重回帰分析との関係について（36p）

2010. 5. 25　石井

いま、変数 $b, x_1, x_2, \cdots, x_d$ があり、変数 b は変数 $x_1, x_2, \cdots, x_d$ の線形結合として、

$$\hat{b} = w_0 + w_1 x_1 + w_2 x_2 + \cdots + w_d x_d \tag{1}$$

の如く推定されると仮定する。ここで $\hat{b}$ は b の推定値である。この b を目的変数、$x_1, x_2, \cdots, x_d$ を説明変数と呼ぶ。実際に、$b, x_1, x_2, \cdots, x_d$ を観測したところ、次のような n 組のデータが得られたとする。

$$\left.\begin{array}{l} b_1, x_{11}, x_{12}, \cdots, x_{1d} \\ b_2, x_{21}, x_{22}, \cdots, x_{2d} \\ \cdots \\ \cdots \\ b_p, x_{p1}, x_{p2}, \cdots, x_{pd} \\ \cdots \\ \cdots \\ b_n, x_{n1}, x_{n2}, \cdots, x_{nd} \end{array}\right\} n \text{ 組} \tag{2}$$

したがって、

$$\left.\begin{array}{l} b_1 = w_0 + w_1 x_{11} + w_2 x_{12} + \cdots + w_d x_{1d} + e_1 \\ b_2 = w_0 + w_1 x_{21} + w_2 x_{22} + \cdots + w_d x_{2d} + e_2 \\ \cdots \\ \cdots \\ b_p = w_0 + w_1 x_{p1} + w_2 x_{p2} + \cdots + w_d x_{pd} + e_p \\ \cdots \\ \cdots \\ b_n = w_0 + w_1 x_{n1} + w_2 x_{n2} + \cdots + w_d x_{nd} + e_n \end{array}\right\} \tag{3}$$

と書ける。ここで $e_1, e_2, \cdots, e_n$ はエラー項である。簡単のため、以下のベクトル表記を用いる。

$$\mathbf{x}_p \stackrel{\text{def}}{=} (1, x_{p1}, x_{p2}, \cdots, x_{pd})^t \qquad (1 \le p \le n) \tag{4}$$

$$\mathbf{w} \stackrel{\text{def}}{=} (w_0, w_1, w_2, \cdots, w_d)^t \tag{5}$$

$$\mathbf{b} \stackrel{\text{def}}{=} (b_1, b_2, \cdots, b_n)^t \tag{6}$$

$$\mathbf{e} \stackrel{\text{def}}{=} (e_1, e_2, \cdots, e_n)^t \tag{7}$$

すなわち、式 (4) は説明変数に関する n 組のデータであり、

$$\left.\begin{array}{l} \mathbf{x}_1 = (1, x_{11}, x_{12}, \cdots, x_{1d})^t \\ \mathbf{x}_2 = (1, x_{21}, x_{22}, \cdots, x_{2d})^t \\ \cdots \\ \cdots \\ \mathbf{x}_p = (1, x_{p1}, x_{p2}, \cdots, x_{pd})^t \\ \cdots \\ \cdots \\ \mathbf{x}_n = (1, x_{n1}, x_{n2}, \cdots, x_{nd})^t \end{array}\right\} \tag{8}$$

を示している。これらより式 (3) は

$$b_p = \mathbf{w}^t \mathbf{x}_p + e_p \qquad (1 \leq p \leq n) \tag{9}$$

と書ける。行列 $\mathbf{X}$ を

$$\mathbf{X} \stackrel{\text{def}}{=} (\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)^t \tag{10}$$

と定義すると、式 (3) は、

$$\mathbf{b} = \mathbf{X}\mathbf{w} + \mathbf{e} \tag{11}$$

と書ける。ここで $\|\mathbf{e}\|^2$ を最小化するため、

$$\frac{\partial}{\partial \mathbf{w}} \|\mathbf{e}\|^2 = 0 \tag{12}$$

となる $\mathbf{w}$ を求める。

$$\begin{aligned} \|\mathbf{e}\|^2 &= \|\mathbf{X}\mathbf{w} - \mathbf{b}\|^2 && (13) \\ &= (\mathbf{X}\mathbf{w} - \mathbf{b})^t (\mathbf{X}\mathbf{w} - \mathbf{b}) && (14) \\ &= (\mathbf{w}^t \mathbf{X}^t - \mathbf{b}^t)(\mathbf{X}\mathbf{w} - \mathbf{b}) && (15) \\ &= \mathbf{w}^t \mathbf{X}^t \mathbf{X}\mathbf{w} + \mathbf{b}^t \mathbf{b} - \mathbf{w}^t \mathbf{X}^t \mathbf{b} - \mathbf{b}^t \mathbf{X}\mathbf{w} && (16) \\ &= \mathbf{w}^t \mathbf{X}^t \mathbf{X}\mathbf{w} + \mathbf{b}^t \mathbf{b} - 2\mathbf{w}^t \mathbf{X}^t \mathbf{b} && (17) \end{aligned}$$

であるから、式 (12) より、

$$\begin{aligned} \frac{\partial}{\partial \mathbf{w}} \|\mathbf{e}\|^2 &= 2\mathbf{X}^t \mathbf{X}\mathbf{w} - 2\mathbf{X}^t \mathbf{b} && (18) \\ &= 2\mathbf{X}^t (\mathbf{X}\mathbf{w} - \mathbf{b}) = 0 && (19) \end{aligned}$$

となり、

$$\mathbf{X}^t \mathbf{X}\mathbf{w} = \mathbf{X}^t \mathbf{b} \tag{20}$$

が成り立つ。これより最終的に、

$$\mathbf{w} = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \mathbf{b} \tag{21}$$

が得られる。これは教科書 36p の式（3.22）と同じである。