5 Least Square Problems

5.1 Introduction

Let the set of data points be (x_i, y_i), $\quad i = 1,\ 2,\ ...,\ m$, and let the curve given by $y = f(x)$ be fitted to this data. If e_i is the error of the approximation at $x = x_i$ due to this fitting then, $e_i = y_i - f(x_i)$. If we write $S = \sum_{i=1}^{m} e_i^2 = \sum_{i=1}^{m} [y_i - f(x_i)]^2$. Then the method of minimizing this error which is the sum of the square of errors is called the method of least squares.

5.1.1 Fitting a Straight Line

Let $Y = a_0 + a_1 x$ be the straight line to be fitted to the given data by the method of least squares. Then the sum of errors is $S = \sum_{i=1}^{m} [y_i - a_0 - a_1 x_i]^2$. S is a function of a_0 and a_1. To minimize S, $\dfrac{\partial S}{\partial a_0} = 0$ and $\dfrac{\partial S}{\partial a_1} = 0$. We have,

$$\frac{\partial S}{\partial a_0} = -2\sum_{i=1}^{m} [y_i - a_0 - a_1 x_i], \qquad \frac{\partial S}{\partial a_1} = -2x_i\sum_{i=1}^{m} [y_i - a_0 - a_1 x_i] \tag{5.1}$$

$$\frac{\partial S}{\partial a_0} = 0 \implies -2\sum_{i=1}^{m} [y_i - a_0 - a_1 x_i] = 0$$

$$\implies \sum_{i=1}^{m} y_i = \sum_{i=1}^{m} a_0 + a_1 \sum_{i=1}^{m} x_i$$

$$\implies \sum_{i=1}^{m} y_i = a_1 \sum_{i=1}^{m} x_i + m a_0 \tag{5.2}$$

$$\frac{\partial S}{\partial a_1} = 0 \implies -2x_i\sum_{i=1}^{m} [y_i - a_0 - a_1 x_i] = 0$$

$$\implies \sum_{i=1}^{m} x_i y_i = a_0 \sum_{i=1}^{m} x_i + a_1 \sum_{i=1}^{m} x_i^2 \tag{5.3}$$

The equations 5.2 and 5.3 are called normal equations which can be solved for a_0 and a_1. From the equation 5.2, it can be easily shown that the fitted line passes through the *means* of x_i and y_i,i.e the line satisfies $\overline{y} = a_0 + a_1 \overline{x}$.

5.1.2 Multiple Linear Least Square

Let $z = a_0+a_1x+a_2y$ be a linear equation to be fitted to the given data $(x_1, y_1, z_1), (x_2, y_2, z_2), ..., (x_m, y, z_m)$ by the method of least squares. Then the sum of errors is $S = \sum_{i=1}^{m} [z_i - a_0 - a_1x_i - a_2y_i]^2$. S is a function of a_0, a_1 and a_2. To minimize S:

$$\frac{\partial S}{\partial a_0} = -2\sum_{i=1}^{m} [y_i - a_0 - a_1x_i - a_2y_i] = 0$$

$$\frac{\partial S}{\partial a_1} = -2x_i\sum_{i=1}^{m} [y_i - a_0 - a_1x_i - a_2y_i] = 0$$

$$\frac{\partial S}{\partial a_2} = -2y_i\sum_{i=1}^{m} [y_i - a_0 - a_1x_i - a_2y_i] = 0$$

So, the normal equations are:

$$\begin{aligned}
\sum_{i=1}^{m} z_i &= a_1\sum_{i=1}^{m} x_i + a_2\sum_{i=1}^{m} y_i + ma_0\\
\sum_{i=1}^{m} z_i\,x_i &= a_1\sum_{i=1}^{m} x_i^2 + a_2\sum_{i=1}^{m} x_iy_i + a_0\sum_{i=1}^{m} x_i\\
\sum_{i=1}^{m} z_i\,y_i &= a_1\sum_{i=1}^{m} x_iy_i + a_2\sum_{i=1}^{m} y_i^2 + a_0\sum_{i=1}^{m} y_i
\end{aligned} \tag{5.4}$$

5.2 Curve Fitting by Polynomials

Let $Y = a_0 + a_1x + a_2x^2 + ... + a_nx^n$ be a polynomial to be fitted to the given data $(x_1, y_1), (x_2, y_2), ..., (x_m, y_m)$ by the method of least squares. Then the sum of errors is $S = \sum_{i=1}^{m} [z_i - a_0 - a_1x_i - a_2y_i]^2$. Then the normal equations are:

$$\begin{aligned}
\sum_{i=1}^{m} y_i &= a_1\sum_{i=1}^{m} x_i + a_2\sum_{i=1}^{m} x_i^2 + ... + ma_0\\
\sum_{i=1}^{m} x_i\,y_i &= a_0\sum_{i=1}^{m} x_i + a_1\sum_{i=1}^{m} x_i^2 + a_2\sum_{i=1}^{m} x_i^3 + ... + a_n\sum_{i=1}^{m} x_i^{n+1}\\
&...\\
\sum_{i=1}^{m} x_i^n\,y_i &= a_0\sum_{i=1}^{m} x_i^n + a_1\sum_{i=1}^{m} x_i^{n+1} + a_2\sum_{i=1}^{m} x_i^{n+2} + ... + a_n\sum_{i=1}^{m} x_i^{2n}
\end{aligned} \tag{5.5}$$

This system constitutes of $(n + 1)$ equations in $(n + 1)$ unknowns. For larger value of n the system is *ill conditioned*, so orthogonal polynomials are used to fit such data points.

5.3 Curve Fitting by a Sum of Exponents

Let $A_1e^{\lambda_1x} + A_2e^{\lambda_2x} + ... + A_ne^{\lambda_nx}$ be the sum of exponents fitted to the given data $(x_1, y_1), (x_2, y_2), ..., (x_m, y_m)$, where m is much larger than $2n$. For the convenience of our presentation, we take $n = 2$ and $m >> 4$. Then we have, e

$$A_1 e^{\lambda_1 x} + A_2 e^{\lambda_2 x} \tag{5.6}$$

From the theory of differential equation theory, 5.6 is a solution of a second order homogeneous differential equation say,

$$y''(x) = a_1 y'(x) + a_2 y(x) \tag{5.7}$$

where a_1, a_2 are some constants. Integrating this equation on $[x_0, x]$ repeatedly, we get,

$$\begin{aligned} y(x_i) + y(x_j) - 2y(x_0) = a_1 \left[\int_{x_0}^{x_i} y(x)\,dx + \int_{x_0}^{x_j} y(x)\,dx \right] \\ + a_2 \left[\int_{x_0}^{x_i} (x_i - x) y(x)\,dx + \int_{x_0}^{x_j} (x_j - x) y(x)\,dx \right] \end{aligned} \tag{5.8}$$

This equation 5.8 can be set for a linear system to find a_1 and a_2. Then the values of λ_1 and λ_2 can be obtained from characteristic equation of 5.7 which is

$$\lambda^2 = a_1 \lambda + a_2 \tag{5.9}$$

Finally A_1, A_2 can be obtained from the least squares or any other suitable method.