Chapter 7 Approximation Theory

Least Square 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 $\sum_{i=1}^{m}[y_i - (ax_i + b)]^2$ with respect to the parameters a and b. For a minimum to occur, it it necessary that

$$0 = \frac{\partial}{\partial a}\sum_{i=1}^{m}[y_i - (ax_i + b)]^2 = 2\sum_{i=1}^{m}(y_i - ax_i - b)(-x_i),$$

and $0 = \frac{\partial}{\partial b}\sum_{i=1}^{m}[y_i - (ax_i + b)]^2 = 2\sum_{i=1}^{m}(y_i - ax_i - b)(-1)$.

These equations simplify to the **normal equations:**

$$a\sum_{i=1}^{m}x_i^2 + b\sum_{i=1}^{m}x_i = \sum_{i=1}^{m}x_i y_i \text{ and } a\sum_{i=1}^{m}x_i + b\cdot m = \sum_{i=1}^{m}y_i .$$

The solution to this system of equations is

$$a = \frac{m\left(\sum_{i=1}^{m}x_i y_i\right) - \left(\sum_{i=1}^{m}x_i\right)\left(\sum_{i=1}^{m}y_i\right)}{m\left(\sum_{i=1}^{m}x_i^2\right) - \left(\sum_{i=1}^{m}x_i\right)^2},$$

and

$$b = \frac{\left(\sum_{i=1}^{m} x_i^2\right)\left(\sum_{i=1}^{m} y_i\right) - \left(\sum_{i=1}^{m} x_i y_i\right)\left(\sum_{i=1}^{m} x_i\right)}{m\left(\sum_{i=1}^{m} x_i^2\right) - \left(\sum_{i=1}^{m} x_i\right)^2}.$$

The general problem of approximating a set of data, $\{(x_i, y_i)\}_{i=1}^{m}$, with an algebraic polynomial $P_n(x) = \sum_{k=0}^{n} a_k x^k$ of degree $n < m - 1$ using the least squares procedure is handled in a similar manner. It requires choosing the constants $a_0, a_1, \ldots, a_n$ to minimize the least squares error

$$\begin{aligned}
E &= \sum_{i=1}^{m} (y_i - P_n(x_i))^2 \\
&= \sum_{i=1}^{m} y_i^2 - 2\sum_{i=1}^{m} P_n(x_i) y_i + \sum_{i=1}^{m} (P_n(x_i))^2 \\
&= \sum_{i=1}^{m} y_i^2 - 2\sum_{i=1}^{m}\left(\sum_{j=0}^{n} a_j x_i^j\right) y_i + \sum_{i=1}^{m}\left(\sum_{j=0}^{n} a_j x_i^j\right)^2 \\
&= \sum_{i=1}^{m} y_i^2 - 2\sum_{j=0}^{n} a_j \left(\sum_{i=1}^{m} y_i x_i^j\right) + \sum_{j=0}^{n}\sum_{k=0}^{n} a_j a_k \left(\sum_{i=1}^{m} x_i^{j+k}\right).
\end{aligned}$$

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,

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

This gives $n+1$ **normal equations** in the $n+1$ unknowns, a_j,

$$\sum_{k=0}^{n} a_k \sum_{i=1}^{m} x_i^{j+k} = \sum_{i=1}^{m} y_i x_i^j \,, \qquad j = 0,1,\dots,n\,.$$

It is helpful to write the equations as follows:

$$a_0 \sum_{i=1}^{m} x_i^0 + a_1 \sum_{i=1}^{m} x_i^1 + a_2 \sum_{i=1}^{m} x_i^2 + \cdots + a_n \sum_{i=1}^{m} x_i^n = \sum_{i=1}^{m} y_i x_i^0,$$

$$a_0 \sum_{i=1}^{m} x_i^1 + a_1 \sum_{i=1}^{m} x_i^2 + a_2 \sum_{i=1}^{m} x_i^3 + \cdots + a_n \sum_{i=1}^{m} x_i^{n+1} = \sum_{i=1}^{m} y_i x_i^1,$$

$$\vdots$$

$$a_0 \sum_{i=1}^{m} x_i^n + a_1 \sum_{i=1}^{m} x_i^{n+1} + a_2 \sum_{i=1}^{m} x_i^{n+2} + \cdots + a_n \sum_{i=1}^{m} x_i^{2n} = \sum_{i=1}^{m} y_i x_i^n$$

It can be shown that the normal equations have a unique solution provided that the x_i, $i = 1,2,\dots,m$ are distinct.

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

$$y = be^{ax} \qquad (1)$$

or

$$y = bx^a \qquad (2)$$

for some constants a and b. Considering the logarithm of the approximating equations:

$\ln y = \ln b + ax$, in the case of Eq. (1)

and $\ln y = \ln b + a \ln x$, in the case of Eq. (2).

In either case, a linear problem now appears and solutions for $\ln b$ and

a can be obtained by appropriately modifying the normal equations (1) and (2).

In fitting a function to data points $\{(x_i, y_i)\}_{i=1}^{m}$, a linear combination of any known functions, including polynomials, may be used:

$$g(x) = c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) + \cdots + c_k f_k(x) \qquad (3)$$

where $f_1, f_2,..$ are prescribed functions, $c_1, c_2,..$ are undetermined coefficients, and k is the total number of prescribed functions. By fitting Eq. (3) to each data point, an over-determined equation is written as

$$\mathbf{Ac} = \mathbf{y},$$

with

$$\mathbf{A} = \begin{bmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_k(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_k(x_2) \\ \vdots & \vdots & & \vdots \\ f_1(x_L) & f_2(x_L) & \cdots & f_k(x_L) \end{bmatrix},$$

$$c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix}, \text{ and } \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_L \end{bmatrix}, \text{ where } L > k.$$

References:

【1】 R. L. Burden and J. D. Faires, *Numerical Analysis*, PWS, Boston, 1993.

【2】 S. Nakamura, *Numerical Analysis and Graphic Visualization with MATLAB*, Prentice-Hall, New Jersey, 1996.