

計量経済学 補足資料

－ 最小2乗パラメータ推定値の導出 －

河田 正樹

2015年5月13日

単純回帰モデル

$$Y_i = a + bX_i + u_i$$

の回帰係数 (パラメータともいう) の推定値を最小2乗法によって求めることは、実績値 (Y_i) から予測値 (\hat{Y}_i) を引いた残差 e_i の2乗和を最小にするような \hat{a}, \hat{b} を求めることである。

予測値は

$$\hat{Y}_i = \hat{a} + \hat{b}X_i$$

となるので、残差の2乗和 (これを G とあらわす) は、

$$G = (Y_1 - \hat{a} - \hat{b}X_1)^2 + (Y_2 - \hat{a} - \hat{b}X_2)^2 + \cdots + (Y_n - \hat{a} - \hat{b}X_n)^2$$

となる。

平方完成による導出

カッコの中を次のように変形する。

$$\begin{aligned} Y_1 - \hat{a} - \hat{b}X_1 &= (Y_1 - \bar{Y}) - \hat{b}(X_1 - \bar{X}) + (\bar{Y} - \hat{b}\bar{X} - \hat{a}) \\ &= y_1 - \hat{b}x_1 + (\bar{Y} - \hat{b}\bar{X} - \hat{a}) \end{aligned}$$

すると、 G は

$$\begin{aligned} G &= \{(y_1 - \hat{b}x_1) + (\bar{Y} - \hat{b}\bar{X} - \hat{a})\}^2 + \cdots + \{(y_n - \hat{b}x_n) + (\bar{Y} - \hat{b}\bar{X} - \hat{a})\}^2 \\ &= \{(y_1 - \hat{b}x_1)^2 + 2(y_1 - \hat{b}x_1)(\bar{Y} - \hat{b}\bar{X} - \hat{a}) + (\bar{Y} - \hat{b}\bar{X} - \hat{a})^2\} \\ &\quad + \cdots + \{(y_n - \hat{b}x_n)^2 + 2(y_n - \hat{b}x_n)(\bar{Y} - \hat{b}\bar{X} - \hat{a}) + (\bar{Y} - \hat{b}\bar{X} - \hat{a})^2\} \\ &= (y_1 - \hat{b}x_1)^2 + \cdots + (y_n - \hat{b}x_n)^2 \\ &\quad + 2(y_1 - \hat{b}x_1)(\bar{Y} - \hat{b}\bar{X} - \hat{a}) + \cdots + 2(y_n - \hat{b}x_n)(\bar{Y} - \hat{b}\bar{X} - \hat{a}) \\ &\quad + (\bar{Y} - \hat{b}\bar{X} - \hat{a})^2 + \cdots + (\bar{Y} - \hat{b}\bar{X} - \hat{a})^2 \\ &= (y_1 - \hat{b}x_1)^2 + \cdots + (y_n - \hat{b}x_n)^2 \\ &\quad + (\bar{Y} - \hat{b}\bar{X} - \hat{a})\{2(y_1 - \hat{b}x_1) + \cdots + 2(y_n - \hat{b}x_n)\} \\ &\quad + n(\bar{Y} - \hat{b}\bar{X} - \hat{a})^2 \end{aligned}$$

となる。ここで、

$$\begin{aligned} (\bar{Y} - \hat{b}\bar{X} - \hat{a})\{2(y_1 - \hat{b}x_1) + \cdots + 2(y_n - \hat{b}x_n)\} &= 2(\bar{Y} - \hat{b}\bar{X} - \hat{a})(y_1 + \cdots + y_n) \\ &\quad - 2\hat{b}(\bar{Y} - \hat{b}\bar{X} - \hat{a})(x_1 + \cdots + x_n) \end{aligned}$$

は、 $(y_1 + \dots + y_n) = 0, (x_1 + \dots + x_n) = 0$ より 0 となるので、

$$G = (y_1 - \hat{b}x_1)^2 + \dots + (y_n - \hat{b}x_n)^2 + n(\bar{Y} - \hat{b}\bar{X} - \hat{a})^2 \quad (1)$$

となる。 G を最小にするような \hat{a} を考えると、

$$\hat{a} = \bar{Y} - \hat{b}\bar{X}$$

となる。このとき、(1) 式の最後の項は 0 となるので、 G を最小にする \hat{b} を求めるには、

$$(y_1 - \hat{b}x_1)^2 + \dots + (y_n - \hat{b}x_n)^2 \quad (2)$$

を最小にする \hat{b} を求めれば良い。この式は \hat{b} の 2 次式と見ることができ、平方完成によって最小値をとる \hat{b} が求められる。

$ax^2 + bx + c$ を最小にする x を求めるとき、 $a(x - \frac{b}{2a})^2 + \dots$ という変形によって $x = \frac{b}{2a}$ のとき最小値をとることがわかる。これが平方完成である。(2) 式を変形すると

$$\begin{aligned} (y_1 - \hat{b}x_1)^2 + \dots + (y_n - \hat{b}x_n)^2 &= (y_1^2 - 2\hat{b}x_1y_1 + \hat{b}^2x_1^2) + \dots + (y_n^2 - 2\hat{b}x_ny_n + \hat{b}^2x_n^2) \\ &= (x_1^2 + \dots + x_n^2)\hat{b}^2 - 2(x_1y_1 + \dots + x_ny_n)\hat{b} + (y_1^2 + \dots + y_n^2) \\ &= (x_1^2 + \dots + x_n^2)\left\{\hat{b} - \frac{(x_1y_1 + \dots + x_ny_n)}{(x_1^2 + \dots + x_n^2)}\right\}^2 + \dots \end{aligned}$$

よって、求める \hat{b} は

$$\hat{b} = \frac{(x_1y_1 + \dots + x_ny_n)}{(x_1^2 + \dots + x_n^2)} = \frac{S_{xy}}{S_x^2}$$

となる。

偏微分による導出

偏微分は高校で学んだ微分を 2 変数以上の場合に拡張したものであり、微分する変数以外の変数はすべて定数とみなしたものである。微分は曲線のある点における傾きを表すものであるから、

$$\text{微分が 0 に等しい} = \text{傾きが 0} = \text{最小値}$$

となる。よって、偏微分したものが 0 に等しいような \hat{a}, \hat{b} が、求める推定値である。

残差平方和 G は

$$G = (Y_1 - \hat{a} - \hat{b}X_1)^2 + \dots + (Y_n - \hat{a} - \hat{b}X_n)^2$$

となるので、 \hat{a} で偏微分すると

$$\begin{aligned} \frac{\partial G}{\partial \hat{a}} &= -2(Y_1 - \hat{a} - \hat{b}X_1) - \dots - 2(Y_n - \hat{a} - \hat{b}X_n) \\ &= -2\{(Y_1 - \hat{a} - \hat{b}X_1) + \dots + (Y_n - \hat{a} - \hat{b}X_n)\} \end{aligned} \quad (3)$$

また G を \hat{b} で偏微分すると

$$\begin{aligned} \frac{\partial G}{\partial \hat{b}} &= -2X_1(Y_1 - \hat{a} - \hat{b}X_1) - \dots - 2X_n(Y_n - \hat{a} - \hat{b}X_n) \\ &= -2\{X_1(Y_1 - \hat{a} - \hat{b}X_1) + \dots + X_n(Y_n - \hat{a} - \hat{b}X_n)\} \end{aligned} \quad (4)$$

(3) 式の {} 内を = 0 とし、 \hat{a} の項と X_i の項を右辺に移行すると、

$$Y_1 + \dots + Y_n = n\hat{a} + \hat{b}(X_1 + \dots + X_n) \quad (5)$$

(4) 式の {} 内を = 0 とし、展開した X_i の項と X_i^2 の項を右辺に移行すると、

$$X_1Y_1 + \dots + X_nY_n = \hat{a}(X_1 + \dots + X_n) + \hat{b}(X_1^2 + \dots + X_n^2) \quad (6)$$

(5) 式に $(X_1 + \dots + X_n)$ をかけたものと、(6) 式を n 倍したものを比べると、

$$(X_1 + \dots + X_n)(Y_1 + \dots + Y_n) = n\hat{a}(X_1 + \dots + X_n) + \hat{b}(X_1 + \dots + X_n)^2 \quad (7)$$

$$n(X_1Y_1 + \dots + X_nY_n) = n\hat{a}(X_1 + \dots + X_n) + n\hat{b}(X_1^2 + \dots + X_n^2) \quad (8)$$

(7) 式から (8) 式を引くと

$$(X_1 + \dots + X_n)(Y_1 + \dots + Y_n) - n(X_1Y_1 + \dots + X_nY_n) = \hat{b}\{(X_1 + \dots + X_n)^2 - n(X_1^2 + \dots + X_n^2)\}$$

となる。よって \hat{b} は

$$\begin{aligned} \hat{b} &= \frac{(X_1 + \dots + X_n)(Y_1 + \dots + Y_n) - n(X_1Y_1 + \dots + X_nY_n)}{(X_1 + \dots + X_n)^2 - n(X_1^2 + \dots + X_n^2)} \\ &= \frac{-\frac{1}{n}(X_1 + \dots + X_n)(Y_1 + \dots + Y_n) + (X_1Y_1 + \dots + X_nY_n)}{-\frac{1}{n}(X_1 + \dots + X_n)^2 + (X_1^2 + \dots + X_n^2)} \end{aligned}$$

となる。ところで、

$$\bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$$

$$\bar{Y} = \frac{1}{n}(Y_1 + \dots + Y_n)$$

を (9) 式に代入すると、

$$\hat{b} = \frac{-n\bar{X}\bar{Y} + (X_1Y_1 + \dots + X_nY_n)}{-n\bar{X}^2 + (X_1^2 + \dots + X_n^2)}$$

よって \hat{b} の最小 2 乗パラメータ推定値は

$$\hat{b} = \frac{S_{xy}}{S_x^2}$$

となる。また \hat{a} の最小 2 乗パラメータ推定値は (5) 式を \hat{a} についてとくと

$$\begin{aligned} \hat{a} &= \frac{1}{n}(Y_1 + \dots + Y_n) - \hat{b}\frac{1}{n}(X_1 + \dots + X_n) \\ &= \bar{Y} - \hat{b}\bar{X} \end{aligned}$$

となる。これが \hat{a} の最小 2 乗パラメータ推定値である。