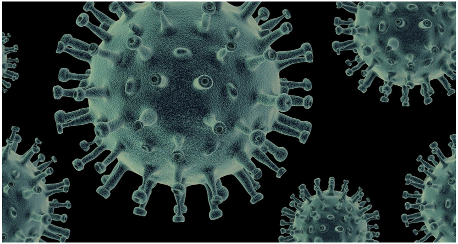
Model regresji wielorakiej

15 marca 2021



Rozważmy następujący model:

$$Y_t = \beta_0 + \beta_1 X_{t,1} + \beta_2 X_{t,2} + \ldots + \beta_k X_{t,k} + \epsilon_t,$$

dla $t = 1, 2, \ldots, n$, gdzie

- $X_{t,1}, X_{t,2}, \ldots, X_{t,k}$ - zaobserwowane wartości zmiennych objaśniających
- Y_t - zaobserwowane wartości zmiennych objaśnianych
- ϵ_t - nieznana wartość składnika losowego
- $\beta_0, \beta_1, \ldots, \beta_k$ - nieznane parametry.

Model liniowy

Model regresji wielorakiej

$$Y_t = \beta_0 + \beta_1 X_{t,1} + \beta_2 X_{t,2} + \ldots + \beta_k X_{t,k} + \epsilon_t,$$

może być przedstawiony jako układ równań liniowych:

$$\begin{cases} Y_1 & = & \beta_0 + \beta_1 X_{1,1} + \beta_2 X_{1,2} + \ldots + \beta_k X_{1,k} + \epsilon_1 \\ Y_2 & = & \beta_0 + \beta_1 X_{2,1} + \beta_2 X_{2,2} + \ldots + \beta_k X_{2,k} + \epsilon_2 \\ \vdots & \vdots & \vdots \\ Y_n & = & \beta_0 + \beta_1 X_{n,1} + \beta_2 X_{n,2} + \ldots + \beta_k X_{n,k} + \epsilon_n \end{cases}$$

lub w formie macierzowej:

$$\underbrace{\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix}}_{\mathbf{X}} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}}_{\beta} + \underbrace{\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}}_{\epsilon}.$$

Innymi słowy, zapiszemy model wg poniższego wzoru:

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon.$$

- **Argumenty wejściowe:** dla każdego $t = 1, 2..., n$ znamy wartości zmiennych objaśniających $X_{t,1}$, $X_{t,2}$, ...,$X_{t,k}$ oraz objaśnianych Y_t;
- **Parametry zakłócające:** nie znamy wartości ϵ_t, ale one również wpływają na wynik obserwacji;
- **Argumenty wyjściowe:** staramy się znaleźć $\beta_0, \beta_1, \ldots, \beta_k$ używając **metodę najmniejszych kwadratów**: znaleźć $\beta_0, \beta_1, \ldots, \beta_k \in \mathbb{R}$ dla których poniższe wyrażenie

$$f(\beta) = \sum_{t=1}^{n} \left(Y_t - \beta_0 - \beta_1 X_{t,1} - \beta_2 X_{t,2} - \ldots - \beta_k X_{t,k}\right)^2$$

ma wartość najmniejszą.

Cel

Podana funkcja celu

$$f(\beta) = \sum_{t=1}^{n} \left(Y_t - \beta_0 - \beta_1 X_{t,1} - \beta_2 X_{t,2} - \ldots - \beta_k X_{t,k}\right)^2$$

może być wyrażona równoważnie w formie macierzowej:

$$f(\beta) = (\mathbf{Y} - \mathbf{X} * \beta)^T * (\mathbf{Y} - \mathbf{X} * \beta)\,.$$

Uwaga 1

Można odnieść wrażenie, że f jest funkcją o wartościach macierzowych. Jednak w rzeczywistości przyjmuje wartość liczbową. Co prawda $\mathbf{Y} - \mathbf{X} * \beta$ jest macierzą rozmiaru $n \times 1$, stąd transpozycja $(\mathbf{Y} - \mathbf{X} * \beta)^T$ jest macierzą rozmiaru $1 \times n$. Stąd $(\mathbf{Y} - \mathbf{X} * \beta)^T * (\mathbf{Y} - \mathbf{X} * \beta)$ jest "macierzą" rozmiaru 1×1. Identyfikując macierz "jednokomórkową" z jej wartością, f przyjmuje wartości liczbowe.

Definicja 1

Niech $\mathbf{A}$ *będzie macierzą o losowych komórkach tzn. wszystkie komórki* A_{ij}, $(i = 1, \ldots, p, j = 1, 2, \ldots, q)$ *są losowe*

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \ldots & A_{1q} \\ A_{21} & A_{22} & A_{23} & \ldots & A_{2q} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & A_{p3} & \ldots & A_{pq} \end{bmatrix}$$

Wtedy ***wartością oczekiwaną macierzy*** $\mathbf{A}$ *jest macierzą wartości oczekiwanych:*

$$E(\mathbf{A}) = \begin{bmatrix} E(A_{11}) & E(A_{12}) & E(A_{13}) & \ldots & E(A_{1q}) \\ E(A_{21}) & E(A_{22}) & E(A_{23}) & \ldots & E(A_{2q}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E(A_{p1}) & E(A_{p2}) & E(A_{p3}) & \ldots & E(A_{pq}) \end{bmatrix}$$

Uwaga 2

Na potrzeby ekonometrii, wektor $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_p)$ rozumiemy jako ***odpowiednia macierz kolumnowa (nie wierszowa)****, tzn.*

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_p) = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \alpha_p \end{bmatrix}.$$

Możemy zapisać również $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_p]^T$*;*

Definicja 2

Dla wektora losowego $\mathbf{A} = [A_1, A_2, \ldots, A_p]^T$ *definujemy* ***macierz kowariancji*** $D^2(\mathbf{A})$ *jako*

$$D^2(\mathbf{A}) := \begin{bmatrix} Cov(A_1, A_1) & Cov(A_1, A_2) & \ldots & Cov(A_1, A_p) \\ Cov(A_2, A_1) & Cov(A_2, A_2) & \ldots & Cov(A_2, A_p) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(A_p, A_1) & Cov(A_p, A_2) & \ldots & Cov(A_p, A_p) \end{bmatrix}.$$

Uwaga 3

Własności macierzy kowariancji

- $D^2(\mathbf{A})$ *jest* ***symetryczna*** *ponieważ* $Cov(A_i, A_j) = Cov(A_j, A_i)$ *dla* $i, j = 1, 2, \ldots, p$;
- *Przekątna macierzy* $D^2(\mathbf{A})$ *zawiera* ***wariancje****, ponieważ* $Cov(A_i, A_i) = Var(A_i)$;
- *Jeśli* $\mathbf{A}$ *jest niezdegenerowanym wektorem losowym (brak stałych składowych), wtedy* $D^2(\mathbf{A})$ *jest macierzą* ***dodatnio określoną*** *tzn. dla wszystkich* $x \in \mathbb{R}^p \setminus \{0\}$, $x^T * D^2(\mathbf{A}) * x > 0$;
- $D^2(\mathbf{A})$ *może być wyrażona jako*

$$D^2(\mathbf{A}) = E\left((\mathbf{A} - E(\mathbf{A})) * (\mathbf{A} - E(\mathbf{A}))^T\right).$$

Fakt 1

Jeśli $E(\mathbf{A}) = \mathbf{0}$ *to*

$$D^2(\mathbf{A}) := \begin{bmatrix} E(A_1^2)) & E(A_1A_2) & \dots & E(A_1A_p) \\ E(A_2A_1) & E(A_2^2) & \dots & E(A_2A_p) \\ \vdots & \vdots & \ddots & \vdots \\ E(A_pA_1) & E(A_pA_2) & \dots & E(A_p^2) \end{bmatrix} = E(\mathbf{A}\mathbf{A}^T).$$

Fact 1

Przypuśćmy, że zmienne losowe $A_1, A_2, \ldots, A_p$ *są parami nieskorelowane (tzn.* $Cov(A_i, A_j) = 0$ *dla wszystkich* $i \neq j$*), oraz* $Var(A_i) = \sigma^2$ *for* $i = 1, 2, \ldots, p$*. Wtedy*

$$D^2(\mathbf{A}) := \sigma^2 \mathbf{I}$$

gdzie

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}.$$

Założenia modelu

Założenie 1

Załóżmy, że mamy więcej obserwacji niż parametrów $k+1 \leq n$. *Ponadto*

- Z1 *Macierz* $\mathbf{X}$ *zmiennych objaśniających jest deterministyczna (nie jest otrzymana w wyniku losowania), tzn. dla dowolnego* t, *zmienne* $[X_{t,1}, X_{t,2}, ..., X_{t,k}]^T$ *nie są losowe;*
- Z2 *Rząd macierzy* $\mathbf{X}$ *jest* $k+1$, *tzn. kolumny są liniowo niezależne;*
- Z3 $E(\epsilon) = \mathbf{0}$, *tzn.* $E\epsilon_t = 0$ *dla wszystkich* t;
- Z4 $D^2(\epsilon) = E\left(\epsilon * \epsilon^T\right) = \sigma^2\mathbf{I}$, *tzn.* ϵ_t *jest ciągiem zmiennych losowych nieskorelowanych, których wariancje są takie same* σ^2;
- Z5 *Zmienne* ϵ_t *mają rozkład normalny* $N(0, \sigma^2)$;

Założenia modelu

W tym modelu
statystyk zakłada Z1-Z4,
czasami również Z5 oraz:

- wartość macierzy jest deterministyczna $\mathbf{X}$, tzn. uzyskał je jako wartość zdeterminowaną (w uproszczeniu nie jest wynikiem losowania);
- wartość wektora $\mathbf{Y}$ jest losowa (w uproszczeniu uzyskana jest w wyniku losowania);

Statystyk nie zna

- wartości (deterministicznej) wektora β i próbuje je oszacować (jest to jego przedmiot zainteresowań);
- realizacji, czyli wylosowanej wartości wektora losowego ϵ, nie zna nawet wariancji (σ^2), ale nie jest to jego przedmiot zainteresowań, jest to **parametr zakłócający**.

Estymacja β

Funkcję celu

$$f(\beta) = \sum_{t=1}^{n} \left(Y_t - \beta_0 - \beta_1 X_{t,1} - \beta_2 X_{t,2} - \ldots - \beta_k X_{t,k} \right)^2$$

można zapisać w formie macierzowej:

$$f(\beta) = (\mathbf{Y} - \mathbf{X} * \beta)^T * (\mathbf{Y} - \mathbf{X} * \beta) .$$

Funkcja celu jest **wypukła**, stąd minimalizacja f sprowadza się do przyrównanie **gradientu** wektora zerowego. Przypomnę, że

$$grad(f) = \begin{bmatrix} \frac{\partial f}{\partial \beta_0} \\ \frac{\partial f}{\partial \beta_1} \\ \vdots \\ \frac{\partial f}{\partial \beta_k} \end{bmatrix}$$

Estymacja β

Zauważmy, że dla każdego wektora $\mathbf{A} = [a_1, a_2, \ldots, a_p]^T$ zachodzi

$$||\mathbf{A}||^2 = \sum_{i=1}^{p} a_i^2 = \mathbf{A}^T * \mathbf{A}.$$

Można więc przekształcić funkcję celu:

$$\begin{aligned}
f(\beta) &= (\mathbf{Y} - \mathbf{X} * \beta)^T * (\mathbf{Y} - \mathbf{X} * \beta) \\
&= \left(\mathbf{Y}^T - (\mathbf{X} * \beta)^T\right) * (\mathbf{Y} - \mathbf{X} * \beta) \\
&= \left(\mathbf{Y}^T - \beta^T * \mathbf{X}^T\right) * (\mathbf{Y} - \mathbf{X} * \beta) \\
&= \mathbf{Y}^T * \mathbf{Y} - \mathbf{Y}^T * \mathbf{X} * \beta - \beta^T * \mathbf{X}^T * \mathbf{Y} + \beta^T * \mathbf{X}^T * \mathbf{X} * \beta \\
&= ||\mathbf{Y}||^2 - \mathbf{Y}^T * \mathbf{X} * \beta - \left(\mathbf{Y}^T * \mathbf{X} * \beta\right)^T + ||\mathbf{X} * \beta||^2.
\end{aligned}$$

Zauważmy, że $\mathbf{Y}^T * \mathbf{X} * \beta$ jest macierzą jeden na jeden. Ten fakt wynika z tego, że $\mathbf{Y}$ jest macierzą rozmiaru $n \times 1$, stąd $\mathbf{Y}^T$ ma rozmiar $1 \times n$. Dalej $\mathbf{X}$ jest macierzą $n \times k$, stąd $\mathbf{Y}^T * \mathbf{X}$ jest macierzą $1 \times k$. Zatem ponieważ β ma rozmiar $k \times 1$, stąd $\mathbf{Y}^T * \mathbf{X} * \beta$ jest rozmiaru 1×1 - *macierz jednokomórkowa*:

$$\left(\underbrace{\mathbf{Y}^T}_{1\times n} * \underbrace{\mathbf{X}}_{n\times k} * \underbrace{\beta}_{k\times 1} \right) \Rightarrow \left(\underbrace{\mathbf{Y}^T * \mathbf{X} * \beta}_{\text{jest } 1\times 1} \right).$$

Transpozycja macierzy 1×1 jest również macierzą 1×1 i ma tą samą wartość

$$\left(\underbrace{\mathbf{Y}^T * \mathbf{X} * \beta}_{\text{jest } 1\times 1} \right)^T = \left(\underbrace{\mathbf{Y}^T * \mathbf{X} * \beta}_{\text{jest } 1\times 1} \right).$$

Estymacja β

Stąd otrzymujemy

$$\begin{aligned} f(\beta) &= ||\mathbf{Y}||^2 - \underbrace{\left(\mathbf{Y}^T * \mathbf{X} * \beta + \left(\mathbf{Y}^T * \mathbf{X} * \beta\right)^T\right)}_{\text{Oba składniki są równe}} \\ &\quad + ||\mathbf{X} * \beta||^2 \\ &= ||\mathbf{Y}||^2 - 2\mathbf{Y}^T * \mathbf{X} * \beta + ||\mathbf{X} * \beta||^2 \end{aligned}$$

Estymacja β

Zatem

$$\begin{aligned} grad(f) &= grad\left(||\mathbf{Y}||^2 - 2\mathbf{Y}^T * \mathbf{X} * \beta + ||\mathbf{X} * \beta||^2\right) \\ &= grad(||\mathbf{Y}||^2) - grad\left(2\mathbf{Y}^T * \mathbf{X} * \beta\right) + grad(||\mathbf{X} * \beta||^2) \\ &= \underbrace{-2\mathbf{X}^T * \mathbf{Y} + 2\mathbf{X}^T * \mathbf{X} * \beta}_{\text{Uwierz na słowo}}. \end{aligned}$$

Teraz otrzymujemy

$$\begin{aligned} (grad(f) = \mathbf{0}) &\Leftrightarrow (-2\mathbf{X}^T * \mathbf{Y} + 2\mathbf{X}^T * \mathbf{X} * \beta = \mathbf{0} \,|: 2) \\ &\Leftrightarrow -\mathbf{X}^T * \mathbf{Y} + \mathbf{X}^T * \mathbf{X} * \beta = \mathbf{0} \\ &\Leftrightarrow (\mathbf{X}^T * \mathbf{X}) * \beta = \mathbf{X}^T * \mathbf{Y}. \end{aligned}$$

Estymacja β

Mamy więc

$$\underbrace{(\mathbf{X}^T * \mathbf{X})}_{\text{Z założenia Z2 macierz jest nieosobliwa}} * \beta = \mathbf{X}^T * \mathbf{Y}$$

Zatem otrzymujemy jedyny estymator metody **najmniejszych kwadratów**

$$\hat{\beta} = (\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T * \mathbf{Y}.$$

Estymator $\hat{\beta}$

Twierdzenie 1 (Gaussa-Markova)

Estymator wyrażony wzorem

$$\hat{\beta} = (\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T * \mathbf{Y}.$$

jest jedynym estymatorem otrzymanym metodą ***najmniejszych kwadratów****. Ponadto*

- *jest liniowy, tzn. jest przekształceniem liniowym wektora losowego* $\mathbf{Y}$*;*
- *jest nieobciążony* $E(\hat{\beta}) = \beta$*;*
- *jest zgodny tzn. jeśli estymator* $\hat{\beta}^n$ *bazuje na n obserwacjach*

$$\forall_{\delta>0} P(|\hat{\beta}^n - \beta| > \delta) \to 0 \textit{ o ile } n \to 0;$$

- *jest efektywnym: tzn. ma najmniejszą wariancję ze wszystkich estymatorów liniowych i nieobciążonych.*

$\hat{\beta}$ jest nieobciążnony tzn. $E(\hat{\beta}) = \beta$

Model jest postaci

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon,$$

gdzie $\mathbf{X}$ jest macierzą deterministyczną spełniaącą (Z2) oraz $E(\epsilon) = \mathbf{0}$ (Z3). Stąd

$$\begin{aligned} E(\hat{\beta}) &= E(\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T * \mathbf{Y}) \\ &= (\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T * E(\mathbf{Y}) \\ &= (\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T * E(\mathbf{X}\beta + \epsilon) \\ &= (\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T (\mathbf{X}\beta + \mathbf{0}) \\ &= (\mathbf{X}^T * \mathbf{X})^{-1} * \left(\mathbf{X}^T\mathbf{X}\right)\beta \\ &= \mathbf{I}\beta = \beta \end{aligned}$$

Twierdzenie 2

Macierz kowariancji $\hat{\beta}$ *spełnia*

$$D^2(\hat{\beta}) = \sigma^2(\mathbf{X}^T\mathbf{X})^{-1}.$$

Dowód twierdzenia 2

Ponieważ $\hat{\beta}$ jest estymatorem nieobciążonym, stąd $E(\hat{\beta}) = \beta$ oraz

$$D^2(\hat{\beta}) := E(\hat{\beta} - E(\hat{\beta}))(\hat{\beta} - E(\hat{\beta}))^T = E(\hat{\beta} - \beta) * (\hat{\beta} - \beta)^T.$$

Ponadto

$$\begin{aligned}
\mathbf{Y} = \mathbf{X}\beta + \epsilon \quad & \Rightarrow \quad \mathbf{X}^T * \mathbf{Y} = \mathbf{X}^T * \mathbf{X} * \beta + \mathbf{X}^T * \epsilon \\
& \Leftrightarrow \quad \mathbf{X}^T(\mathbf{Y} - \epsilon) = (\mathbf{X}^T * \mathbf{X}) * \beta \\
& \Leftrightarrow \quad \beta = (\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T(\mathbf{Y} - \epsilon).
\end{aligned}$$

Stąd

$$\begin{aligned}\hat{\beta}-\beta &= (\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T\mathbf{Y} - (\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T(\mathbf{Y}-\epsilon) \\ &= (\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T * \epsilon.\end{aligned}$$

Stąd

$$
\begin{aligned}
D^2(\hat{\beta}) &= E\left(\left((\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T * \epsilon\right) * \left((\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T * \epsilon\right)^T\right) \\
&= E\left(\left((\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T * \epsilon\right) * \left(\epsilon^T * \mathbf{X} * (\mathbf{X}^T * \mathbf{X})^{-1}\right)\right) \\
&= E\left(\underbrace{(\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T}_{\text{m. deterministyczna}} * (\epsilon * \epsilon^T) * \underbrace{\mathbf{X} * (\mathbf{X}^T * \mathbf{X})^{-1}}_{\text{m. deterministyczna}}\right) \\
&= (\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T * \underbrace{E(\epsilon * \epsilon^T)}_{\text{By Z4 } E(\epsilon*\epsilon^T)=\sigma^2\mathbf{I}} * \mathbf{X} * (\mathbf{X}^T * \mathbf{X})^{-1} \\
&= (\mathbf{X}^T * \mathbf{X})^{-1} * \mathbf{X}^T * \sigma^2\mathbf{I} * \mathbf{X} * (\mathbf{X}^T * \mathbf{X})^{-1} \\
&= \sigma^2 \underbrace{(\mathbf{X}^T * \mathbf{X})^{-1} * (\mathbf{X}^T * \mathbf{X})}_{\text{m. odwrotne}} * (\mathbf{X}^T * \mathbf{X})^{-1} = \sigma^2 (\mathbf{X}^T * \mathbf{X})^{-1}.
\end{aligned}
$$

Przykład

Przykład 1 (Ekonometria, Przykład 2.1. Gruszczyński i Podgórska, Szkoła Główna Handlowa w Warszawie)

Na podstawie danych umieszczonych na następnym slajdzie dopasujemy model między poziomem (stopą) inflacji a poziomem (stopą) bezrobocia. Znajdziemy następnie model postaci:

$$Y_t = \beta_0 + \beta_1 X_{t,1} + \beta_2 X_{t,2} + \epsilon_t$$

gdzie Y_t*-poziom inflacji,* $X_{t,1}$ *poziom bezrobocia i* $X_{t,2}$ *oczekiwany poziom inflacji na lata o indeksie* $t = 1, 2, ..., 15$.

Dane

inflacja (objaśniana)	bezrobocie (objaśniająca)	oczekiwana inflacja (objaśniająca)
5.92	4.9	4.78
4.30	5.9	3.84
3.30	5.6	3.13
6.23	4.9	3.44
10.97	5.6	6.84
9.14	8.5	9.47
5.77	7.7	6.51
6.45	7.1	5.92
7.60	6.1	6.08
11.47	5.8	8.09

Przykład - rozwiąnie

$$\mathbf{X} = \begin{bmatrix} 1 & 4.9 & 4.78 \\ 1 & 5.9 & 3.84 \\ 1 & 5.6 & 3.13 \\ 1 & 4.9 & 3.44 \\ 1 & 5.6 & 6.84 \\ 1 & 8.5 & 9.47 \\ 1 & 7.7 & 6.51 \\ 1 & 7.1 & 5.92 \\ 1 & 6.1 & 6.08 \\ 1 & 5.8 & 8.09 \end{bmatrix} \qquad \mathbf{Y} = \begin{bmatrix} 5.92 \\ 4.30 \\ 3.30 \\ 6.23 \\ 10.97 \\ 9.14 \\ 5.77 \\ 6.45 \\ 7.60 \\ 11.47 \end{bmatrix}$$

Przykład - rozwiązanie

$$\mathbf{X}^T * \mathbf{X} =$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 4.9 & 5.9 & 5.6 & 4.9 & 5.6 & 8.5 & 7.7 & 7.1 & 6.1 & 5.8 \\ 4.78 & 3.84 & 3.13 & 3.44 & 6.84 & 9.47 & 6.51 & 5.92 & 6.08 & 8.09 \end{bmatrix} \begin{bmatrix} 1 & 4.9 & 4.78 \\ 1 & 5.9 & 3.84 \\ 1 & 5.6 & 3.13 \\ 1 & 4.9 & 3.44 \\ 1 & 5.6 & 6.84 \\ 1 & 8.5 & 9.47 \\ 1 & 7.7 & 6.51 \\ 1 & 7.1 & 5.92 \\ 1 & 6.1 & 6.08 \\ 1 & 5.8 & 8.09 \end{bmatrix}.$$

$$= \begin{bmatrix} 10.000 & 62.100 & 58.100 \\ 62.100 & 398.350 & 375.430 \\ 58.100 & 375.430 & 375.532 \end{bmatrix}$$

Przykład-rozwiązanie

Z dokładnością do błędu zaokrągleń mamy:

$$(\mathbf{X}^T\mathbf{X})^{-1} = \begin{bmatrix} 3.219132 & -0.561521 & 0.063325 \\ -0.561521 & 0.141384 & -0.054471 \\ 0.063325 & -0.054471 & 0.047322 \end{bmatrix}$$

Przykład - rozwiązanie

$$\mathbf{X}^T\mathbf{Y} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 4.9 & 5.9 & 5.6 & 4.9 & 5.6 & 8.5 & 7.7 & 7.1 & 6.1 & 5.8 \\ 4.78 & 3.84 & 3.13 & 3.44 & 6.84 & 9.47 & 6.51 & 5.92 & 6.08 & 8.09 \end{bmatrix} \begin{bmatrix} 5.92 \\ 4.30 \\ 3.30 \\ 6.23 \\ 10.97 \\ 9.14 \\ 5.77 \\ 6.45 \\ 7.60 \\ 11.47 \end{bmatrix}$$

$$= \begin{bmatrix} 71.150 \\ 445.617 \\ 452.907 \end{bmatrix}$$

Następnie podstawiamy do wzoru

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} =$$

$$= \begin{bmatrix} 3.219132 & -0.561521 & 0.063325 \\ -0.561521 & 0.141384 & -0.054471 \\ 0.063325 & -0.054471 & 0.047322 \end{bmatrix} \begin{bmatrix} 71.150 \\ 445.617 \\ 452.907 \end{bmatrix} = \begin{bmatrix} 7.4979 \\ -1.6192 \\ 1.6648 \end{bmatrix}$$

Daje to następujące oszacowanie modelu:

$$\underbrace{Y_t}_{\text{inflacja}} = 7.4979 - 1.6192 * \underbrace{X_{t,2}}_{\text{bezrobocie}} + 1.6648 * \underbrace{X_{t,2}}_{\text{oczekiwana inflacja}}$$

$$+\epsilon_t$$