Narayana, B. Mahaboob, B.Venkateswarlu, J. Ravi sankar and P. Balasiddamuni, “A Study on Misspecification and Predictive Accuracy of Stochastic Linear Regression Models”, International Journal of Engineering and Technology, vol. vol. 7(4.10)(2018) ,532-535.
29. C. Narayana, B. Mahaboob, B.Venkateswarlu, J. Ravi sankar and P. Balasiddamuni, “A Discourse on Modified Likelihood Ratio (LR), Wald and Lagrange Multipliers (LM) Tests for Testing General Linear Hypothesis in Stochastic Linear Regression Model” , International Journal of Engineering and Technology, vol. 7(4.10)(2018) 536-538
30. C. Narayana, B. Mahaboob, B.Venkateswarlu, J. Ravi sankar and P. Balasiddamuni, “A Treatise on Testing General Linear Hypothesis in Stochastic Linear Regression Model, International Journal of Engineering and Technology, vol. 7(4.10)(2018) 539-542..
31. B Mahaboob, B Venkateswarlu and J Ravi Sankar, “Estimation of parameters of constant elasticity of substitution production functional model” IOP Conference Series: Materials Science and Engineering, 263 (2017) 042121.
32. B Mahaboob, B Venkateswarlu, J Ravi Sankar and P Balasiddamuni, “Computation of nonlinear least squares estimator and maximum likelihood using principles in matrix calculus”, IOP Conference Series: Materials Science and Engineering, 263 (2017) 042125
33. B. Venkateswarlu, B. Mahaboob, C Subbarami Reddy and J Ravi Sankar, “A study on technical efficiency of a DMU (review of literature) “, IOP Conference Series: Materials Science and Engineering, 263 (2017) 042124
34. B Mahaboob, B Venkateswarlu, G Mokeshrayalu and P Balasiddamuni, “New methods of testing nonlinear hypothesis using iterative NLLS estimator” , IOP Conference Series: Materials Science and Engineering, 263 (2017) 042126
35. B Venkateswarlu, B Mahaboob, C Subbarami Reddy and B Madhusudhana Rao, “Fitting of Full Cobb-Douglas and Full VRTS Cost Frontiers by solving Goal Programming Problem” IOP Conference Series: Materials Science and Engineering, 263 (2017) 042127
36. B Mahaboob, B Venkateswarlu, G Mokeshrayalu and P Balasiddamuni, “A Different Approach to Estimate Nonlinear Regression Model using Numerical Methods”, IOP Conference Series: Materials Science and Engineering, 263 (2017) 042122
37. K. Lakshmi, B. Mahaboob,, M. Rajaiiah, C. Narayana,” Ordinary least squares estimation of parameters of linear model”, Journal of Mathematical and Computational Science, Vol 11 (2021), No. 2, 2015-2030 ISSN: 1927-5307
38. B. Venkateswarlu, M. Mubashir Unnissa, and B. Mahaboob” Estimating Cost Analysis using Goal Programming”, Indian Journal of Science and Technology, Vol.9 (44), ISSN:0974-5645,November 2016
- 39.C.Narayana , P.Sreehari Reddy , B.Mahaboob, Y.Hari Krishna, K.Lakshmi and B.Sivaram, “A brief literature on problems and perspectives of mathematical and stochastic modelling”, Turkish Journal of Computer and Mathematics Education, Vol.12 No.4 (2021), 1267-1275\_e-ISSN 1309-4653

40.K.Lakshmi,B.Mahaboob,M.Rajaiah,B.Sivaram,Y.Hari Krishna and C.Narayana,” A study on mathematical and statistical aspects of linear models”, Turkish Journal of Computer and Mathematics Education, Vol.12 No.4 (2021), 1328-1338, , e-ISSN 1309-4653