

単回帰曲線の求め方

ここでは例として、2次曲線の一般型として次の曲線から考えます。

$$y = b_0 + b_1x + b_2x^2 + b_3x^3 + b_4x^4 + b_5x^5 (= f(x))$$

この曲線について最小2乗法を当てはめていきます。

実際の値と曲線との距離 $L = |y_i - f(x_i)|$ の2乗和の最小値を求める。実際には、偏微分

$$\frac{\partial L}{\partial b_0} = 0, \frac{\partial L}{\partial b_1} = 0, \frac{\partial L}{\partial b_2} = 0, \frac{\partial L}{\partial b_3} = 0, \frac{\partial L}{\partial b_4} = 0, \frac{\partial L}{\partial b_5} = 0$$

を解けばよい。

$$\frac{\partial L}{\partial b_0} = \frac{\partial}{\partial b_0} \sum_{i=0}^{\infty} (y_i - b_0 - b_1x - b_2x^2 - b_3x^3 - b_4x^4 - b_5x^5)^2 = 0$$

$$\frac{\partial L}{\partial b_1} = \frac{\partial}{\partial b_1} \sum_{i=0}^{\infty} (y_i - b_0 - b_1x - b_2x^2 - b_3x^3 - b_4x^4 - b_5x^5)^2 = 0$$

$$\frac{\partial L}{\partial b_2} = \frac{\partial}{\partial b_2} \sum_{i=0}^{\infty} (y_i - b_0 - b_1x - b_2x^2 - b_3x^3 - b_4x^4 - b_5x^5)^2 = 0$$

$$\frac{\partial L}{\partial b_3} = \frac{\partial}{\partial b_3} \sum_{i=0}^{\infty} (y_i - b_0 - b_1x - b_2x^2 - b_3x^3 - b_4x^4 - b_5x^5)^2 = 0$$

$$\frac{\partial L}{\partial b_4} = \frac{\partial}{\partial b_4} \sum_{i=0}^{\infty} (y_i - b_0 - b_1x - b_2x^2 - b_3x^3 - b_4x^4 - b_5x^5)^2 = 0$$

$$\frac{\partial L}{\partial b_5} = \frac{\partial}{\partial b_5} \sum_{i=0}^{\infty} (y_i - b_0 - b_1x - b_2x^2 - b_3x^3 - b_4x^4 - b_5x^5)^2 = 0$$

偏微分すると、

$$2 \sum_{i=0}^{\infty} (y_i - b_0 - b_1x_i - b_2x_i^2 - b_3x_i^3 - b_4x_i^4 - b_5x_i^5) = 0$$

$$2 \sum_{i=0}^{\infty} (y_i - b_0 - b_1x_i - b_2x_i^2 - b_3x_i^3 - b_4x_i^4 - b_5x_i^5)x_i = 0$$

$$2 \sum_{i=0}^{\infty} (y_i - b_0 - b_1x_i - b_2x_i^2 - b_3x_i^3 - b_4x_i^4 - b_5x_i^5)x_i^2 = 0$$

$$2 \sum_{i=0}^{\infty} (y_i - b_0 - b_1x_i - b_2x_i^2 - b_3x_i^3 - b_4x_i^4 - b_5x_i^5)x_i^3 = 0$$

$$2 \sum_{i=0}^{\infty} (y_i - b_0 - b_1x_i - b_2x_i^2 - b_3x_i^3 - b_4x_i^4 - b_5x_i^5)x_i^4 = 0$$

$$2 \sum_{i=0}^{\infty} (y_i - b_0 - b_1x_i - b_2x_i^2 - b_3x_i^3 - b_4x_i^4 - b_5x_i^5)x_i^5 = 0$$

更に整理して、

$$\sum_{i=0}^{\infty} y_i = nb_0 + b_1 \sum_{i=0}^{\infty} x_i + b_2 \sum_{i=0}^{\infty} x_i^2 + b_3 \sum_{i=0}^{\infty} x_i^3 + b_4 \sum_{i=0}^{\infty} x_i^4 + b_5 \sum_{i=0}^{\infty} x_i^5 \quad (1)$$

$$\sum_{i=0}^{\infty} x_i y_i = b_0 \sum_{i=0}^{\infty} x_i + b_1 \sum_{i=0}^{\infty} x_i^2 + b_2 \sum_{i=0}^{\infty} x_i^3 + b_3 \sum_{i=0}^{\infty} x_i^4 + b_4 \sum_{i=0}^{\infty} x_i^5 + b_5 \sum_{i=0}^{\infty} x_i^6 \quad (2)$$

$$\sum_{i=0}^{\infty} x_i^2 y_i = b_0 \sum_{i=0}^{\infty} x_i^2 + b_1 \sum_{i=0}^{\infty} x_i^3 + b_2 \sum_{i=0}^{\infty} x_i^4 + b_3 \sum_{i=0}^{\infty} x_i^5 + b_4 \sum_{i=0}^{\infty} x_i^6 + b_5 \sum_{i=0}^{\infty} x_i^7 \quad (3)$$

$$\sum_{i=0}^{\infty} x_i^3 y_i = b_0 \sum_{i=0}^{\infty} x_i^3 + b_1 \sum_{i=0}^{\infty} x_i^4 + b_2 \sum_{i=0}^{\infty} x_i^5 + b_3 \sum_{i=0}^{\infty} x_i^6 + b_4 \sum_{i=0}^{\infty} x_i^7 + b_5 \sum_{i=0}^{\infty} x_i^8 \quad (4)$$

$$\sum_{i=0}^{\infty} x_i^4 y_i = b_0 \sum_{i=0}^{\infty} x_i^4 + b_1 \sum_{i=0}^{\infty} x_i^5 + b_2 \sum_{i=0}^{\infty} x_i^6 + b_3 \sum_{i=0}^{\infty} x_i^7 + b_4 \sum_{i=0}^{\infty} x_i^8 + b_5 \sum_{i=0}^{\infty} x_i^9 \quad (5)$$

$$\sum_{i=0}^{\infty} x_i^5 y_i = b_0 \sum_{i=0}^{\infty} x_i^5 + b_1 \sum_{i=0}^{\infty} x_i^6 + b_2 \sum_{i=0}^{\infty} x_i^7 + b_3 \sum_{i=0}^{\infty} x_i^8 + b_4 \sum_{i=0}^{\infty} x_i^9 + b_5 \sum_{i=0}^{\infty} x_i^{10} \quad (6)$$

これらの (1),(2),(3),(4),(5),(6) の連立方程式を解けば、 $b_0, b_1, b_2, b_3, b_4, b_5$ の値がわかるので、5次の回帰曲線が特定できる。