

問 1

(x,y) の実験データがあり、その誤差の二乗の和は、以下のように表すことができる。

$$E(a,b) = \sum_{i=1}^N (y_i - ax_i - b)^2$$

(x_i, y_i) は i 番目のデータ、 N はデータの個数である。この時誤差が最小になる a, b を求めよ。(最小二乗法)

解答

a も b も 2 次式でその係数は正の値なので最小値がある。誤差の最小値は、それぞれ偏微分した値がゼロとなるときに得ることができる。

$$\frac{\partial E}{\partial a} = -\sum_{i=1}^N 2(y_i - ax_i - b)x_i = 0$$

$$\frac{\partial E}{\partial b} = -\sum_{i=1}^N 2(y_i - ax_i - b) = 0$$

問 2

次の四点 $(2, 2), (3, 4), (5, 6), (7, 9)$ を近似する直線が $y=ax+b$ とする。最小二乗法を用いて係数 a と b を求めよ。

解答

問 1 の結果を利用する。

$$\sum_{i=1}^N x_i = 17$$

$$\sum_{i=1}^N y_i = 21$$

これは a と b の連立方程式である。すなわち、

$$\begin{cases} a \sum_{i=1}^N x_i^2 + b \sum_{i=1}^N x_i = \sum_{i=1}^N x_i y_i \\ a \sum_{i=1}^N x_i + Nb = \sum_{i=1}^N y_i \end{cases}$$

である。これを解くと、

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

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

となる。

以上

$$\sum_{i=1}^N x_i y_i = 109$$

$$\sum_{i=1}^N x_i^2 = 87$$

以上の値を問 1 の結果に代入する。

$$a=1.34 \quad , \quad b=0.44$$

以上