

NO.

0083

4点データに対する直線の最小二乗フィット

データ点 (x_i, y_i) ($i = 1, \dots, n$) に対して、モデル関数 $f(x)$ (未知パラメータを含む) を当てはめる最小二乗法では、誤差二乗和

$$S = \sum_{i=1}^n (y_i - f(x_i))^2$$

を最小にするようにパラメータを定める。 f が未知パラメータについて線形 (ここでは $f(x) = ax + b$) である場合、 S は各パラメータの2次関数であるから、最小値では各パラメータに関する偏微分がすべて0になる。

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

とおくと、

$$\frac{\partial S}{\partial a} = -2 \sum_i x_i (y_i - ax_i - b) = 0, \quad \frac{\partial S}{\partial b} = -2 \sum_i (y_i - ax_i - b) = 0$$

これを整理すると、正規方程式

$$\begin{cases} a \sum_i x_i^2 + b \sum_i x_i = \sum_i x_i y_i \\ a \sum_i x_i + b n = \sum_i y_i \end{cases}$$

を得る。

$n = 4$ とし、必要な和を計算すると

$$\sum x_i = 6, \quad \sum x_i^2 = 14, \quad \sum y_i = 9, \quad \sum x_i y_i = 18$$

正規方程式は

$$14a + 6b = 18, \quad 6a + 4b = 9$$

第1式を2で割って $7a + 3b = 9$ 。これを4倍、第2式を3倍すると

$$28a + 12b = 36, \quad 18a + 12b = 27$$

辺々引いて $10a = 9$ 、すなわち $a = \frac{9}{10}$ 。これを $6a + 4b = 9$ に代入すると

$$6 \cdot \frac{9}{10} + 4b = 9 \implies 4b = 9 - \frac{27}{5} = \frac{18}{5} \implies b = \frac{9}{10}$$

したがって求める直線は

$$y = \frac{9}{10}x + \frac{9}{10}$$

(検算: 残差 $r_i = y_i - (0.9x_i + 0.9)$ は $r_1 = 0.1$, $r_2 = 0.2$, $r_3 = -0.7$, $r_4 = 0.4$ で、 $\sum r_i = 0$ かつ $\sum x_i r_i = 0 \cdot 0.1 + 1 \cdot 0.2 + 2 \cdot (-0.7) + 3 \cdot 0.4 = 0.2 - 1.4 + 1.2 = 0$ となり、正規方程式を満たすことが確認できる。)