

where

$$a_{ij} = \sum_{k=1}^m x_k^{i+j-2}$$

and x_1, x_2, \dots, x_m are distinct with $n < m - 1$. Suppose A is singular and that $\mathbf{c} \neq \mathbf{0}$ is such that $\mathbf{c}' A \mathbf{c} = 0$. Show that the n th-degree polynomial whose coefficients are the coordinates of \mathbf{c} has more than n roots, and use this to establish a contradiction.]

8.2 Orthogonal Polynomials and Least Squares Approximation

The previous section considered the problem of least squares approximation to fit a collection of data. The other approximation problem mentioned in the introduction concerns the approximation of functions.

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

$$\int_a^b [f(x) - P_n(x)]^2 dx.$$

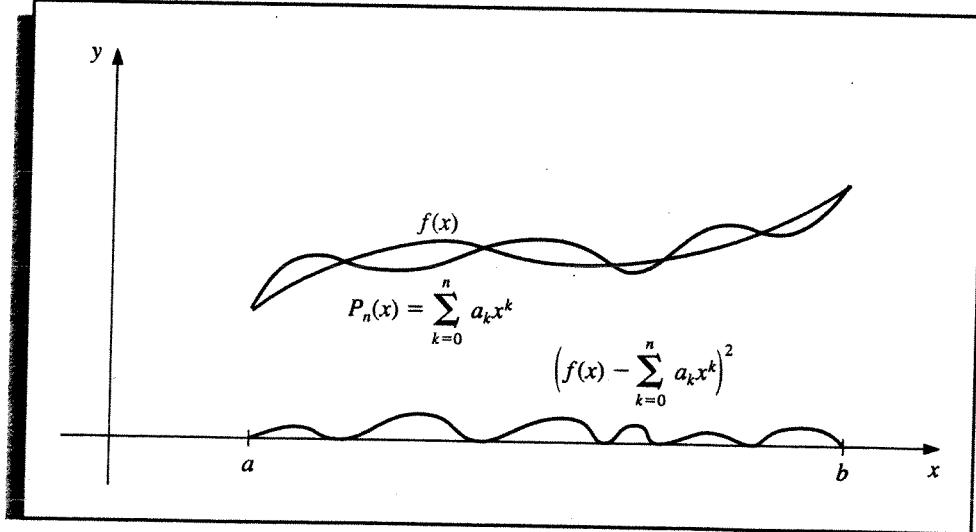
To determine a least squares approximating polynomial, that is, a polynomial to minimize this expression, let

$$P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = \sum_{k=0}^n a_k x^k,$$

and define, as shown in Figure 8.5,

$$E(a_0, a_1, \dots, a_n) = \int_a^b \left(f(x) - \sum_{k=0}^n a_k x^k \right)^2 dx.$$

Figure 8.5



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

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

Since

$$E = \int_a^b [f(x)]^2 dx - 2 \sum_{k=0}^n a_k \int_a^b x^k f(x) dx + \int_a^b \left(\sum_{k=0}^n a_k x^k \right)^2 dx,$$

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**

$$(8.6) \quad \sum_{k=0}^n a_k \int_a^b x^{j+k} dx = \int_a^b x^j f(x) dx, \quad \text{for each } j = 0, 1, \dots, n,$$

must be solved for the $(n + 1)$ unknowns a_j . It can be shown that the normal equations always have a unique solution provided $f \in C[a, b]$. (See Exercise 15.)

EXAMPLE 1 Find the least squares approximating polynomial of degree two 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

$$\begin{aligned} a_0 \int_0^1 1 dx + a_1 \int_0^1 x dx + a_2 \int_0^1 x^2 dx &= \int_0^1 \sin \pi x dx, \\ a_0 \int_0^1 x dx + a_1 \int_0^1 x^2 dx + a_2 \int_0^1 x^3 dx &= \int_0^1 x \sin \pi x dx, \\ a_0 \int_0^1 x^2 dx + a_1 \int_0^1 x^3 dx + a_2 \int_0^1 x^4 dx &= \int_0^1 x^2 \sin \pi x dx. \end{aligned}$$

Performing the integration yields

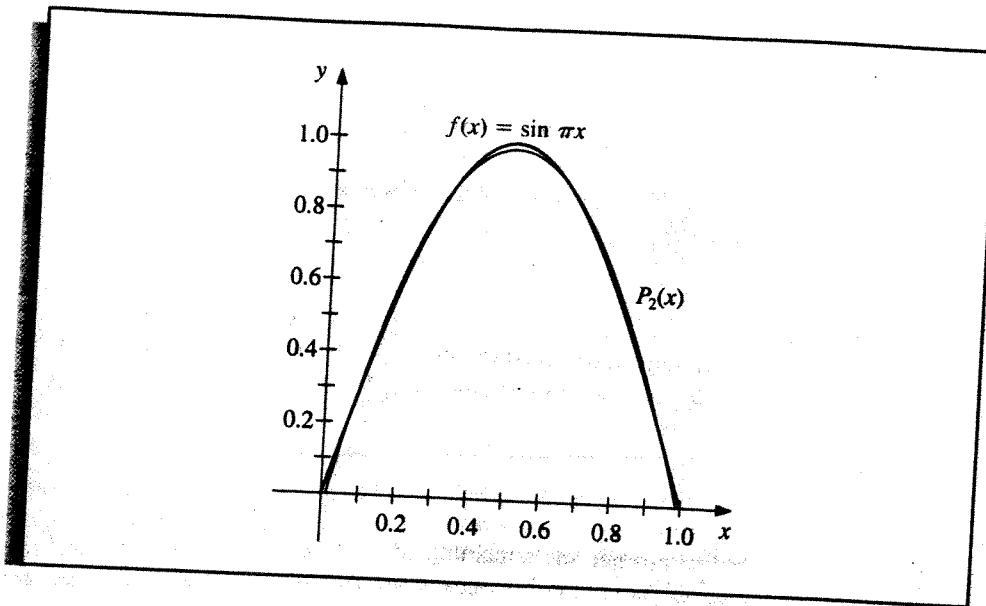
$$a_0 + \frac{1}{2}a_1 + \frac{1}{3}a_2 = \frac{2}{\pi}, \quad \frac{1}{2}a_0 + \frac{1}{3}a_1 + \frac{1}{4}a_2 = \frac{1}{\pi}, \quad \frac{1}{3}a_0 + \frac{1}{4}a_1 + \frac{1}{5}a_2 = \frac{720 - 60\pi^2}{\pi^3}.$$

These three equations in three unknowns can be solved to obtain

$$a_0 = \frac{12\pi^2 - 120}{\pi^3} \approx -0.050465 \quad \text{and} \quad a_1 = -a_2 = \frac{720 - 60\pi^2}{\pi^3} \approx 4.17$$

Consequently, the least squares polynomial approximation of degree two for $f(x) = \sin \pi x$ on $[0, 1]$ is $P_2(x) = -4.12251x^2 + 4.12251x - 0.050465$. (See Figure 8.6.)

Figure 8.6



Example 1 illustrates the difficulty in obtaining a least squares polynomial approximation. An $(n + 1) \times (n + 1)$ linear system must be solved for the coefficients a_0, \dots, a_n of $P_n(x)$. The coefficients in the linear system are of the form

$$\int_a^b x^{j+k} dx = \frac{b^{j+k+1} - a^{j+k+1}}{j+k+1},$$

a linear system that does not have an easily computed numerical solution. The matrix in the linear system is known as a **Hilbert matrix**. This ill-conditioned matrix is a classic example for demonstrating roundoff error difficulties. (See Exercise 6 of Section 7.4.) Another disadvantage is similar to the situation that occurred when the Lagrange polynomials were first introduced in Section 3.1. The calculations that were performed in obtaining the best n th-degree polynomial, $P_n(x)$, do not lessen the amount of work required to obtain $P_{n+1}(x)$, the polynomial of next higher degree.

A different technique to obtain least squares approximations will now be considered. This turns out to be computationally efficient, and once $P_n(x)$ is known, it is easy to determine $P_{n+1}(x)$. To facilitate the discussion, we need some new concepts.

Definition 8.1 The set of functions $\{\phi_0, \dots, \phi_n\}$ is said to be **linearly independent** on $[a, b]$ if, whenever

$$c_0\phi_0(x) + c_1\phi_1(x) + \dots + c_n\phi_n(x) = 0, \quad \text{for all } x \in [a, b],$$

then $c_0 = c_1 = \dots = c_n = 0$.

Otherwise the set of functions is said to be **linearly dependent**.