



EKONOMETRIA

Temat wykładu:

*Liniowy model ekonometryczny (regresji)
z wieloma zmiennymi objaśniającymi –
estymacja parametrów*

Prowadzący: dr inż. Zbigniew TARAPATA
e-mail: Zbigniew.Tarapata@isi.wat.edu.pl
http:// zbigniew.tarapata.akcja.pl/p_ekonometria/
tel.: 0-606-45-54-80

Plan wykładu

- Wprowadzenie w zagadnienia weryfikacji modeli regresji liniowej
- Zapis modelu regresji wielu zmiennych
- Metoda najmniejszych kwadratów
- Przykład estymacji parametrów modelu
- Mierniki jakości modelu
- Przykład liczenia mierników jakości modelu

Wprowadzenie w zagadnienia weryfikacji modeli regresji liniowej

Model ekonometryczny z wieloma zmiennymi objaśniającymi:

$$y_t = f_t(x_{1t}, x_{2t}, \dots, x_{Kt}, \xi_t), \quad t=1, 2, \dots, n$$

- y - zmienna objaśniana, regresant;
- x_{it} - i -ta zmienna objaśniająca ($i=1, \dots, K$) w chwili t , regresor;
- ξ - składnik losowy;
- t - numer kolejnej obserwacji;
- K - liczba zmiennych objaśniających.

Zapis modelu regresji wielu zmiennych

W dalszych rozważaniach dotyczących ekonometrycznych modeli jednorównaniowych, ograniczymy się do funkcji regresji liniowej :

$$y_t = \sum_{i=0}^K \alpha_i \cdot x_{it} + \xi_t \quad (1)$$

gdzie:

y_t – wartość zmiennej objaśnianej (endogenicznej) w chwili t ($t = \overline{1, n}$);

x_{it} – wartość i -tej zmiennej objaśniającej w chwili t , przy czym $x_{0t}=1$;

α_i – parametry modelu, $i = \overline{0, K}$;

n – liczba obserwacji.

Estymacja parametrów modelu - wprowadzenie

Zakładamy, że w modelu (1) spełnione są założenia Gaussa-Markowa (patrz wykład nr 3, slajdy 8-11) definiujące

standardowy model liniowy

Wówczas parametry modelu (1) estymuje się klasyczną metodą najmniejszych kwadratów (dla modelu z wieloma zmiennymi objaśniającymi).

Estymacja parametrów modelu – metodą najmniejszych kwadratów (MNK)

Funkcja kryterium, którą należy zminimalizować, ma postać (2):

$$\Psi = \sum_{t=1}^n e_t^2 = \sum_{t=1}^n (y_t - \hat{y}_t)^2 = \sum_{t=1}^n (y_t - a_0 - a_1 \cdot x_{1t} - a_2 \cdot x_{2t} - \dots - a_K \cdot x_{Kt})^2$$

e_t – reszta modelu, ;

$a_0, a_1, \dots, a_K$ – estymatory (oszacowania) parametrów odpowiednio $\alpha_0, \alpha_1, \dots, \alpha_K$ modelu (1);

y_t – wartość rzeczywista (empiryczna) zmiennej endogenicznej w chwili t ;

$\hat{y}_t$ – wartość teoretyczna (z modelu) zmiennej endogenicznej w chwili t ;

$x_{1t}, x_{2t}, \dots, x_{Kt}$ – wartości zmiennych egzogenicznych (objaśniających) w chwili t ;

Estymacja parametrów modelu – metodą najmniejszych kwadratów (MNK), c.d.

W przypadku dużych modeli wygodniej jest przedstawić funkcję kryterium (1) w zapisie macierzowym:

$$\begin{aligned}\Psi &= \mathbf{e}^T \cdot \mathbf{e} = (\mathbf{y} - \mathbf{X} \cdot \mathbf{a})^T \cdot (\mathbf{y} - \mathbf{X} \cdot \mathbf{a}) = \\ &= \mathbf{y}^T \cdot \mathbf{y} - \mathbf{a}^T \cdot \mathbf{X}^T \cdot \mathbf{y} - \mathbf{y}^T \cdot \mathbf{X} \cdot \mathbf{a} + \mathbf{a}^T \cdot \mathbf{X}^T \cdot \mathbf{X} \cdot \mathbf{a} = \\ &= \mathbf{y}^T \cdot \mathbf{y} - 2 \cdot \mathbf{a}^T \cdot \mathbf{X}^T \cdot \mathbf{y} + \mathbf{a}^T \cdot \mathbf{X}^T \cdot \mathbf{X} \cdot \mathbf{a}\end{aligned}\quad (2)$$

gdzie:

$\mathbf{e}$ – wektor reszt modelu:

UWAGA!

$\mathbf{y}^T$ – oznacza transponowanie wektora $\mathbf{y}$
(analogicznie w przypadku macierzy $\mathbf{X}^T$
wektora $\mathbf{a}^T$, $\mathbf{e}^T$)

$$\mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

$$\text{, przy czym } \mathbf{e}^T \cdot \mathbf{e} = \sum_{t=1}^n e_t^2$$

Estymacja parametrów modelu – metodą najmniejszych kwadratów (MNK), c.d.

$\mathbf{y}$ – wektor empirycznych wartości zmiennych endogenicznych;

$\mathbf{X}$ – macierz zaobserwowanych wartości zmiennych objaśniających;

$\mathbf{a}$ – wektor ocen parametrów modelu;

Należy zauważyć, że $\hat{\mathbf{y}} = \mathbf{X} \cdot \mathbf{a}$, gdzie $\hat{\mathbf{y}}$ oznacza wektor teoretycznych wartości zmiennych endogenicznych, tzn.

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{K1} \\ 1 & x_{12} & \cdots & x_{K2} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_{1n} & \cdots & x_{Kn} \end{bmatrix}$$

$$\mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_K \end{bmatrix}$$

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix}$$

Estymacja parametrów modelu – metodą najmniejszych kwadratów (MNK), c.d.

Aby wyznaczyć wektor **a**, należy obliczyć pochodną funkcji (2) względem wektora **a** i przyrównać ją do zera. Pochodna ta jest równa:

$$\frac{\partial \Psi}{\partial \mathbf{a}} = -2 \cdot \mathbf{X}^T \cdot \mathbf{y} + 2 \cdot \mathbf{X}^T \cdot \mathbf{X} \cdot \mathbf{a} \quad (3)$$

co po przyrównaniu do zera daje:

$$-2 \cdot \mathbf{X}^T \cdot \mathbf{y} + 2 \cdot \mathbf{X}^T \cdot \mathbf{X} \cdot \mathbf{a} = \mathbf{0} \quad (4)$$

gdzie **0** jest wektorem kolumnowym o $(K+1)$ wierszach którego składowymi są zera.

Ostatecznie otrzymujemy:

$$\mathbf{X}^T \cdot \mathbf{X} \cdot \mathbf{a} = \mathbf{X}^T \cdot \mathbf{y} \quad (5)$$

Estymacja parametrów modelu – metodą najmniejszych kwadratów (MNK), c.d.

Po lewostronnym pomnożeniu lewej i prawej strony równania (5) przez $(\mathbf{X}^T \cdot \mathbf{X})^{-1}$ otrzymujemy wzór na wektor **a** ocen parametrów:

$$\mathbf{a} = (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \cdot \mathbf{y} \quad (6)$$

gdzie:

$$\mathbf{X}^T \cdot \mathbf{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{12} & \cdots & x_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ x_{K1} & x_{K2} & \cdots & x_{Kn} \end{bmatrix} \cdot \begin{bmatrix} 1 & x_{11} & \cdots & x_{K1} \\ 1 & x_{12} & \cdots & x_{K2} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_{1n} & \cdots & x_{Kn} \end{bmatrix} = \begin{bmatrix} n & \sum_{t=1}^n x_{1t} & \cdots & \sum_{t=1}^n x_{Kt} \\ \sum_{t=1}^n x_{1t} & \sum_{t=1}^n x_{1t}^2 & \cdots & \sum_{t=1}^n x_{1t} \cdot x_{Kt} \\ \vdots & \vdots & \cdots & \vdots \\ \sum_{t=1}^n x_{Kt} & \sum_{t=1}^n x_{Kt} \cdot x_{1t} & \cdots & \sum_{t=1}^n x_{Kt}^2 \end{bmatrix} \quad (7)$$

Estymacja parametrów modelu – metodą najmniejszych kwadratów (MNK), c.d.

Oraz

$$\mathbf{X}^T \cdot \mathbf{y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{12} & \cdots & x_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ x_{K1} & x_{K2} & \cdots & x_{Kn} \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n x_{1t} \cdot y_t \\ \vdots \\ \sum_{t=1}^n x_{Kt} \cdot y_t \end{bmatrix} \quad (8)$$

Estymacja parametrów modelu – metodą najmniejszych kwadratów (MNK), c.d. Przykład 1

Weźmy pod uwagę następujący model:

$$y_t = \alpha_0 + \alpha_1 \cdot x_{1t} + \alpha_2 \cdot x_{2t} + \zeta_t$$

gdzie:

y_t – koszt produkcji pewnego wyrobu
w chwili t ;

x_{1t} – wielkość produkcji w chwili t ;

x_{2t} – wielkość zatrudnienia w chwili t ;

Wartości zmiennych w kolejnych 10-ciu
latach przedstawia tabela.

t	y_t	x_{1t}	x_{2t}
1	1,4	1	5
2	1,7	1,5	4
3	1,5	1	5
4	2	1,5	3,5
5	2,2	2	3,5
6	2,3	2,5	3,5
7	2,5	2	2
8	2,8	3	2
9	2,6	3	2,5
10	2,9	3,5	1,5

Oszacować parametry modelu na
podstawie danych historycznych.

**Estymacja parametrów modelu –
metoda najmniejszych kwadratów (MNK), c.d.**
Przykład 1, c.d.

Rozwiązanie:

Korzystając z (7) i (8) otrzymujemy:

$$\mathbf{X}^T \cdot \mathbf{X} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & \dots & x_{1,10} \\ x_{21} & x_{22} & \dots & x_{2,10} \end{bmatrix} \cdot \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{1,10} & x_{2,10} \end{bmatrix} = \begin{bmatrix} 10 & \sum_{t=1}^{10} x_{1t} & \sum_{t=1}^{10} x_{2t} \\ \sum_{t=1}^{10} x_{1t} & \sum_{t=1}^{10} x_{1t}^2 & \sum_{t=1}^{10} x_{1t} \cdot x_{2t} \\ \sum_{t=1}^{10} x_{2t} & \sum_{t=1}^{10} x_{1t} \cdot x_{2t} & \sum_{t=1}^{10} x_{2t}^2 \end{bmatrix}$$

$$\mathbf{X}^T \cdot \mathbf{y} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & \dots & x_{1,10} \\ x_{21} & x_{22} & \dots & x_{2,10} \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{10} \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^{10} y_t \\ \sum_{t=1}^{10} x_{1t} \cdot y_t \\ \sum_{t=1}^{10} x_{2t} \cdot y_t \end{bmatrix}$$

**Estymacja parametrów modelu –
metoda najmniejszych kwadratów (MNK), c.d.**
Przykład 1, c.d.

Dla naszego przykładu mamy:

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 5 \\ 1 & 1,5 & 4 \\ 1 & 1 & 5 \\ 1 & 1,5 & 3,5 \\ 1 & 2 & 3,5 \\ 1 & 2,5 & 3,5 \\ 1 & 2 & 2 \\ 1 & 3 & 2 \\ 1 & 3 & 2,5 \\ 1 & 3,5 & 1,5 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} 1,4 \\ 1,7 \\ 1,5 \\ 2 \\ 2,2 \\ 2,3 \\ 2,5 \\ 2,8 \\ 2,6 \\ 2,9 \end{bmatrix}$$

Estymacja parametrów modelu – metoda najmniejszych kwadratów (MNK), c.d. Przykład 1, c.d.

t	y_t	x_{1t}	x_{2t}	x_{1t}^2	$x_{1t} \cdot x_{2t}$	x_{2t}^2	$x_{1t} \cdot y_t$	$x_{2t} \cdot y_t$
1	1,4	1	5	1	5	25	1,4	7
2	1,7	1,5	4	2,25	6	16	2,55	6,8
3	1,5	1	5	1	5	25	1,5	7,5
4	2	1,5	3,5	2,25	5,25	12,25	3	7
5	2,2	2	3,5	4	7	12,25	4,4	7,7
6	2,3	2,5	3,5	6,25	8,75	12,25	5,75	8,05
7	2,5	2	2	4	4	4	5	5
8	2,8	3	2	9	6	4	8,4	5,6
9	2,6	3	2,5	9	7,5	6,25	7,8	6,5
10	2,9	3,5	1,5	12,25	5,25	2,25	10,15	4,35
Suma ->	21,9	21	32,5	51	59,75	119,2	49,95	65,5

$$X^T \cdot y = \begin{bmatrix} \sum_{t=1}^{10} y_t \\ \sum_{t=1}^{10} x_{1t} \cdot y_t \\ \sum_{t=1}^{10} x_{2t} \cdot y_t \end{bmatrix} =$$

$$= X^T \cdot y = \begin{bmatrix} 21,9 \\ 49,95 \\ 65,5 \end{bmatrix}$$

$$X^T \cdot X = \begin{bmatrix} 10 & \sum_{t=1}^{10} x_{1t} & \sum_{t=1}^{10} x_{2t} \\ \sum_{t=1}^{10} x_{1t} & \sum_{t=1}^{10} x_{1t}^2 & \sum_{t=1}^{10} x_{1t} \cdot x_{2t} \\ \sum_{t=1}^{10} x_{2t} & \sum_{t=1}^{10} x_{1t} \cdot x_{2t} & \sum_{t=1}^{10} x_{2t}^2 \end{bmatrix} = \begin{bmatrix} 10 & 21 & 32,5 \\ 21 & 51 & 59,75 \\ 32,5 & 59,75 & 119,25 \end{bmatrix}$$

Z. Tarapata, Ekonometria, wykład nr 5

15

Estymacja parametrów modelu – metoda najmniejszych kwadratów (MNK), c.d. Przykład 1, c.d.

Odwracając macierz $X^T \cdot X$ otrzymamy:

$$(X^T \cdot X)^{-1} = \begin{bmatrix} 11,54 & -2,58 & -1,85 \\ -2,58 & 0,63 & 0,39 \\ -1,85 & 0,39 & 0,32 \end{bmatrix}$$

Korzystając z (6) możemy wyznaczyć wektor a ocen parametrów $\alpha_0, \alpha_1, \alpha_2$ modelu :

$$a = (X^T \cdot X)^{-1} \cdot X^T \cdot y = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 11,54 & -2,58 & -1,85 \\ -2,58 & 0,63 & 0,39 \\ -1,85 & 0,39 & 0,32 \end{bmatrix} \cdot \begin{bmatrix} 21,9 \\ 49,95 \\ 65,5 \end{bmatrix} = \begin{bmatrix} 2,459 \\ 0,263 \\ -0,253 \end{bmatrix}$$

Oszacowany model przybiera więc postać: $\hat{y}_t = 2,459 + 0,263 \cdot x_{1t} - 0,253 \cdot x_{2t}$

Z. Tarapata, Ekonometria, wykład nr 5

16

Mienniki jakości modelu

+ Wariancja s^2 składnika resztowego modelu:

$$s^2 = \frac{1}{n-K} \cdot \sum_{t=1}^n (y_t - \hat{y}_t)^2 = \frac{1}{n-K} \cdot (\mathbf{y}^T \cdot \mathbf{y} - \mathbf{y}^T \cdot \mathbf{X} \cdot \mathbf{a}) \quad (9)$$

+ Odchylenie standardowe s składnika resztowego modelu:

$$s = \sqrt{s^2} \quad (10)$$

+ Współczynnik determinacji R^2 :

$$R^2 = 1 - \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2} = 1 - \frac{\mathbf{y}^T \cdot \mathbf{y} - \mathbf{y}^T \cdot \mathbf{X} \cdot \mathbf{a}}{(\mathbf{y} - \bar{\mathbf{y}})^T \cdot (\mathbf{y} - \bar{\mathbf{y}})} \quad R^2 \in [0,1] \quad (11)$$

Im R^2 bliższe jedności, tym model jest lepiej dopasowany.

Mienniki jakości modelu, c.d.

+ Macierz wariancji – kowariancji estymatorów parametrów:

$$\mathbf{D}^2(\mathbf{a}) = s^2 \cdot (\mathbf{X}^T \cdot \mathbf{X})^{-1} \quad (12)$$

Szczególne znaczenie mają elementy znajdujące się na głównej przekątnej tej macierzy. Określają one wariancje estymatorów parametrów. Pierwiastki z nich, to błędy średnie szacunku parametrów modelu. Poza główną przekątną znajdują się kowariancje estymatorów parametrów.

+ Współczynnik wyrazistości w (zmienności) reszt modelu:

$$w = \frac{s}{\bar{y}} \cdot 100\% \quad (13)$$

Mierniki jakości modelu

Przykład 2

Oceń jakość dopasowania modelu

$$\hat{y}_t = 2,459 + 0,263 \cdot x_{1t} - 0,253 \cdot x_{2t}$$

oszacowanego w Przykładzie 1.

Wykorzystać wyliczoną tam macierz

$$(\mathbf{X}^T \cdot \mathbf{X})^{-1} = \begin{bmatrix} 11,54 & -2,58 & -1,85 \\ -2,58 & 0,63 & 0,39 \\ -1,85 & 0,39 & 0,32 \end{bmatrix}$$

Mierniki jakości modelu

Przykład 2, c.d.

t	y_t	$\hat{y}_t$	$y_t - \hat{y}_t$	$(y_t - \hat{y}_t)^2$	$y_t - \bar{y}$	$(y_t - \bar{y})^2$
1	1,4	1,457	-0,057	0,0032	-0,79	0,6241
2	1,7	1,842	-0,142	0,0202	-0,49	0,2401
3	1,5	1,457	0,043	0,0018	-0,69	0,4761
4	2	1,968	0,032	0,001	-0,19	0,0361
5	2,2	2,1	0,1	0,01	0,01	0,0001
6	2,3	2,231	0,069	0,0048	0,11	0,0121
7	2,5	2,479	0,021	0,0004	0,31	0,0961
8	2,8	2,742	0,058	0,0034	0,61	0,3721
9	2,6	2,616	-0,016	0,0003	0,41	0,1681
10	2,9	3	-0,1	0,01	0,71	0,5041
Suma	21,9	21,89	0,008	0,0551	0	2,529

$$R^2 = 1 - \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2} = 1 - \frac{0,0551}{2,529} = 1 - 0,0218 = 0,9782$$

$$s^2 = \frac{1}{n-K} \cdot \sum_{t=1}^n (y_t - \hat{y}_t)^2 = \frac{1}{7} \cdot 0,0551 = 0,0069$$

$$s = \sqrt{s^2} = \sqrt{0,0069} = 0,0831$$

Mierniki jakości modelu

Przykład 2, c.d.

t	y_t	$\hat{y}_t$	$y_t - \hat{y}_t$	$(y_t - \hat{y}_t)^2$	$y_t - \bar{y}$	$(y_t - \bar{y})^2$
1	1,4	1,457	-0,057	0,0032	-0,79	0,6241
2	1,7	1,842	-0,142	0,0202	-0,49	0,2401
3	1,5	1,457	0,043	0,0018	-0,69	0,4761
4	2	1,968	0,032	0,001	-0,19	0,0361
5	2,2	2,1	0,1	0,01	0,01	0,0001
6	2,3	2,231	0,069	0,0048	0,11	0,0121
7	2,5	2,479	0,021	0,0004	0,31	0,0961
8	2,8	2,742	0,058	0,0034	0,61	0,3721
9	2,6	2,616	-0,016	0,0003	0,41	0,1681
10	2,9	3	-0,1	0,01	0,71	0,5041
Suma	21,9	21,89	0,008	0,0551	0	2,529

$$w = \frac{s}{y} \cdot 100\% = \frac{0,0831}{2,19} \cdot 100\% \approx 3,8\%$$

Wnioski

R^2 jest bliskie jedności, więc model jest bardzo dobrze dopasowany do danych empirycznych. Świadczy o tym również bardzo mała wartość wariancji s^2 składnika resztowego, jak również bardzo małe wartości elementów macierzy wariancji-kowariancji.

$$\mathbf{D}^2(\mathbf{a}) = s^2 \cdot (\mathbf{X}^T \cdot \mathbf{X})^{-1} = 0,0069 \cdot \begin{bmatrix} 11,54 & -2,58 & -1,85 \\ -2,58 & 0,63 & 0,39 \\ -1,85 & 0,39 & 0,32 \end{bmatrix} = \begin{bmatrix} 0,0796 & -0,0178 & -0,0127 \\ -0,0178 & 0,0043 & 0,0027 \\ -0,0128 & 0,0027 & 0,0022 \end{bmatrix}$$