Rozwiązania zadań z listy do wykładu 5 (Interpretacja geometryczna)

Rozwiązanie zadania 1.

Oznaczmy przez $\mathbb{X}$ i $\mathbb{Z}$ macierze o trzech kolumnach $\mathbb{X}=[x_1\ x_2\ x_3]$ $\mathbb{Z}=[z_1\ z_2\ z_3]$. odpowiadających wektorom modelu pierwszego i drugiego z zadania. Przypomnijmy związki między tymi wektorami

$$
\begin{aligned}
z_1 &= x_1 - 2x_2 \\
z_2 &= x_2 + 4x_3 \qquad\qquad (*) \\
z_3 &= 2x_1 - 3x_2 + 5x_3.
\end{aligned}
$$

Biorąc ten związek pod uwagę otrzymujemy

$$\mathbb{Z} = \mathbb{X}\mathbb{A}$$

gdzie

$$\mathbb{A} = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 1 & -3 \\ 0 & 4 & 5 \end{bmatrix}$$

Rozwiązując układ równań (*) możemy wyrazić kolumny macierzy $\mathbb{X}$ w postaci kombinacji liniowych kolumn macierzy $\mathbb{Z}$.

$$
\begin{aligned}
x_1 &= 17z_1 + 10z_2 - 8z_3 \\
x_2 &= 8z_1 + 5z_2 - 4z_3 \\
x_3 &= -2z_1 - z_2 + z_3
\end{aligned}
$$

Tak więc macierz $\mathbb{A}$ jest odwracalna i jej macierz odwrotna jest postaci

$$\mathbb{A}^{-1} = \begin{bmatrix} 17 & 8 & -2 \\ 10 & 5 & -1 \\ -8 & -4 & 1 \end{bmatrix}$$

Oba modele regresji z zadania dają w wyniku te same dopasowania parametrów i te same wartości reszt ponieważ

$$P_{\mathbb{X}} = P_{\mathbb{Z}} \quad \text{oraz} \quad M_{\mathbb{X}} = M_{\mathbb{Z}}.$$

a więc dopasowanie $P_{\mathbb{X}} y$ dla pierwszego modelu są takie same jak dopasowanie $P_{\mathbb{Z}} y$ dla drugiego modelu i podobnie dla reszt w obu modelach.

Ponieważ dopasowania w obu modelach są takie same więc

$$\mathbb{X}\mathbb{A}\hat{\alpha} = \mathbb{X}\hat{\beta}$$

co implikuje $\hat{\beta} = \mathbb{A}\hat{\alpha}$ i $\hat{\alpha} = \mathbb{A}^{-1}\hat{\beta}$

Te związki między $\hat{\beta}$ i $\hat{\alpha}$ pozwalają nam wyrazić elementy wektora $\hat{\beta}$ w terminach elementów wektora $\hat{\alpha}$ i elementy wektora $\hat{\alpha}$ w terminach elementów wektora $\hat{\beta}$.

W szczególności

$$\hat{\beta}_1 = \hat{\alpha}_1 + 2\hat{\alpha}_3$$

$$\hat{\alpha}_1 = 17\hat{\beta}_1 + 8\hat{\beta}_2 - 2\hat{\beta}_3.$$

Rozwiązanie zadania 2.

Estymator parametru β_2 w modelu z zadania

$$y = X_1\beta_1 + X_2\beta_2 + u \qquad (*)$$

jest postaci

$$\hat{\beta}_2 = (X_2^T M_1 X_2)^{-1} X_2^T M_1 y$$

a reszta jest postaci

$$M_X y = M_1(I - X_2(X_2^T M_1 X_2)^{-1} X_2^T M_1) y$$

gdzie P_1 oznacza projekcję na podprzestrzeń rozpiętą przez X_1 a $M_1 = I - P_1$.

Dla (1) ani estymator ani reszta nie będą takie jak dla modelu (*) ponieważ X_2 nie rozpina tej samej przestrzeni co X.

Dla (2) i (3) estymatory nie są takie same. Reszta najprawdopodobniej nie będzie taka sama bo $P_1 y$ nie będzie raczej równe y. W przypadku modelu (3) reszty są równe zero ponieważ zmienna objaśniana i zmienne objaśniające leżą w przestrzeni liniowej rozpiętej przez X_1.

Dla (4) estymatory są takie same jak w modelu wyjściowym (*)

$$\tilde{\beta} = (X^T X)^{-1} X^T P_X y = (X^T X)^{-1} X^T y$$

Reszty nie będą takie same bo model (4) idealnie dopasowuje się do danych.

Dla (5) ani estymatory ani reszty nie są takie same. Dla estymatorów możemy to sprawdzić gdyż

$$\tilde{\beta}_2 = (X_2^T X_2)^{-1} X_2^T P_X y = (X_2^T X_2)^{-1} X_2^T y$$

co oczywiście nie jest takie samo jak w przypadku modelu (*)

Dla (6) również estymatory i reszty nie są takie same. Estymator dla (6) ma postać

$$\tilde{\beta}_2 = (X_2^T X_2)^{-1} X_2^T M_1 y$$

~~istotnie~~ różną od estymatora dla (*)

Dla (7) estymatory i reszty są takie same jak dla (*)

Dla (8) otrzymamy ten sam wynik jak dla modelu (7). Dodatkowo X_1 jest ortogonalny zarówno do zmiennej objaśnianej i zmiennej $M_1 X_2$. Tak więc $\tilde{\beta}_1 = 0$ i $\tilde{\beta}_2$ oraz reszty nie zmieniają się gdy dodamy do modelu X_1 jako zmienną objaśniającą.

Dla (9) zarówno $\tilde{\beta}_2$ jak i reszty są takie same jak w (*).

Dla (10) reszty nie będą najprawdopodobniej takie same jak dla (*) ponieważ zmienną objaśnianą jest $P_X y$. Estymatory jednak będą takie same. Wynika to z faktu, że

$$P_1 P_X \overset{(1)}{=} P_1 \quad i \quad X_2 P_X \overset{(2)}{=} X_2$$

a więc

$$\begin{aligned}
\tilde{\beta} &= (X_2^T M_1 X_2)^{-1} X_2^T M_1 P_X y \\
&= (X_2^T M_1 X_2)^{-1} X_2^T (I - P_1) P_X y \\
&\overset{(1)}{=} (X_2^T M_1 X_2)^{-1} X_2^T (P_X - P_1) y \\
&\overset{(2)}{=} (X_2^T M_1 X_2)^{-1} (X_2^T - X_2^T P_1) y \\
&= (X_2^T M_1 X_2)^{-1} X_2^T M_1 y
\end{aligned}$$