Summary of last session

-Chapter 6. Approximating Function

- **6.0 Introduction**

- **6.1. Polynomial Interpolation:**

  method of undetermined coefficents,

  the error of the interpolation,

  lagrange interpolation method,

- **6.2. Polynomial Interpolation:**

  Newton interpolation (divided difference),

  difference and the interpolation for even spaced nodes

6.4 Best Approximation – Least Squares Method

First, let us look at an example.

Hooke's law states that when a force is applied to a spring constructed of uniform material, the length of the spring is a linear function of that force. We can write the linear function as $F(l)=k(l-E)$, where $F(l)$ represents the force required to stretch the spring l units, the constant E is the length of the spring with no force applied, and the k is the spring constant.

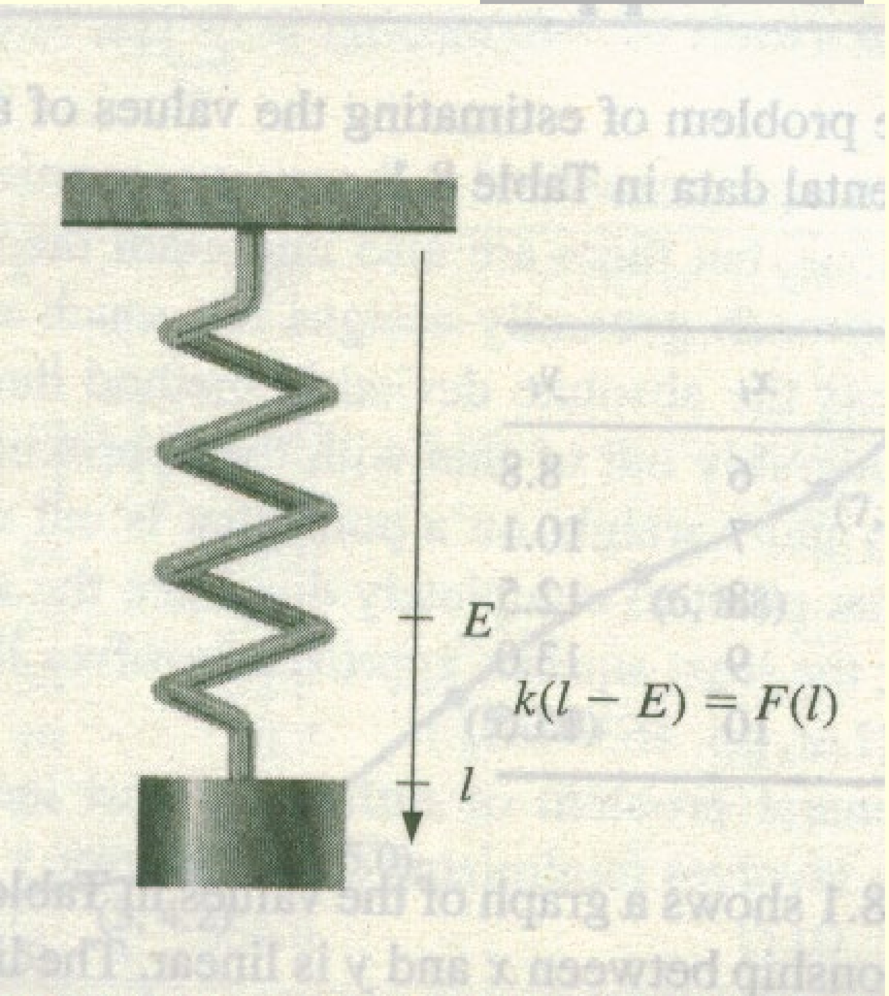

Figure 6.1

6.4 Best Approximation – Least Squares Method

Suppose we want to determine the spring constant for a spring that has initial length 5.3cm. We apply forces 2, 4, and 6N to it and find that its length increases to 7.0, 9.4, and 12.3cm, respectively. A quick examination shows that the points (0, 5.3), (2, 7.0), (4, 9.4) and (6, 12.3) do not quite lie in a straight line. Although we could simply use one random pair of these data points to approximate the spring constant, *it would seem more reasonable to find the line that best approximates all the data points to determine the constant.* This type of approximation will be considered in this section.

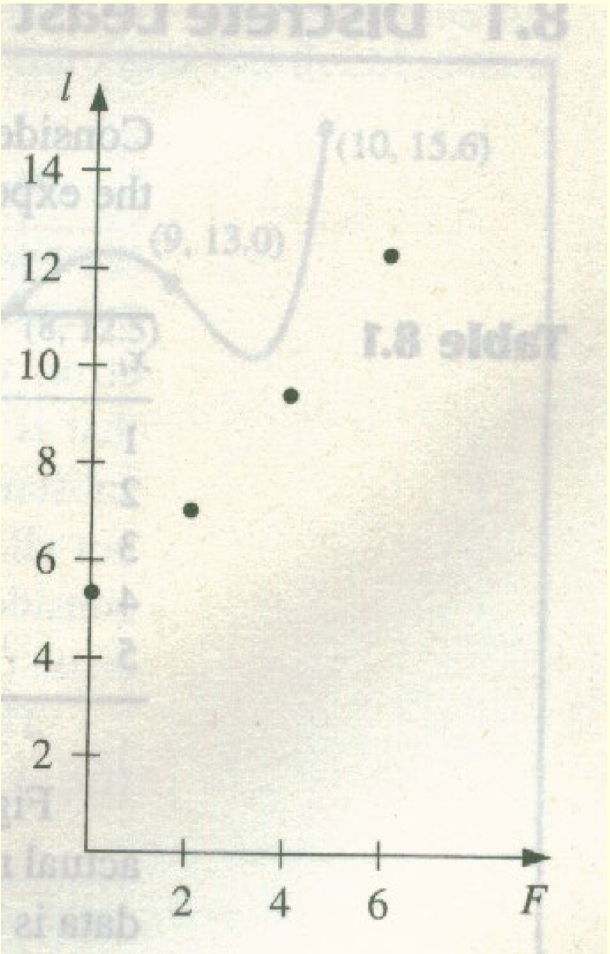

Figure 6.2

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

Consider the problem of estimating the values of a function at non-tabulated points, given in the experimental data in the table.

From the graph, it appears that the actual relationship between x and y is linear. The likely reason that no line precisely fits the data is because of errors in the data. *So it is unreasonable to require that the approximating function agree exactly with the data.*

x_i	y_i
1	1.3
2	3.5
3	4.2
4	5.0
5	7.0
6	8.8
7	10.1
8	12.5
9	13.0
10	15.6

Table 6.1

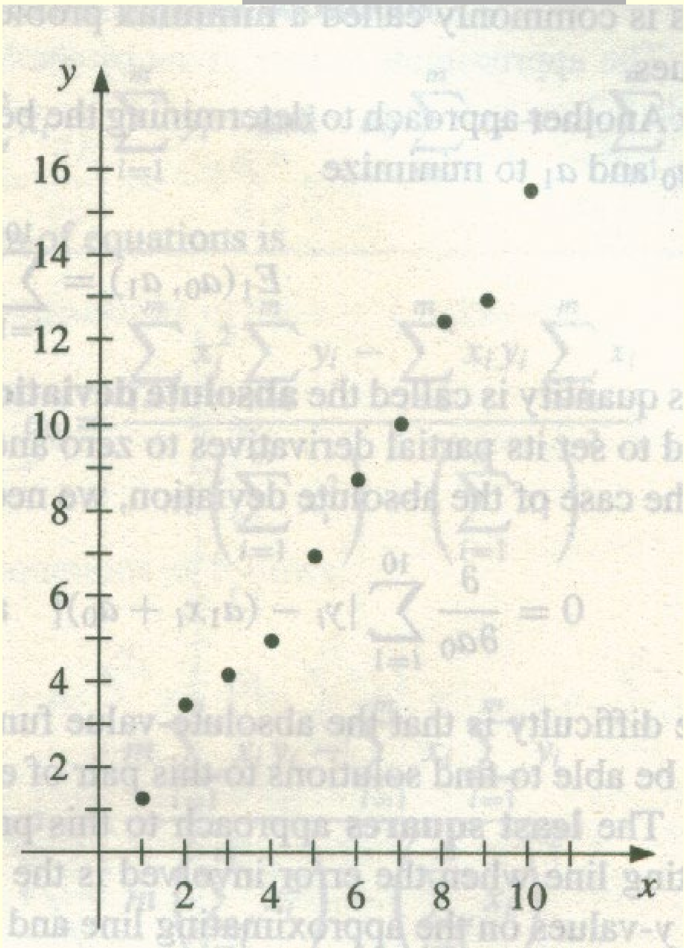

Figure 6.3

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

In fact, such a function would introduce oscillations that were not originally present. For example, the **ninth degree** interpolating polynomial on the data shown in Figure 6.4 is clearly a poor predictor of information between a number of the data points. A better approach would be to find the "best" (in some sense) approximating line, even if it does not agree precisely with the data at any point.

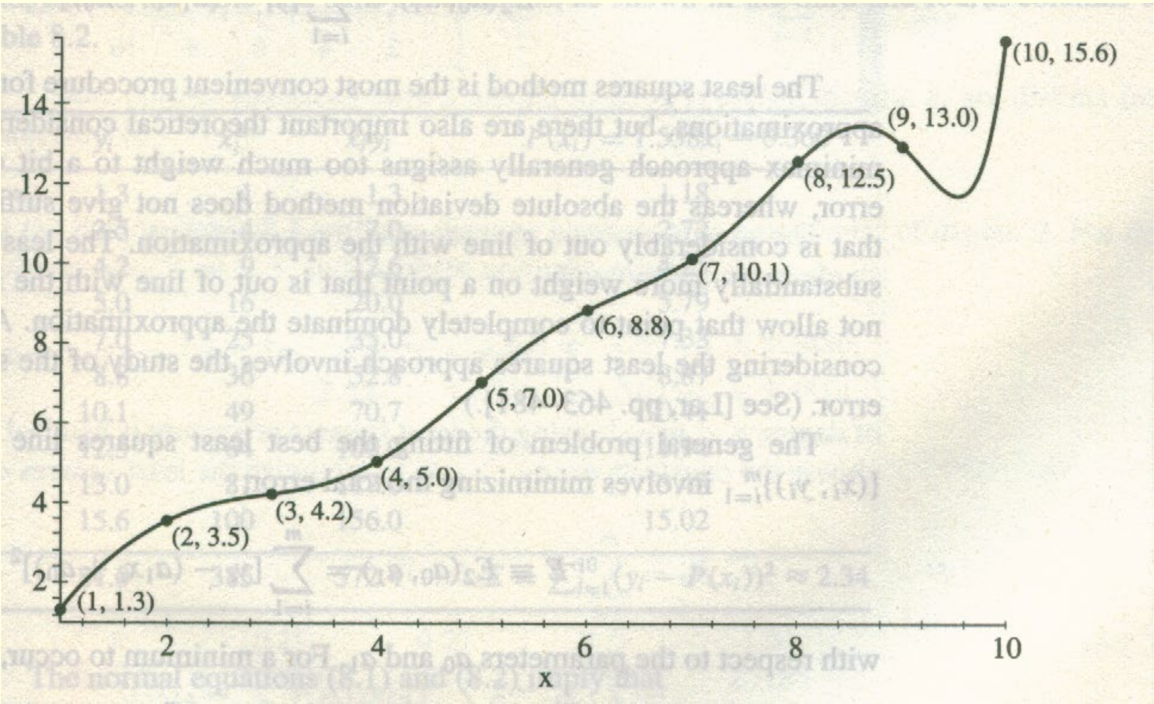

Figure 6.4

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

Let $a_1x_i+a_0$ denote the ith value on the approximating line and y_i be the ith given y-value. The problem of finding the equation of the best linear approximation in the absolute sense requires that values of a_0 and a_1 be found to minimize

$$E_\infty(a_0,a_1)=\max_{1\le i\le 10}\left\{\ |y_i-(a_1x_i+a_0)|\right\} \qquad (1)$$

This is commonly called a **minimax** problem. Another approach to determine the best linear approximation involves finding values of a_0 and a_1 to minimize

$$E_1(a_0,a_1)=\sum_{i=1}^{10}|y_i-(a_1x_i+a_0)| \qquad (2)$$

This quantity is called the **absolute deviation**. To minimize a function of two variables, we need to set its patial derivatives to zero and simultaneously solve the resulting equations.

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

In the case of the absolute deviation, we need to find a_0 and a_1 with

$$\begin{cases} \dfrac{\partial}{\partial a_0}\displaystyle\sum_{i=1}^{10} |\, y_i - (a_1 x_i + a_0)\,| = 0 \\ \dfrac{\partial}{\partial a_1}\displaystyle\sum_{i=1}^{10} |\, y_i - (a_1 x_i + a_0)\,| = 0 \end{cases} \qquad (3)$$

The difficulty is that the absolute-value function is not differentiable at some points, and we may not be able to find solutions to equations. The **least squares approach** to this problem involves finding the best approximating line when the error involved is the sum of the squares of the differences between the y-values on the approximating line and the given y-values. Hence, constants a_0 and a_1 must be found that minimize the **least squares error**:

$$E_2(a_0, a_1) = \sum_{i=1}^{10} [y_i - (a_1 x_i + a_0)]^2 \qquad (4)$$

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

The least squares method is the most convenient procedure for determining best linear approximations, but there are also important theoretical considerations that favor it. The minimax approach generally assigns too much weight to a bit of data that is badly in error, whereas the absolute deviation method does not give sufficient weight to a point that is considerably out of line with the approximation. The least squares approach puts substantially more weight on a point that is out of line with the rest of the data but will not allow that point to completely dominate the approximation.

$$E_\infty(a_0, a_1) = \max_{1\le i\le 10}\{\ |y_i - (a_1 x_i + a_0)|\} \qquad (1)$$

$$E_1(a_0, a_1) = \sum_{i=1}^{10} |y_i - (a_1 x_i + a_0)| \qquad (2)$$

$$E_2(a_0, a_1) = \sum_{i=1}^{10} [y_i - (a_1 x_i + a_0)]^2 \qquad (4)$$

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

The general problem of fitting the best least squares line to a collection of data $\{(x_i, y_i)\}_{i=1}^{m}$ involves minimizing the total error,

$$E = E_2(a_0, a_1) = \sum_{i=1}^{m} [y_i - (a_1 x_i + a_0)]^2 \qquad (6)$$

with repect to a_0 and a_1. For a minimum to occur, we need

$$\begin{cases} \dfrac{\partial}{\partial a_0} \sum_{i=1}^{m} [y_i - (a_1 x_i + a_0)]^2 = 2\sum_{i=1}^{m} [y_i - a_1 x_i - a_0](-1) = 0 \\ \dfrac{\partial}{\partial a_1} \sum_{i=1}^{m} [y_i - (a_1 x_i + a_0)]^2 = 2\sum_{i=1}^{m} [y_i - a_1 x_i - a_0](-x_i) = 0 \end{cases} \qquad (7)$$

These equations can be simplified to the normal equations:

$$\begin{cases} a_0 \cdot m + a_1 \sum_{i=1}^{m} x_i = \sum_{i=1}^{m} y_i \\ a_0 \sum_{i=1}^{m} x_i + a_1 \sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} x_i y_i \end{cases} \qquad (8)$$

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

The solution to this system of equations is

$$a_0 = \frac{\sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} y_i - \sum_{i=1}^{m} x_i y_i \sum_{i=1}^{m} x_i}{m(\sum_{i=1}^{m} x_i^2) - (\sum_{i=1}^{m} x_i)^2} \qquad (9)$$

and

$$a_1 = \frac{m\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i}{m(\sum_{i=1}^{m} x_i^2) - (\sum_{i=1}^{m} x_i)^2} \qquad (10)$$

$$a_0 = \frac{\sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} y_i - \sum_{i=1}^{m} x_i y_i \sum_{i=1}^{m} x_i}{m(\sum_{i=1}^{m} x_i^2) - (\sum_{i=1}^{m} x_i)^2} \quad (9) \qquad a_1 = \frac{m\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i}{m(\sum_{i=1}^{m} x_i^2) - (\sum_{i=1}^{m} x_i)^2} \quad (10)$$

Example 1. Consider the data presented in Table 1 and find the least squares line approximating this data. We extend the table to sum the columns, as shown in the third and fourth columns of the table.

x_i	y_i	x_i^2	$x_i y_i$	$P(x_i)=1.538x_i-0.360$
1	1.3	1	1.3	1.18
2	3.5	4	7.0	2.72
3	4.2	9	12.6	4.25
4	5.0	16	20.0	5.79
5	7.0	25	35.0	7.33
6	8.8	36	52.8	8.87
7	10.1	49	70.7	10.41
8	12.5	64	100.0	11.94
9	13.0	81	117.0	13.48
10	15.6	100	156.0	15.02
55	81.0	385	572.4	$E=\sum_{i=1}^{10}(y_i - P(x_i))^2 \approx 2.34$

Table 6.2

$$a_0 = \frac{385(81) - 55(572.4)}{10(385) - (55)^2} = -0.360$$

$$a_1 = \frac{10(572.4) - 55(81)}{10(385) - (55)^2} = 1.538$$

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

So the least squares approximation is

$$P(x) = 1.538x - 0.360$$

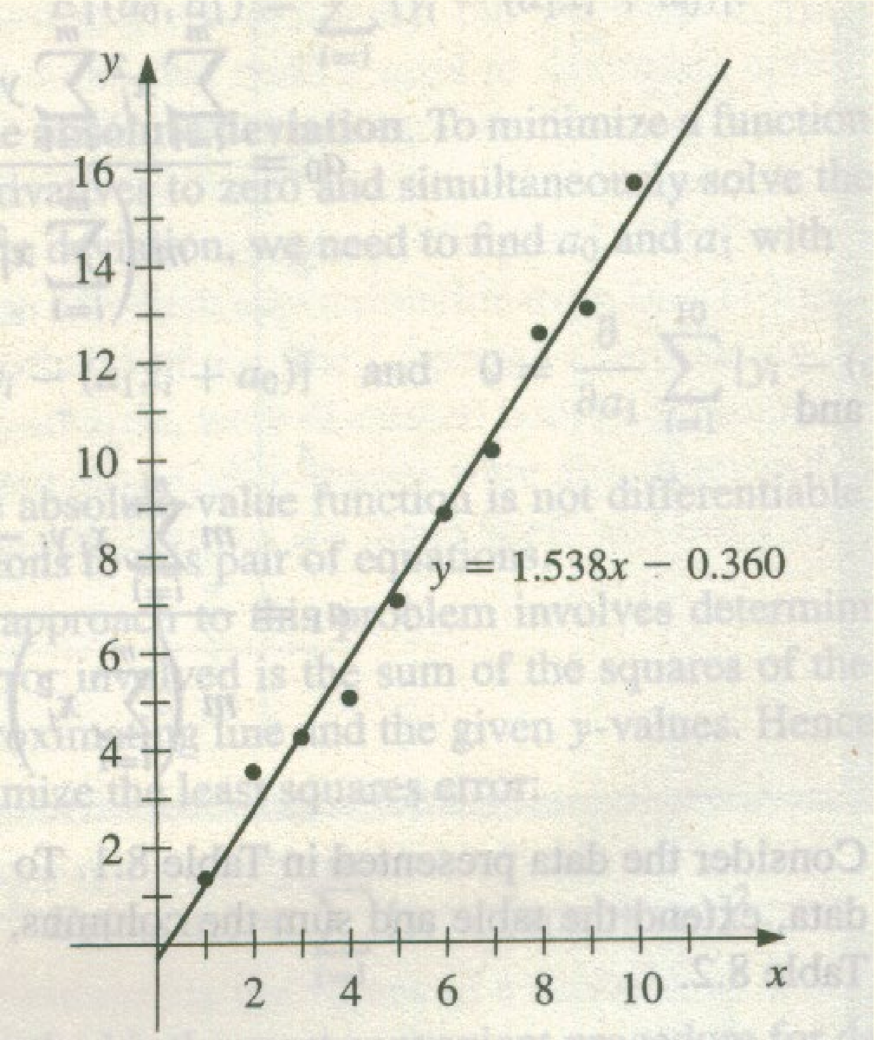

Figure 6.5

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

The general problem of approximating a set of data, $\{(x_i, y_i),\ i = 1,2,....,m\}$, with an algebraic polynomial

$$P_n(x) = a_n x^n + a_{n\text{-}1} x^{n\text{-}1} + \ldots + a_1 x + a_0 = \sum_{j=0}^{n} a_j x^j$$

of degree $n < m\text{-}1$, using the least squares procedure is handled in a similar manner. We choose $a_0, a_1, \ ..., a_n$ to minimize the least squares error

$$E_2 = \sum_{i=1}^{m} (y_\mathrm{i} \text{-} P_n(x_i))^2 \qquad \left(\because (\sum_{j=0}^{n} a_j x_i^j)^2 = \sum_{k=0}^{n} a_k x_i^k (\sum_{j=0}^{n} a_j x_i^j) \right.$$

$$= \sum_{i=1}^{m} y_i^2 - 2\sum_{i=1}^{m} P_n(x_i) y_\mathrm{i} + \sum_{i=1}^{m} (P_n(x_i))^2 \qquad \left. = \sum_{j=0}^{n}\sum_{k=0}^{n} a_j a_k x_i^{j+k} \right)$$

$$= \sum_{i=1}^{m} y_i^2 - 2\sum_{i=1}^{m} (\sum_{j=0}^{n} a_j x_i^j) y_\mathrm{i} + \sum_{i=1}^{m} (\sum_{j=0}^{n} a_j x_i^j)^2$$

$$= \sum_{i=1}^{m} y_i^2 - 2\sum_{j=0}^{n} (a_j (\sum_{i=1}^{m} y_\mathrm{i} x_i^j)) + \sum_{j=0}^{n}\sum_{k=0}^{n} a_j a_k (\sum_{i=1}^{m} x_i^{j+k}) \qquad (11)$$

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

As in the linear case, for E to be minimized it is necessary that $\partial E/\partial a_j = 0$, for each $j = 0, 1, \ldots, n$. Thus, for each j,

$$\frac{\partial E}{\partial a_j} = -2\sum_{i=1}^{m} y_i x_i^j + 2\sum_{k=0}^{n} a_k \sum_{i=1}^{m} x_i^{j+k} = 0 \qquad (12)$$

This gives $n+1$ normal equations in the $n+1$ unknows a_j.

$$\sum_{k=0}^{n} a_k \sum_{i=1}^{m} x_i^{j+k} = \sum_{i=1}^{m} y_i x_i^j, \quad \text{for each } j = 0,1,\ldots,n \qquad (13)$$

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

It is helpful to write the equation as follows:

$$j = 0: \qquad a_0 \sum_{i=1}^{m} x_i^0 + a_1 \sum_{i=1}^{m} x_i^1 + a_2 \sum_{i=1}^{m} x_i^2 + \ldots + a_n \sum_{i=1}^{m} x_i^n = \sum_{i=1}^{m} y_i x_i^0$$

$$j = 1: \qquad a_0 \sum_{i=1}^{m} x_i^1 + a_1 \sum_{i=1}^{m} x_i^2 + a_2 \sum_{i=1}^{m} x_i^3 + \ldots + a_n \sum_{i=1}^{m} x_i^{n+1} = \sum_{i=1}^{m} y_i x_i^1$$

$$\vdots$$

$$j = n: \qquad a_0 \sum_{i=1}^{m} x_i^n + a_1 \sum_{i=1}^{m} x_i^{n+1} + a_2 \sum_{i=1}^{m} x_i^{n+2} + \ldots + a_n \sum_{i=1}^{m} x_i^{2n} = \sum_{i=1}^{m} y_i x_i^n$$

These normal equations have a unique solution provided that the x_i are distinct.

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

Example 2 Fit the data in Table 3 with the discrete least squares polynomial of degree 2.

Solution : For this problem, $n=2$, $m=5$, and the three normal equations are

$$5a_0 + 2.5a_1 + 1.875a_2 = 8.7680$$
$$2.5a_0 + 1.875a_1 + 1.5625a_2 = 5.4514$$
$$1.875a_0 + 1.5625a_1 + 1.3828a_2 = 4.4015.$$

Table 6.3

i	1	2	3	4	5
x_i	0.00	0.25	0.50	0.75	1.00
y_i	1.0000	1.2840	1.6487	2.1170	2.7183

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

We solve this equation system and get:

$a_0 = 1.0051,\ \ a_1 = 0.86468,\ \ a_2 = 0.84316$

Thus the least squares polynomial of degree 2 fitting the preceding data is

$$P_2(x) = 1.0051 + 0.86468x + 0.84316x^2$$

The graph is shown in Figure 6.6. At the given values of x_i, we have the approximations shown in Table 6.4.

The total error:

$$E_2 = \sum_{i=1}^{5} (y_i - P(x_i))^2 = 2.74 \times 10^{-4}$$

is the least that can be obtained by using a polynomial of degree at most 2.

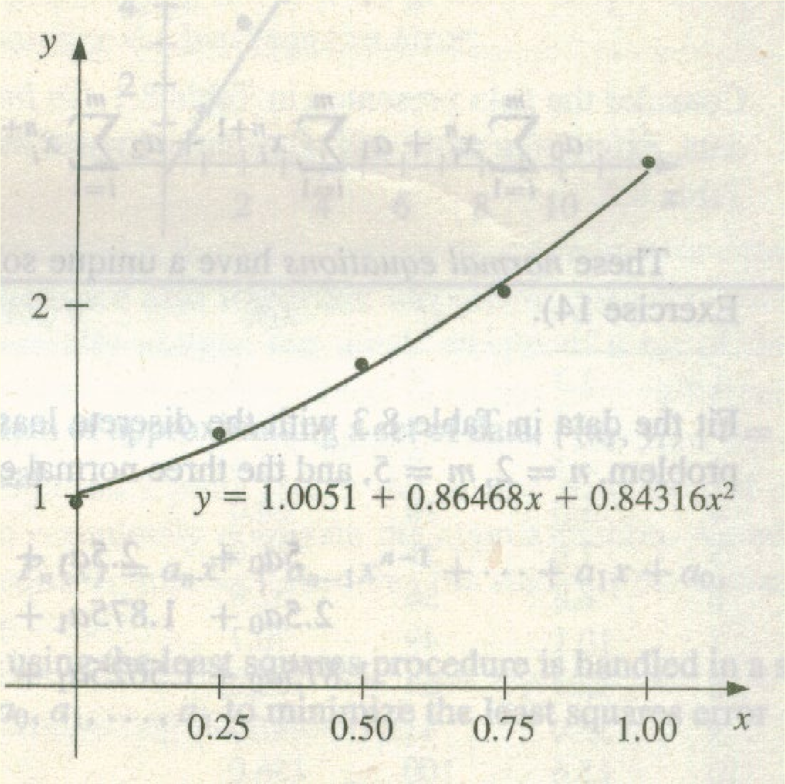

Figure 6.6

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

Table 6.4

i	1	2	3	4	5
x_i	0	0.25	0.5	0.75	1.00
y_i	1.0000	1.2840	1.6487	2.1170	2.7183
$P(x_i)$	1.0051	1.2740	1.6482	2.1279	2.7129
y_i-$P(x_i)$	-0.0051	0.0100	0.0004	-0.0109	0.0054

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

Occasionally it is appropriate to assume that the data are exponentially related. This requires the approximating function to be of the form

$$y = be^{ax} \qquad (14)$$

or

$$y = bx^{a} \qquad (15)$$

for some constants *a* and *b*. The difficulty with applying the least squares procedure in a situation of this type comes from attempting to minimize

$$E = \sum_{i=1}^{m} (y_{\mathrm{i}} - be^{ax_i})^2, \quad \text{in the case of (14)}$$

or

$$E = \sum_{i=1}^{m} (y_{\mathrm{i}} - bx_i^{a})^2, \quad \text{in the case of (15)}$$

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

The normal equations associated with these procedures are obtained from

$$\begin{cases} \dfrac{\partial E}{\partial b} = 2\sum_{i=1}^{m}(y_{\mathrm{i}} - be^{ax_i})(-e^{ax_i}) = 0 \\ \dfrac{\partial E}{\partial a} = 2\sum_{i=1}^{m}(y_{\mathrm{i}} - be^{ax_i})(-bx_i e^{ax_i}) = 0 \end{cases} \quad \text{in the case of (14)}$$

or

$$\begin{cases} \dfrac{\partial E}{\partial b} = 2\sum_{i=1}^{m}(y_{\mathrm{i}} - bx_i^a)(-x_i^a) = 0 \\ \dfrac{\partial E}{\partial a} = 2\sum_{i=1}^{m}(y_{\mathrm{i}} - bx_i^a)(-b(\ln x_i)x_i^a) = 0 \end{cases} \quad \text{in the case of (15)}$$

No exact solution to either of these systems in a and b can generally be found.

$$a_0 = \frac{\sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} y_i - \sum_{i=1}^{m} x_i y_i \sum_{i=1}^{m} x_i}{m(\sum_{i=1}^{m} x_i^2) - (\sum_{i=1}^{m} x_i)^2} \quad (9) \quad a_1 = \frac{m\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i}{m(\sum_{i=1}^{m} x_i^2) - (\sum_{i=1}^{m} x_i)^2} \quad (10)$$

The method that is commonly used when the data are suspected to be exponentially related is to consider the logarithm of the approximating equation

$\ln y = \ln b + ax,$ in the case of (14) $\quad y = be^{ax} \quad (14)$

and

$\ln y = \ln b + a \ln x,$ in the case of (15) $\quad y = bx^{a} \quad (15)$

In either case, a linear problem now appears, and solutions for $\ln b$ and a can be obtained by appropriately modifying the normal equations (9) and (10). However, the approximation obtained in this manner is not the least squares approximation for the original poblem, and this approximation can in some cases differ significantly from the least squares approximation to the original problem.

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

Example 3 Consider the collection of data in the first three columns of Table 5.

Table 6.5

i	x_i	y_i	$\ln y_i$	x_i^2	$x_i \ln y_i$
1	1.00	5.10	1.629	1.0000	1.629
2	1.25	5.79	1.756	1.5625	2.195
3	1.50	6.53	1.876	2.2500	2.814
4	1.75	7.45	2.008	3.0625	3.514
5	2.00	8.46	2.135	4.0000	4.270
	7.50		9.404	11.875	14.422

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

If x_i is graphed with $\ln y_i$, the data appear to have a linear relation, so it is reasonable to assume an approximation of the form

$$y = be^{ax} \quad \text{or} \quad \ln y = \ln b + ax$$

Extending the table and summing the appropriate columns gives the data in Table 5. Using the normal equtions (9) and (10), we have

$$a = \frac{(5)(14.422) - (7.5)(9.404)}{(5)(11.875) - (7.5)^2} = 0.5056$$

and

$$\ln b = \frac{(11.875)(9.404) - (14.422)(7.5)}{(5)(11.875) - (7.5)^2} = 1.122$$

Since $b = e^{1.122} = 3.071$, the approximation assumes the form

$$y = 3.071\, e^{0.5056x},$$

which at the data points, gives the values in Table 6 (see Figure 6.7).

6.4 Best Approximation – Least Squares Method

- Discrete least squares approximation

Table 6.6

i	x_i	y_i	$3.071\,e^{0.5056x_i}$
1	1.00	5.10	5.09
2	1.25	5.79	5.78
3	1.50	6.53	6.56
4	1.75	7.45	7.44
5	2.00	8.46	8.44

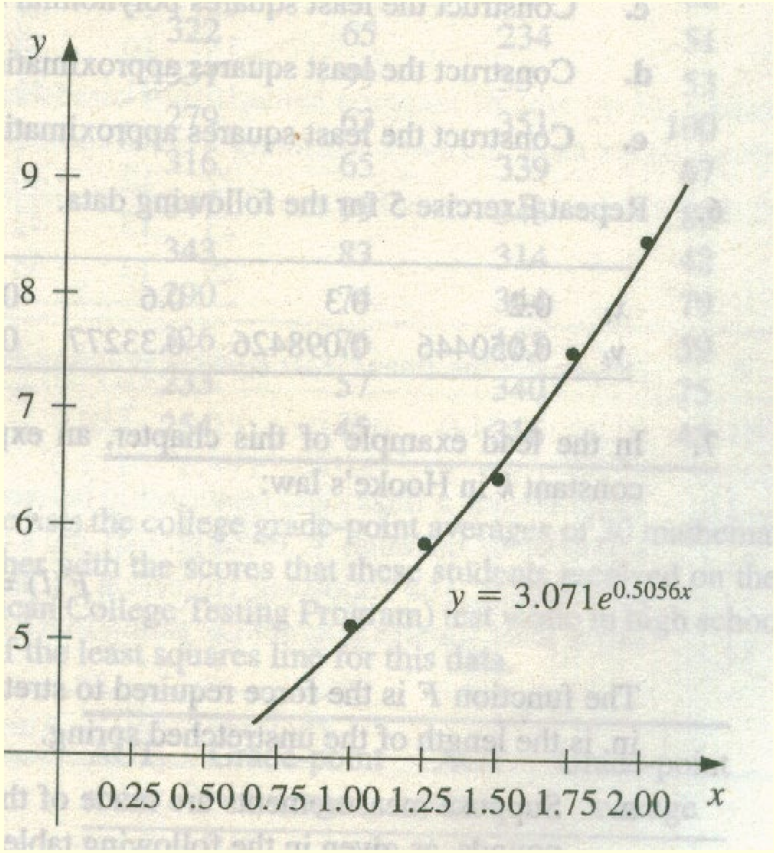

Figure 6.7

6.4 Best Approximation – Least Squares Method

- Orthogonal polynomials and least squares approximation

Up to now we discussed the problem of least squares approximation to fit a collection of data. The other approximation problem concerns the approximation of function.

Suppose $f(x) \in C[a,b]$ and that a polynomial $P_{\mathrm{n}}(x)$ of degree at most n is required that will minimize the error

$$\int_a^b [f(x) - P_{\mathrm{n}}(x)]^2 dx$$

To determine a least squares approximating polynomial; that is, a polynomial to minimize this expression, let

$$P_{\mathrm{n}}(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = \sum_{k=0}^{n} a_k x^k \qquad (16)$$

and define, as shown in Figure 6.8

$$E \equiv E(a_0, a_1, \cdots, a_n) = \int_a^b (f(x) - \sum_{k=0}^{n} a_k x^k)^2 dx \qquad (17)$$

6.4 Best Approximation – Least Squares Method

- Orthogonal polynomials and least squares approximation

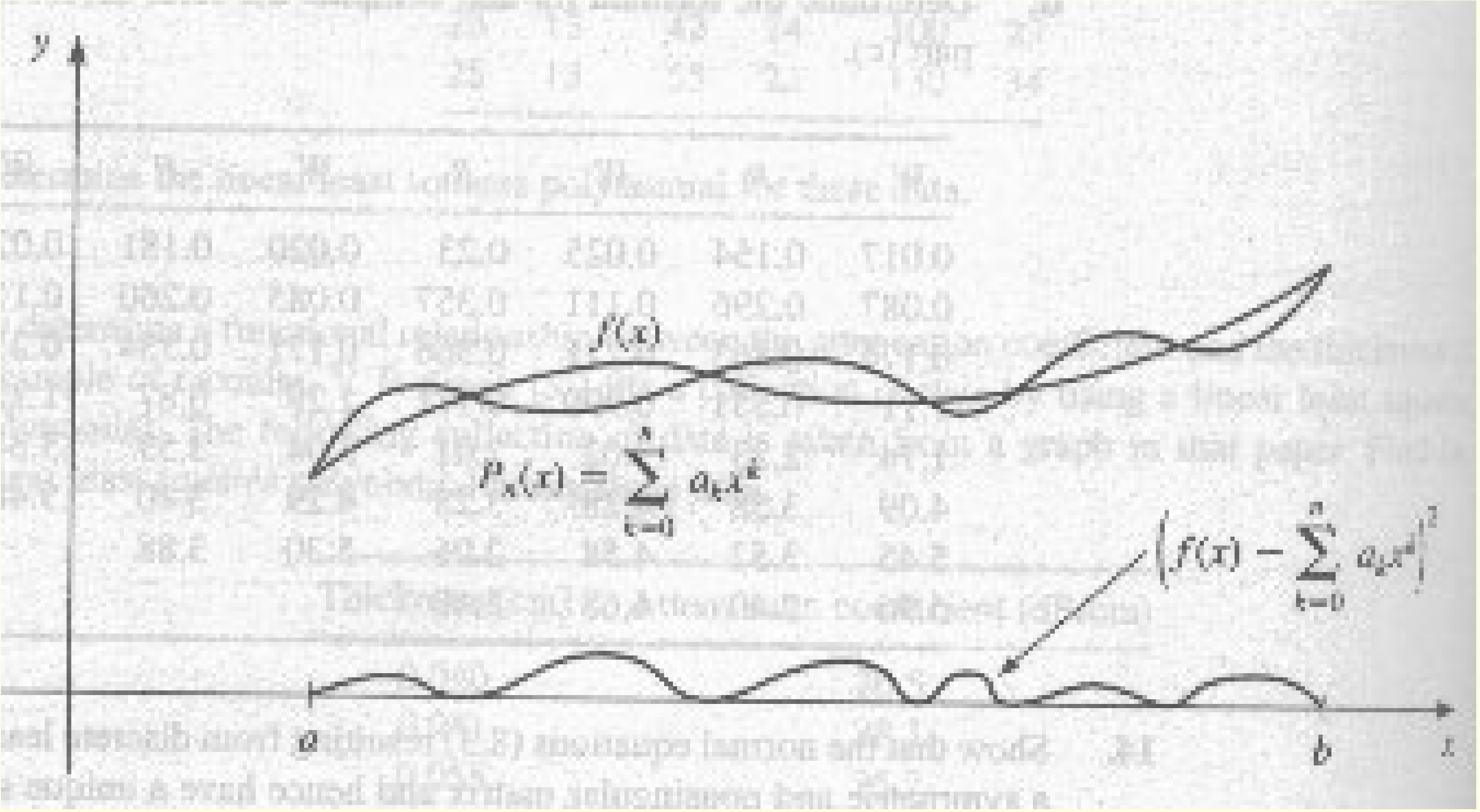

Figure 6.8

6.4 Best Approximation – Least Squares Method

- Orthogonal polynomials and least squares approximation

The problem is to find real coefficients $a_0, a_1, \cdots, a_n$ that will minimize E. A necessary condition for the numbers $a_0, a_1, \cdots, a_n$ to minimize E is that

$$\frac{\partial E}{\partial a_j} = 0, \quad \text{for each } j = 0, 1, \ldots, n.$$

Since $$E = \int_a^b (f(x))^2\,dx - 2\sum_{k=0}^{n} a_k \int_a^b x^k f(x)\,dx + \int_a^b \Big(\sum_{k=0}^{n} a_k x^k\Big)^2 dx$$

we have

$$\frac{\partial E}{\partial a_j} = -2\int_a^b x^j f(x)\,dx + 2\sum_{k=0}^{n} a_k \int_a^b x^{j+k}\,dx$$

Hence, to find $P_n(x)$, the $(n+1)$ linear normal equations

$$\sum_{k=0}^{n} a_k \int_a^b x^{j+k}\,dx = \int_a^b x^j f(x)\,dx, \quad \text{for each } j = 0, 1, \ldots, n \qquad (17)$$

must be solved for the $(n+1)$ unknowns a_j.

$$\left(\text{Discrete: } \sum_{k=0}^{n} a_k \sum_{i=1}^{m} x_i^{j+k} = \sum_{i=1}^{m} y_i x_i^j, \quad \text{for each } j = 0,1,\ldots,n \qquad (13)\right)$$

$$\left(\sum_{k=0}^{n} a_k \int_a^b x^{j+k} dx = \int_a^b x^j f(x) dx, \quad \text{for each } j = 0, 1, \ldots, n \qquad (17)\right)$$

Example 4 Find the least squares appoximating polynomial of degree 2 for the function $f(x) = \sin \pi x$ on the interval [0, 1]. The normal equations for $P_2(x) = a_2 x^2 + a_1 x + a_0$ are

$j = 0$
$$a_0 \int_0^1 1\, dx + a_1 \int_0^1 x\, dx + a_2 \int_0^1 x^2\, dx = \int_0^1 \sin \pi x\, dx,$$

$j = 1$
$$a_0 \int_0^1 x\, dx + a_1 \int_0^1 x^2\, dx + a_2 \int_0^1 x^3\, dx = \int_0^1 x \sin \pi x\, dx,$$

$j = 2$
$$a_0 \int_0^1 x^2\, dx + a_1 \int_0^1 x^3\, dx + a_2 \int_0^1 x^4\, dx = \int_0^1 x^2 \sin \pi x\, dx.$$

Performing the integration yields

$$a_0 + \frac{1}{2} a_1 + \frac{1}{3} a_2 = \frac{2}{\pi},$$

$$\frac{1}{2} a_0 + \frac{1}{3} a_1 + \frac{1}{4} a_2 = \frac{1}{\pi},$$

$$\frac{1}{3} a_0 + \frac{1}{4} a_1 + \frac{1}{5} a_2 = \frac{\pi^2 - 4}{\pi^3}$$

6.4 Best Approximation – Least Squares Method

- Orthogonal polynomials and least squares approximation

Solving these three equations, we obtain

$$a_0 = \frac{12\pi^2 - 120}{\pi^3} \approx -0.050465,$$

$$a_1 = -a_2 = \frac{720\text{-}60\pi^2}{\pi^3} \approx 4.12251$$

Consequently, the least squares polynomial approximation of degree 2 for $f(x) = \sin \pi x$ on $[0, 1]$ is

$$P_2(x) = -4.12251x^2 + 4.12251x - 0.050465.$$

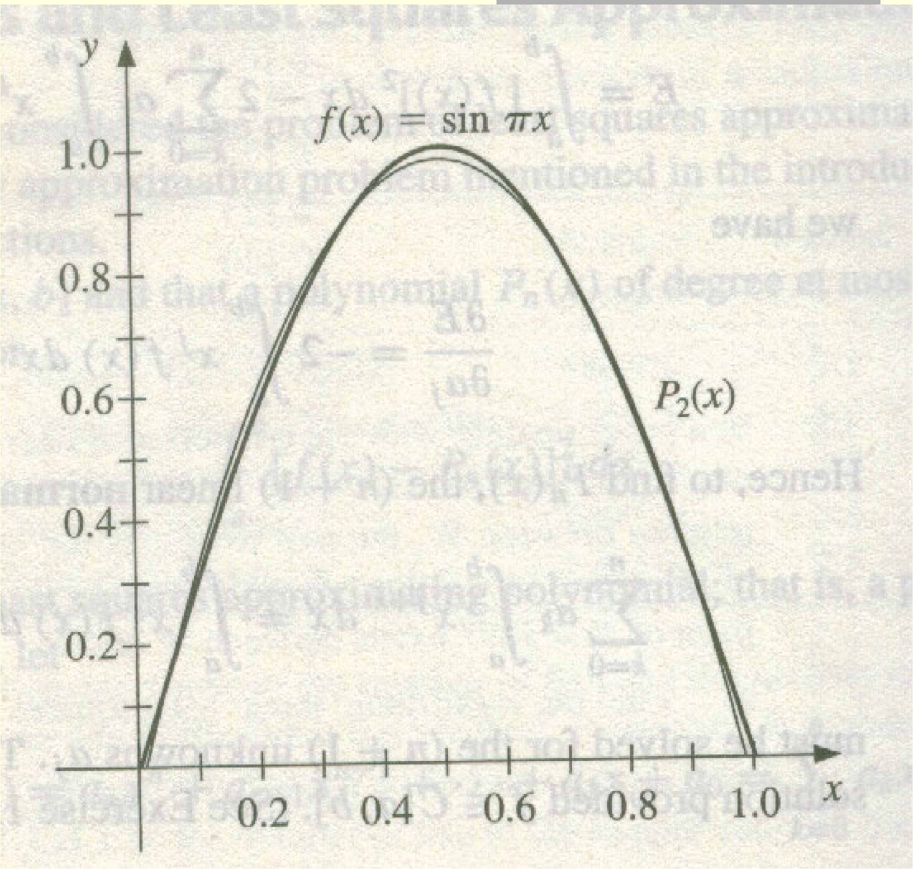

Figure 6.9

6.5 Orthogonal polynomials

- Orthogonal functions family and orthogonal polynomials

Definition 5 If $f(x), g(x) \in C[a,b]$, $\rho(x)$ is the weight function on $[a,b]$ such that

$$\left(f(x), g(x)\right)=\int_a^b \rho(x) f(x) g(x) dx=0, \qquad (2.1)$$

then we call that $f(x)$ and $g(x)$ are **orthogonal** with weight $\rho(x)$. If a family of functions $\varphi_0(x), \varphi_1(x), \cdots, \varphi_n(x), \cdots$ satisfy

$$\left(\varphi_j, \varphi_k\right)=\int_a^b \rho(x) \varphi_j(x) \varphi_k(x) dx=\begin{cases} 0, & j \neq k \\ A_k > 0, & j = k \end{cases} \qquad (2.2)$$

then $\left\{\varphi_k(x)\right\}$ is called an **orthogonal functions family** with weight $\rho(x)$. If $A_k \equiv 1$, then it is called a **standard orthogonal functions family**.

For example, the trigonometric functions family

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \cdots$$

is an orthogonal functions family on $\left[-\pi, \pi\right]$.

6.5 Orthogonal polynomials

- Orthogonal functions family and orthogonal polynomials

Definition 6 Assume $\varphi_n(x)$ is a polynomial of degree n on $[a,b]$ with the leading coefficient $a_n \neq 0$, $\rho(x)$ is the weight function on $[a,b]$, if $\{\varphi_n(x)\}_0^\infty$ satisfies (2.2), then we call that $\{\varphi_n(x)\}_0^\infty$ is orthogonal with weight $\rho(x)$ on $[a,b]$, $\varphi_n(x)$ is the orthogonal polynomial of degree n with weight $\rho(x)$ on $[a,b]$.

Given the interval $[a,b]$ and weight function $\rho(x)$, from $\{1, x, \cdots, x^n, \cdots\}$, we can construct the following sequence of orthogonal polynomials $\{\varphi_n(x)\}_0^\infty$:

$$\varphi_0(x) = 1, \quad \varphi_n(x) = x^n - \sum_{j=0}^{n-1} \frac{\left(x^n, \varphi_j(x)\right)}{\left(\varphi_j(x), \varphi_j(x)\right)} \varphi_j(x) \ \ (n = 1, 2, \cdots). \qquad (2.3)$$

One Important Property

$\left(\varphi_j(x), \varphi_k(x)\right) = 0,\ k \neq j,$ and $\varphi_k(x)$ is orthogonal to any polynomial of degree less than k.

6.5 Orthogonal polynomials

- Legendre polynomials and Chebyshev polynomials

Legendre Polynomials

When the interval is $[-1,1]$, $\rho(x)\equiv 1$, the orthogonal polynomials constructed from $\{1, x, \cdots, x^n, \cdots\}$ are called the **Legendre Polynomials**, which are represented by $P_0(x), P_1(x), \cdots, P_n(x), \cdots$.

$$P_0(x)=1, \quad P_n(x)=\frac{1}{2^n n!}\frac{d^n}{dx^n}\left\{\left(x^2-1\right)^n\right\} \ (n=1,2,\cdots).$$

We present the first several Legendre Polynomials as follows:

$P_0(x)=1,\ P_1(x)=x,\ P_2(x)=(3x^2-1)/2,\ P_3(x)=(5x^3-3x)/2,\ P_4(x)=(35x^4-30x^2+3)/8$

Chebyshev Polynomials

When the interval is $[-1,1]$, $\rho(x)=\dfrac{1}{\sqrt{1-x^2}}$, the orthogonal polynomials constructed from $\{1, x, \cdots, x^n, \cdots\}$ are called the **Chebyshev Polynomials**, which are represented by

$$T_n(x)=\cos(n\arccos x), \quad |x|\le 1.$$

Let $x=\cos\theta$, then $T_n(x)=\cos n\theta, 0\le\theta\le\pi$.

We present the first several Chebyshev Polynomials as follows:

$T_0(x)=1,\ T_1(x)=x,\ T_2(x)=2x^2-1,\ T_3(x)=4x^3-3x,\ T_4(x)=8x^4-8x^2+1$

Exercises

$Ex1.$ 1.Find the least squares polynomials of degree 1, 2 and 3 for the data in the following table. Compute the error E in each case.

x_i	0.00	0.15	0.31	0.50	0.60	0.75
y_i	1.000	1.004	1.031	1.117	1.223	1.422

$Ex2.$ 1.Find the least squares polynomial approximations of degree 1 and 2 on the interval [-1, 1] for the following functions, and compute the error E for each case.

(1) $f(x) = x^2 - 2x + 3$ (2) $f(x) = x^3$ (3) $f(x) = 1/(x+2)$

(4) $f(x) = e^x$ (5) $f(x) = \frac{1}{2}\cos x + \frac{1}{3}\sin 2x$