Chapter 4

Least Squares and Fourier Transforms

4.1 INTRODUCTION

In experimental work, we often encounter the problem of fitting a curve to data which are subject to errors. The strategy for such cases is to derive an approximating function that *broadly* fits the data without necessarily passing through the given points. The curve drawn is such that the discrepancy between the data points and the curve is least. In the method of least squares, the sum of the squares of the errors is minimized. For continuous functions, the method is discussed in Section 4.4.

The problem of approximating a function by means of Chebyshev polynomials is described in Section 4.5. This is important from the standpoint of digital computation.

In Chapter 3, we concentrated on polynomial interpolation, i.e., interpolation based on a linear combination of functions 1, x, x^2, …, x^n. On the other hand, trigonometric interpolation, i.e., interpolation based on trigonometric functions such as, cos x, sin x, cos $2x$, sin $2x$, … plays an important role in modelling vibrating systems. The Fourier series is a useful tool for dealing with periodic systems; but for aperiodic systems, the Fourier transform is the primary tool available. The computations of discrete Fourier transform and the Fast Fourier Transform (FFT) are discussed in detail in Section 4.6.

4.2 LEAST SQUARES CURVE FITTING PROCEDURES

Let the set of data points be (x_i, y_i), $i = 1, 2, \ldots, m$, and let the curve given by $Y = f(x)$ be fitted to this data. At $x = x_i$, the given ordinate is y_i and the corresponding value on the fitting curve is $f(x_i)$. If e_i is the error of approximation at $x = x_i$, then we have

$$e_i = y_i - f(x_i) \tag{4.1}$$

If we write

$$\begin{aligned} S &= [y_1 - f(x_1)]^2 + [y_2 - f(x_2)]^2 + \cdots + [y_m - f(x_m)]^2 \\ &= e_1^2 + e_2^2 + \cdots + e_m^2, \end{aligned} \tag{4.2}$$

then the method of least squares consists in minimizing S, i.e., the sum of the squares of the errors. In the following sections, we shall study the linear and nonlinear least squares fitting to given data (x_i, y_i), $i = 1, 2, \ldots, m$.

4.2.1 Fitting a Straight Line

Let $Y = a_0 + a_1 x$ be the straight line to be fitted to the given data, viz. (x_i, y_i), $i = 1, 2, \ldots, m$. Then, corresponding to Eq. (4.2), we have

$$\begin{aligned} S = {} & [y_1 - (a_0 + a_1 x_1)]^2 + [y_2 - (a_0 + a_1 x_2)]^2 \\ & + \cdots + [y_m - (a_0 + a_1 x_m)]^2 \end{aligned} \tag{4.3}$$

For S to be minimum, we have

$$\begin{aligned} \frac{\partial S}{\partial a_0} = 0 = {} & -2[y_1 - (a_0 + a_1 x_1)] - 2[y_2 - (a_0 + a_1 x_2)] \\ & - \cdots -2[y_m - (a_0 + a_1 x_m)] \end{aligned} \tag{4.4a}$$

and

$$\begin{aligned} \frac{\partial S}{\partial a_1} = 0 = {} & -2x_1[y_1 - (a_0 + a_1 x_1)] - 2x_2\,[y_2 - (a_0 + a_1 x_2)] \\ & - \cdots -2x_m\,[y_m - (a_0 + a_1 x_m)] \end{aligned} \tag{4.4b}$$

The above equations simplify to

$$m a_0 + a_1(x_1 + x_2 + \cdots + x_m) = y_1 + y_2 + \cdots + y_m \tag{4.5a}$$

and

$$a_0(x_1 + x_2 + \cdots + x_m) + a_1\left(x_1^2 + x_2^2 + \cdots + x_m^2\right) = x_1 y_1 + x_2 y_2 + \cdots + x_m y_m \tag{4.5b}$$

or more compactly to

$$m a_0 + a_1 \sum_{i=1}^{m} x_i = \sum_{i=1}^{m} y_i \tag{4.6a}$$

and

$$a_0 \sum_{i=1}^{m} x_i + a_1 \sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} x_i y_i \tag{4.6b}$$

Equations (4.6) are called the *normal equations*, and can be solved for a_0 and a_1, since x_i and y_i are known quantities.

We can obtain easily

$$a_1 = \frac{m\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \cdot \sum_{i=1}^{m} y_i}{m\sum_{i=1}^{m} x_i^2 - \left(\sum_{i=1}^{m} x_i\right)^2} \tag{4.7}$$

and then

$$a_0 = \bar{y} - a_1\bar{x}. \tag{4.8}$$

Since $\frac{\partial^2 S}{\partial a_0^2}$ and $\frac{\partial^2 S}{\partial a_1^2}$ are both positive at the points a_0 and a_1, it follows that these values provide a *minimum* of S. In Eq. (4.8), $\bar{x}$ and $\bar{y}$ are the means of x and y, respectively. From Eq. (4.8), we have

$$\bar{y} = a_0 + a_1\bar{x},$$

which shows that the fitted straight line passes through the centroid of the data points.

Sometimes, a goodness of fit is adopted. The correlation coefficient (cc) is defined as

$$\text{cc} = \sqrt{\frac{S_y - S}{S_y}}, \tag{4.9}$$

where

$$S_y = \sum_{i=1}^{m} (y_i - \bar{y})^2 \tag{4.10}$$

and S defined by Eq. (4.3).

If cc is close to 1, then the fit is considered to be good, although this is not always true.

Example 4.1 Find the best values of a_0 and a_1 if the straight line $Y = a_0 + a_1 x$ is fitted to the data (x_i, y_i):

(1, 0.6), (2, 2.4), (3, 3.5), (4, 4.8), (5, 5.7)

Find also the correlation coefficient.

From the table of values given below, we find $\bar{x} = 3$, $\bar{y} = 3.4$, and

$$a_1 = \frac{5(63.6) - 15(17)}{5(55) - 225} = 1.26$$

Therefore,

$$a_0 = \bar{y} - a_1\bar{x} = -0.38.$$

x_i	y_i	x_i^2	$x_i y_i$	$(y_i - \bar{y})^2$	$(y_i - a_0 - a_1 x_i)^2$
1	0.6	1	0.6	7.84	0.0784
2	2.4	4	4.8	1.00	0.0676
3	3.5	9	10.5	0.01	0.0100
4	4.8	16	19.2	1.96	0.0196
5	5.7	25	28.5	5.29	0.0484
15	17.0	55	63.6	16.10	0.2240

The correlation coefficient = $\sqrt{\dfrac{16.10 - 0.2240}{16.10}} = 0.9930.$

Example 4.2 Certain experimental values of x and y are given below:

$$(0, -1),\ (2, 5),\ (5, 12),\ (7, 20)$$

If the straight line $Y = a_0 + a_1 x$ is fitted to the above data, find the approximate values of a_0 and a_1.

The table of values is given below.

x	y	x^2	xy
0	–1	0	0
2	5	4	10
5	12	25	60
7	20	49	140
14	36	78	210

The normal equations are

$$4a_0 + 14a_1 = 36$$

and

$$14a_0 + 78a_1 = 210$$

Solving the two equations, we obtain

$$a_0 = -1.1381 \quad \text{and} \quad a_1 = 2.8966$$

Hence the best straight line fit is given by

$$Y = -1.1381 + x(2.8966).$$

4.2.2 Multiple Linear Least Squares

Suppose that z is a linear function of two variables x and y. If the function $z = a_0 + a_1 x + a_2 y$ is fitted to the data (z_1, x_1, y_1), (z_2, x_2, y_2), …(z_m, x_m, y_m), then the sum

$$S = \sum_{i=1}^{m} (z_i - a_0 - a_1 x_i - a_2 y_i)^2$$

should be minimum. For this, we have

$$\frac{\partial S}{\partial a_0} = -2\sum (z_i - a_0 - a_1 x_i - a_2 y_i) = 0,$$

$$\frac{\partial S}{\partial a_1} = -2x_i \sum (z_i - a_0 - a_1 x_i - a_2 y_i) = 0,$$

and

$$\frac{\partial S}{\partial a_2} = -2y_i \sum (z_i - a_0 - a_1 x_i - a_2 y_i) = 0.$$

These equations simplify to

$$\left.\begin{aligned} ma_0 + a_1 \Sigma x_i + a_2 \Sigma y_i &= \Sigma z_i \\ a_0 \Sigma x_i + a_1 \Sigma x_i^2 + a_2 \Sigma x_i y_i &= \Sigma z_i x_i \\ a_0 \Sigma y_i + a_1 \Sigma y_i x_i + a_2 \Sigma y_i^2 &= \Sigma z_i y_i \end{aligned}\right\} \tag{4.11}$$

from which a_0, a_1 and a_2 can be determined.

Example 4.3 Find the values of a_0, a_1 and a_2, so that the function $z = a_0 + a_1 x + a_2 y$ is fitted to the data (x, y, z) given below.

(0, 0, 2), (1, 1, 4), (2, 3, 3), (4, 2, 16) and (6, 8, 8).

We form the following table of values

x	y	z	x^2	xy	zx	y^2	yz
0	0	2	0	0	0	0	0
1	1	4	1	1	4	1	4
2	3	3	4	6	6	9	9
4	2	16	16	8	64	4	32
6	8	8	36	48	48	64	64
13	14	33	57	63	122	78	109

The normal equations are

$$5a_0 + 13a_1 + 14a_2 = 33$$

$$13a_0 + 57a_1 + 63a_2 = 122$$

$$14a_0 + 63a_1 + 78a_2 = 109$$

The solution of the above system is

$$a_0 = 2,\ a_1 = 5 \text{ and } a_2 = -3.$$

4.2.3 Linearization of Nonlinear Laws

The given data may not always follow a linear relationship. This can be ascertained from a plot of the given data. If a nonlinear model is to be fitted, it can be conveniently transformed to a linear relationship. Some nonlinear laws and their transformations are given as follows.

(a) $y = ax + \dfrac{b}{x}$

This can be written as

$$xy = ax^2 + b$$

Put $xy = Y$, $x^2 = X$. With these transformations, it becomes a linear model.

(b) $xy^a = b$

Taking logarithms of both sides, we get

$$\log_{10}x + a\log_{10}y = \log_{10}b.$$

In this case, we put

$$\log_{10}y = Y,\ \log_{10}x = X,$$

$$\frac{1}{a}\log_{10}b = A_0 \text{ and } -\frac{1}{a} = A_1,$$

so that

$$Y = A_0 + A_1X.$$

(c) $y = ab^x$

Taking logarithms of both sides, we obtain

$$\log_{10}y = \log_{10}a + x\log_{10}b$$

$$\Rightarrow Y = A_0 + A_1X,$$

where

$$Y = \log_{10}y,\ A_0 = \log_{10}a,$$

$$X = x, \text{ and } A_1 = \log_{10}b$$

(d) $y = ax^b$

We have

$$\log_{10}y = \log_{10}a + b\log_{10}x$$

$$\Rightarrow Y = A_0 + A_1X,$$

where

$$Y = \log_{10}y,\ A_0 = \log_{10}a,\ A_1 = b$$

and

$$X = \log_{10}x.$$

(e) $y = ae^{bx}$

In this case, we write

$$\ln y = \ln a + bx$$

$$\Rightarrow Y = A_0 + A_1X,$$

where

$$Y = \ln y,\ A_0 = \ln a,\ A_1 = b$$

and

$$X = x.$$

Example 4.4 Using the method of least squares, find constants a and b such that the function $y = ae^{bx}$ fits the following data:

(1.0, 2.473), (3.0, 6.722), (5.0, 18.274), (7.0, 49.673), (9.0, 135.026).

We have

$$y = ae^{bx}$$

Therefore,

$$\ln y = \ln a + bx$$

$$\Rightarrow Y = A_0 + A_1 X,$$

where

$$Y = \ln y,\ A_0 = \ln a,\ A_1 = b \text{ and } X = x.$$

The table of values is given below

X	$Y = \ln y$	X^2	XY
1	0.905	1	0.905
3	1.905	9	5.715
5	2.905	25	14.525
7	3.905	49	27.335
9	4.905	81	44.145
25	14.525	165	92.625

We obtain

$$\bar{X} = 5,\ \bar{Y} = 2.905$$

$$A_1 = \frac{5(92.625) - 25(14.525)}{5(165) - 625} = 0.5 = b.$$

Then

$$A_0 = \bar{Y} - A_1\bar{X} = 2.905 - 0.5(5) = 0.405.$$

Hence,

$$a = e^{A_0} = e^{0.405} = 1.499.$$

It follows that the required curve is of the form

$$y = 1.499e^{0.5x}$$

Example 4.5 Using the method of least squares, fit a curve of the form $y = \dfrac{x}{a + bx}$ to the following data

(3, 7.148), (5, 10.231), (8, 13.509), (12, 16.434).

We have

$$y = \frac{x}{a + bx}$$

$$\Rightarrow \frac{1}{y} = \frac{a + bx}{x} = b + \frac{a}{x}$$

$$\Rightarrow Y = A_0 + A_1 X,$$

where

$$A_0 = b,\ A_1 = a,\ X = \frac{1}{x} \text{ and } Y = \frac{1}{y}.$$

The table of values is

X	Y	X^2	XY
0.333	0.140	0.111	0.047
0.200	0.098	0.040	0.020
0.125	0.074	0.016	0.009
0.083	0.061	0.007	0.005
0.741	0.373	0.174	0.081

We obtain

$$A_1 = a = \frac{4(0.081) - 0.741(0.373)}{4(0.174) - (0.741)^2} = 0.324,\ \bar{X} = 0.185,\ \bar{Y} = 0.093$$

and $A_0 = b = \bar{Y} - a\bar{X} = 0.0331$.

Hence the required fit is $Y = 0.0331 + 0.324(X)$, which simplifies to

$$y = \frac{x}{0.324 + 0.0331(x)}.$$

$$\left[\text{Note: The given data is obtained from the relation } y = \frac{x}{0.3162 + 0.0345x}\right]$$

4.2.4 Curve Fitting by Polynomials

Let the polynomial of the nth degree,

$$Y = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \tag{4.12}$$

be fitted to the data points (x_i, y_i), $i = 1, 2, \ldots, m$. We then have

$$\begin{aligned} S &= \left[y_1 - (a_0 + a_1x_1 + a_2x_1^2 + \cdots + a_nx_1^n)\right]^2 \\ &\quad + \left[y_2 - (a_0 + a_1x_2 + a_2x_2^2 + \cdots + a_nx_2^n)\right]^2 \\ &\quad + \cdots + \left[y_m - (a_0 + a_1x_m + a_2x_m^2 + \cdots + a_nx_m^n)\right]^2 \end{aligned} \tag{4.13}$$

Equating to zero the first partial derivatives and simplifying, we obtain the normal equations:

$$\left.\begin{aligned} ma_0 + a_1\Sigma x_i + a_2\Sigma x_i^2 + \cdots + a_n\Sigma x_i^n &= \Sigma y_i, \\ a_0\Sigma x_i + a_1\Sigma x_i^2 + \cdots + a_n\Sigma x_i^{n+1} &= \Sigma x_i y_i, \\ &\vdots \\ a_0\Sigma x_i^n + a_1\Sigma x_i^{n+1} + \cdots + a_n\Sigma x_i^{2n} &= \Sigma x_i^n y_i \end{aligned}\right\} \tag{4.14}$$

where the summations are performed from $i = 1$ to $i = m$.

The system (4.14) constitutes $(n + 1)$ equations in $(n + 1)$ unknowns, and hence can be solved for a_0, a_1, …, a_n. Equation (4.12) then gives the required polynomial of the nth degree.

For larger values of n, system (4.14) becomes unstable with the result that round off errors in the data may cause large changes in the solution. Such systems occur quite often in practical problems and are called *ill-conditioned* systems. Orthogonal polynomials are most suited to solve such systems and one particular form of these polynomials, the Chebyshev polynomials, will be discussed later in this chapter.

Example 4.6 Fit a polynomial of the second degree to the data points (x, y) given by

$$(0, 1), (1, 6) \text{ and } (2, 17).$$

For $n = 2$, Eq. (4.14) requires $\Sigma x_i, \Sigma x_i^2, \Sigma x_i^3, \Sigma x_i^4, \Sigma y_i, \Sigma x_i y_i$ and $\Sigma x_i^2 y_i$. The table of values is as follows:

x	y	x^2	x^3	x^4	xy	x^2y
0	1	0	0	0	0	0
1	6	1	1	1	6	6
2	17	4	8	16	34	68
3	24	5	9	17	40	74

The normal equations are

$$3a_0 + 3a_1 + 5a_2 = 24$$
$$3a_0 + 5a_1 + 9a_2 = 40$$
$$5a_0 + 9a_1 + 17a_2 = 74$$

Solving the above system, we obtain

$$a_0 = 1, a_1 = 2 \text{ and } a_2 = 3.$$

The required polynomial is given by $Y = 1 + 2x + 3x^2$, and it can be seen that this fitting is *exact.*

Example 4.7 Fit a second degree parabola $y = a_0 + a_1x + a_2x^2$ to the data (x_i, y_i):

$$(1, 0.63), (3, 2.05), (4, 4.08), (6, 10.78).$$

The table of values is

x	y	x^2	x^3	x^4	xy	x^2y
1	0.63	1	1	1	0.63	0.63
3	2.05	9	27	81	6.15	18.45
4	4.08	16	64	256	16.32	65.28
6	10.78	36	216	1296	64.68	388.08
14	17.54	62	308	1634	87.78	472.44

The normal equations are

$$4a_0 + 14a_1 + 62a_2 = 17.54$$
$$14a_0 + 62a_1 + 308a_2 = 87.78$$
$$62a_0 + 308a_1 + 1634a_2 = 472.44,$$

from which we obtain

$$a_0 = 1.24, \ a_1 = -1.05 \quad \text{and} \quad a_2 = 0.44$$

4.2.5 Curve Fitting by a Sum of Exponentials

A frequently encountered problem in engineering and physics is that of fitting a sum of exponentials of the form

$$y = f(x) = A_1 e^{\lambda_1 x} + A_2 e^{\lambda_2 x} + \cdots + A_n e^{\lambda_n x} \tag{4.15}$$

to a set of data points (x_i, y_i), $i = 1, 2, \ldots, m$, where m is *much greater than* $2n$.

We describe here a computational technique due to Moore [1974]. For easy of presentation, we assume $n = 2$.

Then the function

$$y = A_1 e^{\lambda_1 x} + A_2 e^{\lambda_2 x} \tag{4.16}$$

is to be fitted to the data (x_i, y_i), $i = 1, 2, \ldots, m$, and $m \gg 4$. It is known that $y(x)$ satisfies a differential equation of the form

$$\frac{d^2 y}{dx^2} = a_1 \frac{dy}{dx} + a_2 y \tag{4.17}$$

where the constants a_1 and a_2 have to be determined. Integrating Eq. (4.17), we obtain

$$y'(x) - y'(x_0) = a_1 [y(x) - y(0)] + a_2 \int_{x_0}^{x} y(x)\, dx, \tag{4.18}$$

where x_0 is the initial value of x and $y'(x) = \dfrac{dy}{dx}$. Integrating Eq. (4.18), we get

$$y(x) - y(0) - (x - x_0)\, y'(x_0) = a_1 \int_{x_0}^{x} y(x)\, dx - a_1 (x - x_0)\, y(x_0)$$
$$+ a_2 \int_{x_0}^{x} \int_{x_0}^{x} y(x)\, dx\, dx \tag{4.19}$$

Now,

$$\int_{x_0}^{x}\int_{x_0}^{x} y(x)\,dx\,dx = \int_{x_0}^{x}(x-t)y(t)\,dt$$

Hence, Eq. (4.19) becomes

$$y(x) - y(0) - (x - x_0)\,y'(x_0) = a_1\int_{x_0}^{x} y(x)\,dx - a_1(x - x_0)\,y(x_0)$$

$$+\, a_2\int_{x_0}^{x}(x-t)\,y(t)\,dt \qquad (4.20)$$

In Eq. (4.20), $y'(x_0)$ is eliminated in the following way. Let x_1 and x_2 be two data points such that

$$x_0 - x_1 = x_2 - x_0 \qquad (4.21)$$

Then Eq. (4.20) gives

$$y(x_1) - y(x_0) - (x_1 - x_0)\,y'(x_0) = a_1\int_{x_0}^{x_1} y(x)\,dx - a_1(x_1 - x_0)\,y(x_0)$$

$$+\, a_2\int_{x_0}^{x_1}(x_1-t)\,y(t)\,dt \qquad (4.22)$$

and

$$y(x_2) - y(x_0) - (x_2 - x_0)\,y'(x_0) = a_1\int_{x_0}^{x_2} y(x)\,dx - a_1(x_2 - x_0)\,y(x_0)$$

$$+\, a_2\int_{x_0}^{x_2}(x_2-t)\,y(t)\,dt \qquad (4.23)$$

Adding Eqs. (4.22) and (4.23) and using Eq. (4.21), we obtain

$$y(x_1) + y(x_2) - 2y(x_0) = a_1\left[\int_{x_0}^{x_1} y(x)dx + \int_{x_0}^{x_2} y(x)\,dx\right]$$

$$+\, a_2\left[\int_{x_0}^{x_1}(x_1-t)y(t)dt + \int_{x_0}^{x_2}(x_2-t)\,y(t)\,dt\right] \qquad (4.24)$$

Equation (4.24) can now be used to set up a linear system of equations for a_1 and a_2, and then we obtain λ_1 and λ_2 from the characteristic equation

$$\lambda^2 = a_1\lambda + a_2 \tag{4.25}$$

Finally, A_1 and A_2 can be obtained by the method of least squares or by the method of averages.

Example 4.8 Fit a function of the form

$$y = A_1 e^{\lambda_1 x} + A_2 e^{\lambda_2 x} \tag{i}$$

to the data defined by (x, y)

(1, 1.54), (1.1, 1.67), (1.2, 1.81), (1.3, 1.97), (1.4, 2.15),

(1.5, 2.35), (1.6, 2.58), (1.7, 2.83), (1.8, 3.11).

Let $x_0 = 1.2$, $x_1 = 1.0$, $x_2 = 1.4$. Then, Eq. (4.24) gives

$$0.07 = a_1\left[-\int_{1.0}^{1.2} y(x)\,dx + \int_{1.2}^{1.4} y(x)\,dx\right]$$

$$+\, a_2\left[-\int_{1.0}^{1.2} (1.0 - t)\,y(t)\,dt + \int_{1.2}^{1.4} (1.4 - t)\,y(t)\,dt\right]$$

Evaluating the integrals by Simpson's rule* and simplifying, the above equation becomes

$$1.81a_1 + 2.180a_2 = 2.10 \tag{ii}$$

Again, choosing $x_1 = 1.4$, $x_0 = 1.6$ and $x_2 = 1.8$, and evaluating the integrals as before, we obtain the equation

$$2.88a_1 + 3.104a_2 = 3.00 \tag{iii}$$

Solving Eqs. (ii) and (iii), we get

$$a_1 = 0.03204 \quad \text{and} \quad a_2 = 0.9364.$$

Equation (4.25) now gives

$$\lambda^2 - 0.03204\lambda - 0.9364 = 0,$$

from which we obtain

$$\lambda_1 = 0.988 = 0.99,$$

and

$$\lambda_2 = -0.96.$$

Using the method of least squares, we finally obtain

$$A_1 = 0.499 \quad \text{and} \quad A_2 = 0.491.$$

The above data was actually constructed from the function $y = \cosh x$ so that $A_1 = A_2 = 0.5$, $\lambda_1 = 1.0$ and $\lambda_2 = -1.0$.

---

*See Section 6.4.2.

4.3 WEIGHTED LEAST SQUARES APPROXIMATION

In the previous section, we have minimized the sum of squares of the errors. A more general approach is to minimize the weighted sum of the squares of the errors taken over all data points. If this sum is denoted by S, then instead of Eq. (4.2), we have

$$S = W_1 [y_1 - f(x_1)]^2 + W_2 [y_2 - f(x_2)]^2 + \cdots + W_m [y_m - f(x_m)]^2$$

$$= W_1 e_1^2 + W_2 e_2^2 + \cdots + W_m e_m^2. \tag{4.26}$$

In Eq. (4.26), the W_i are prescribed positive numbers and are called *weights*. A weight is prescribed according to the relative accuracy of a data point. If all the data points are accurate, we set $W_i = 1$ for all i. We consider again the linear and nonlinear cases below.

4.3.1 Linear Weighted Least Squares Approximation

Let $Y = a_0 + a_1 x$ be the straight line to be fitted to the given data points, viz. $(x_1, y_1),\ldots,(x_m, y_m)$. Then

$$S(a_0, a_1) = \sum_{i=1}^{m} W_i [y_i - (a_0 + a_1 x_i)]^2. \tag{4.27}$$

For maxima or minima, we have

$$\frac{\partial S}{\partial a_0} = \frac{\partial S}{\partial a_1} = 0, \tag{4.28}$$

which give

$$\frac{\partial S}{\partial a_0} = -2 \sum_{i=1}^{m} W_i [y_i - (a_0 + a_1 x_i)] = 0 \tag{4.29}$$

and

$$\frac{\partial S}{\partial a_1} = -2 \sum_{i=1}^{m} W_i [y_i - (a_0 + a_1 x_i)]\, x_i = 0. \tag{4.30}$$

Simplification yields the system of equations for a_0 and a_1:

$$a_0 \sum_{i=1}^{m} W_i + a_1 \sum_{i=1}^{m} W_i x_i = \sum_{i=1}^{m} W_i y_i \tag{4.31}$$

and

$$a_0 \sum_{i=1}^{m} W_i x_i + a_1 \sum_{i=1}^{m} W_i x_i^2 = \sum_{i=1}^{m} W_i x_i y_i, \tag{4.32}$$

which are the *normal equations* in this case and are solved to obtain a_0 and a_1. We consider Example 4.2 again to illustrate the use of weights.

Example 4.9 Suppose that in the data of Example 4.2, the point (5, 12) is known to be more reliable than the others. Then we prescribe a weight (say, 10) corresponding to this point only and all other weights are taken as unity. The following table is then obtained.

x	y	W	Wx	Wx^2	Wy	Wxy
0	–1	1	0	0	–1	0
2	5	1	2	4	5	10
5	12	10	50	250	120	600
7	20	1	7	49	20	140
14	36	13	59	303	144	750

The normal Eqs. (4.31) and (4.32) then give

$$13a_0 + 59a_1 = 144 \quad \text{(i)}$$

$$59a_0 + 303a_1 = 750. \quad \text{(ii)}$$

Solution to Eqs. (i) and (ii) gives

$$a_0 = -1.349345 \quad \text{and} \quad a_1 = 2.73799.$$

The 'linear least squares approximation' is, therefore, given by

$$y = -1.349345 + 2.73799x.$$

Example 4.10 We consider Example 4.9 again with an increased weight, say 100, corresponding to $y(5.0)$. The following table is then obtained.

x	y	W	Wx	Wx^2	Wy	Wxy
0	–1	1	0	0	–1	0
2	5	1	2	4	5	10
5	12	100	500	2500	1200	6000
7	20	1	7	49	20	140
14	36	103	509	2553	1224	6150

The normal equations in this case are

$$103a_0 + 509a_1 = 1224 \quad \text{(i)}$$

and

$$509a_0 + 2553a_1 = 6150. \quad \text{(ii)}$$

Solving the preceding equations, we obtain

$$a_0 = -1.41258 \quad \text{and} \quad a_1 = 2.69056.$$

The required 'linear least squares approximation' is therefore given by

$$y = -1.41258 + 2.69056x,$$

and the value of $y(5) = 12.0402$.
It follows that the approximation becomes better when the weight is increased.

4.3.2 Nonlinear Weighted Least Squares Approximation

We now consider the least squares approximation of a set of m data points (x_i, y_i), $i = 1, 2, \ldots, m$, by a polynomial of degree $n < m$. Let

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \tag{4.33}$$

be fitted to the given data points. We then have

$$S(a_0, a_1, \ldots, a_n) = \sum_{i=1}^{m} W_i \, [y_i - (a_0 + a_1 x_i + \cdots + a_n x_i^n)]^2. \tag{4.34}$$

If a minimum occurs at $(a_0, a_1, \ldots, a_n)$, then we have

$$\frac{\partial S}{\partial a_0} = \frac{\partial S}{\partial a_1} = \frac{\partial S}{\partial a_2} = \cdots = \frac{\partial S}{\partial a_n} = 0. \tag{4.35}$$

These conditions yield the normal equations

$$\left.\begin{aligned}
a_0 \sum_{i=1}^{m} W_i + a_1 \sum_{i=1}^{m} W_i x_i + \cdots + a_n \sum_{i=1}^{m} W_i x_i^n &= \sum_{i=1}^{m} W_i y_i \\
a_0 \sum\nolimits_{i=1}^{m} W_i x_i + a_1 \sum_{i=1}^{m} W_i x_i^2 + \cdots + a_n \sum_{i=1}^{m} W_i x_i^{n+1} &= \sum_{i=1}^{m} W_i x_i y_i \\
&\vdots \\
a_0 \sum_{i=1}^{m} W_i x_i^n + a_1 \sum_{i=1}^{m} W_i x_i^{n+1} + \cdots + a_n \sum_{i=1}^{m} W_i x_i^{2n} &= \sum_{i=1}^{m} W_i x_i^n y_i.
\end{aligned}\right\} \tag{4.36}$$

Equations (4.36) are $(n + 1)$ equations in $(n + 1)$ unknowns $a_0, a_1, \ldots, a_n$. If the x_i are distinct with $n < m$, then the equations possess a 'unique' solution.

4.4 METHOD OF LEAST SQUARES FOR CONTINUOUS FUNCTIONS

In the previous sections, we considered the least squares approximations of discrete data. We shall, in the present section, discuss the least squares approximation of a continuous function on $[a, b]$. The summations in the normal equations are now replaced by definite integrals.

Let

$$y(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \tag{4.37}$$

be chosen to minimize

$$S(a_0, a_1, \ldots, a_n) = \int_a^b W(x)\,[y(x) - (a_0 + a_1 x + \cdots + a_n x^n)]^2\, dx. \tag{4.38}$$

The necessary conditions for a minimum are given by

$$\frac{\partial S}{\partial a_0} = \frac{\partial S}{\partial a_1} = \cdots = \frac{\partial S}{\partial a_n} = 0, \tag{4.39}$$

which yield

$$\left.\begin{aligned}
-2\int_a^b W(x)\,[y(x) - (a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n)]\, dx &= 0\\
-2\int_a^b W(x)\,[y(x) - (a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n)]\, x\, dx &= 0\\
-2\int_a^b W(x)\,[y(x) - (a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n)]\, x^2\, dx &= 0\\
&\ \vdots\\
-2\int_a^b W(x)\,[y(x) - (a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n)]\, x^n\, dx &= 0.
\end{aligned}\right\} \tag{4.40}$$

Rearrangement of terms in Eq. (4.40) gives the system

$$\left.\begin{aligned}
a_0\int_a^b W(x)\,dx + a_1\int_a^b xW(x)\,dx + \cdots + a_n\int_a^b x^n W(x)\,dx &= \int_a^b W(x)y(x)\,dx\\
a_0\int_a^b xW(x)\,dx + a_1\int_a^b x^2 W(x)\,dx + \cdots + a_n\int_a^b x^{n+1} W(x)\,dx &= \int_a^b x\,W(x)y(x)\,dx\\
&\ \vdots\\
a_0\int_a^b x^n W(x)\,dx + a_1\int_a^b x^{n+1} W(x)\,dx + \cdots + a_n\int_a^b x^{2n} W(x)\,dx &= \int_a^b x^n W(x)y(x)\,dx.
\end{aligned}\right\} \tag{4.41}$$

The system in Eq. (4.41) comprises $(n + 1)$ normal equations in $(n + 1)$ unknowns, viz. $a_0, a_1, a_2, \ldots, a_n$ and they always possess a 'unique' solution.

Example 4.11 Construct a least squares quadratic approximation to the function $y(x) = \sin x$ on $[0, \pi/2]$ with respect to the weight function $W(x) = 1$.

Let

$$y = a_0 + a_1 x + a_2 x^2 \tag{i}$$

be the required quadratic approximation. Then using Eq. (4.41), we obtain the system

$$\left.\begin{aligned}
a_0 \int_0^{\pi/2} dx + a_1 \int_0^{\pi/2} x\,dx + a_2 \int_0^{\pi/2} x^2 dx &= \int_0^{\pi/2} \sin x\,dx \\
a_0 \int_0^{\pi/2} x\,dx + a_1 \int_0^{\pi/2} x^2\,dx + a_2 \int_0^{\pi/2} x^3 dx &= \int_0^{\pi/2} x \sin x\,dx \\
a_0 \int_0^{\pi/2} x^2\,dx + a_1 \int_0^{\pi/2} x^3\,dx + a_2 \int_0^{\pi/2} x^4 dx &= \int_0^{\pi/2} x^2 \sin x\,dx.
\end{aligned}\right\} \tag{ii}$$

Simplifying Eq. (ii), we obtain

$$a_0 \frac{\pi}{2} + a_1 \frac{\pi^2}{8} + a_2 \frac{\pi^3}{24} = 1$$

$$a_0 \frac{\pi^2}{8} + a_1 \frac{\pi^3}{24} + a_2 \frac{\pi^4}{64} = 1$$

$$a_0 \frac{\pi^3}{24} + a_1 \frac{\pi^4}{64} + a_2 \frac{\pi^5}{160} = 2\left(\frac{\pi}{2} - 1\right),$$

whose solution is

$$\left.\begin{aligned}
a_0 &= \frac{18}{\pi} + \frac{96}{\pi^2} - \frac{480}{\pi^3} \\
a_1 &= -\frac{144}{\pi^2} - \frac{1344}{\pi^3} + \frac{5760}{\pi^4} \\
a_2 &= \frac{240}{\pi^3} + \frac{2880}{\pi^4} - \frac{11520}{\pi^5}.
\end{aligned}\right\} \tag{iii}$$

The required quadratic approximation to $y = \sin x$ on $[0, \pi/2]$ is then given by (i) and (iii),
As a check, we obtain, at $x = \pi/4$,

$$\sin x \approx -\frac{3}{\pi} - \frac{60}{\pi^2} + \frac{240}{\pi^3} = 0.706167587.$$

The true value of $\sin(\pi/4) = 0.707106781$, so that the error in the preceding solution is 0.000939194.

4.4.1 Orthogonal Polynomials

In the previous section, we have seen that the method of determining a least square approximation to a continuous function gives satisfactory results. However, this method possesses the disadvantage of solving a large linear system of equations. Besides, such a system may exhibit a peculiar tendency called *ill-conditioning*, which means that small change in any of its parameters introduces large errors in the solution—the degree of *ill-conditioning* increasing with the order of the system. Hence, alternative methods of solving the aforesaid least-squares problem have gained importance, and of these the method that employs 'orthogonal polynomials' is currently in use. This method possessess the great advantage that it does not require a linear system to be solved and is described below.

We choose the approximation in the form:

$$Y(x) = a_0 f_0(x) + a_1 f_1(x) + \cdots + a_n f_n(x), \tag{4.42}$$

where $f_j(x)$ is a polynomial in x of degree j.

Then, we write

$$S(a_0, a_1, \ldots, a_n) = \int_a^b W(x)\{y(x) - [a_0 f_0(x) + a_1 f_1(x) + \cdots + a_n f_n(x)]\}^2 \, dx. \tag{4.43}$$

For S to be minimum, we must have

$$\left.\begin{aligned}
\frac{\partial S}{\partial a_0} &= 0 = -2\int_a^b W(x)\{y(x) - [a_0 f_0(x) + a_1 f_1(x) + \cdots + a_n f_n(x)]\} f_0(x)\, dx \\
\frac{\partial S}{\partial a_1} &= 0 = -2\int_a^b W(x)\{y(x) - [a_0 f_0(x) + a_1 f_1(x) + \cdots + a_n f_n(x)]\} f_1(x)\, dx \\
&\vdots \\
\frac{\partial S}{\partial a_n} &= 0 = -2\int_a^b W(x)\{y(x) - [a_0 f_0(x) + a_1 f_1(x) + \cdots + a_n f_n(x)]\} f_n(x)\, dx
\end{aligned}\right\} \tag{4.44}$$