

REGRESJA (zapis macierzowy)

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_k X_k$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \quad b = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix} \quad Y = Xb$$

$$b = (X^T X)^{-1} X^T Y$$

Własność

$$Y^T Y = \sum_{i=1}^n y_i^2$$

$$\text{Dla } k=1 \quad X^T X = \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix}, \quad X^T Y = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n y_i x_i \end{bmatrix},$$

$$\text{Dla } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \text{ gdy } ad - cb \neq 0$$

$$\text{to } A^{-1} = \frac{1}{ad - cb} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Wariancja resztowa.

Niech, $e = Y - \hat{Y}$, gdzie $\hat{Y} = Xb$ wtedy

$$S_e^2 = S^2 = \frac{e^T e}{n - (k+1)} = \frac{\sum_{i=1}^n e_i^2}{n - (k+1)} = \frac{1}{n - (k+1)} (Y^T Y - b^T X^T Y)$$

Współczynnik determinacji

$$R^2 = \frac{b^T X^T Y - n \bar{Y}^2}{Y^T Y - n \bar{Y}^2} \quad R^2 \in \langle 0, 1 \rangle$$

Standardowe błędy oszacowania parametrów strukturalnych

Rozpatrujemy macierz $S_e^2 (X^T X)^{-1}$,

na głównej przekątnej tej macierzy mamy wariancje tych błędów tzn. $S^2(b_i)$, $i = 0, 1, \dots, k$.

Zatem

$$S(b_i) = \sqrt{S^2(b_i)}, \quad i = 0, 1, \dots, k.$$

Prognoza.

Prognoza punktowa.

Niech

$$x_\tau = [1 \quad x_{\tau 1} \quad \dots \quad x_{\tau k}]^T$$

będzie wektorem którego pierwszą składową jest 1, a pozostałe to przewidywane wartości kolejnych zmiennych X_i w okresie prognozy τ .

Prognoza punktowa y_τ^* to przewidywana wartość Y odpowiadająca wartościom $x_{\tau i}$ zmiennych X_i .

$$y_\tau^* = x_\tau^T b$$

Standardowy błąd prognozy

$$S_\tau = S_e \sqrt{1 + x_\tau^T (X^T X)^{-1} x_\tau}$$

Uwaga: zawsze $s_\tau > s_e$.

Zatem należy traktować wartość prognozy jako $y_\tau^* \pm s_\tau$

Jakość prognozy punktowej możemy ocenić względnym błędem prognozy punktowej

$$\delta_{punkt} = \frac{s_\tau}{y_\tau^*} \cdot 100\%$$

Prognoza przedziałowa.

Prognoza przedziałowa dla poziomu ufności $1 - \alpha$.

$$\left\langle y_\tau^* - u_\alpha S_\tau; y_\tau^* + u_\alpha S_\tau \right\rangle$$

gdzie u_α odczytujemy z tablicy rozkładu Studenta:

$$P\left(|T_{n-(k+1)}| > u_\alpha\right) = \alpha.$$

Jakość prognozy przedziałowej możemy ocenić względnym błędem prognozy przedziałowej

$$\delta_{prz} = \frac{u_\alpha S_\tau}{y_\tau^*} \cdot 100\%$$