Twierdzenie Frischa-Waugha

Marcin Pawlik

16 marca 2006 roku

Spis treści

1 Rzutowanie 2

2 Twierdzenie Frischa-Waugha 3

3 Przykład liczbowy 4

1 Rzutowanie

Wektor reszt empirycznych dany jest wzorem:

$$\mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\beta} \tag{1}$$

podstawiając $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ otrzymujemy:

$$\mathbf{e} = \mathbf{y} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \left(\mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right)\mathbf{y} = \mathbf{M}\mathbf{y} \tag{2}$$

Macierz $\mathbf{M}_{\mathbf{n}\times\mathbf{n}}$ jest symetryczna:

$$\begin{aligned} \mathbf{M}' &= \left(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right)' = \\ &= \mathbf{I}' - \left(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right)' = \\ &= \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{M} \end{aligned} \tag{3}$$

oraz idempotentna (tzn. pomnożona przez samą siebie daje w wyniku samą siebie):

$$\begin{aligned} \mathbf{M}\mathbf{M} &= \left(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right)\left(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right) = \\ &= \mathbf{I} - \mathbf{2}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' + \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \\ &= \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{M} \end{aligned} \tag{4}$$

Macierz $\mathbf{M}$ można traktować jako macierz, która przez wetkor $\mathbf{y}$ daje wektor reszt empirycznych z regresji $\mathbf{y}$ względem $\mathbf{X}$. Ponadto zachodzi następująca zależność:

$$\mathbf{M}\mathbf{X} = \mathbf{0} \tag{5}$$

Jedną z interpretacji tego wyniku jest fakt, że regresja $\mathbf{X}$ względem $\mathbf{X}$ da zerowy wektor reszt. Wartości otrzymane w wyniku regresji można zdekomponować na wartości obserwowane oraz reszty empiryczne:

$$\hat{\mathbf{y}} = \mathbf{y} - \mathbf{e} = \mathbf{y} - \mathbf{M}\mathbf{y} = (\mathbf{I} - \mathbf{M})\mathbf{y} = \mathbf{P}\mathbf{y} \tag{6}$$

gdzie:

$$\mathbf{P} = \mathbf{I} - \mathbf{M} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \tag{7}$$

Macierz $\mathbf{P}$ jest macierzą projekcji (inaczej rzutowania) wektora $\mathbf{y}$ na podprzestrzeń rozpiętą przez $\mathbf{X}$. $\mathbf{P}$ jest macierzą skonstruowaną z macierzy $\mathbf{X}$ taką, że po przemnożeniu przez $\mathbf{y}$ daje w wyniku wartości regresji $\mathbf{y}$ względem $\mathbf{X}$. Podobnie jak $\mathbf{M}\mathbf{P}$ jest symetryczna i idempotentna:

$$\begin{aligned} \mathbf{P}' = & \left(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right)' = \\ = & \left(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right) = \mathbf{P} \end{aligned} \tag{8}$$

$$\begin{aligned}\mathbf{PP} &= \left(\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\right)\left(\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\right) = \\ &= \left(\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\right) = \mathbf{P}\end{aligned} \tag{9}$$

Macierze **M** i **P** są ortogonalne:

$$\mathbf{PM} = \mathbf{MP} = \mathbf{0} \tag{10}$$

$$\mathbf{PM} = \quad \mathbf{P}\left(\mathbf{I} - \mathbf{P}\right) = \mathbf{P} - \mathbf{P^2} \quad = \mathbf{P} - \mathbf{P} = \mathbf{0} \tag{11}$$

$$\mathbf{MP} = \quad \mathbf{M}\left(\mathbf{I} - \mathbf{M}\right) = \mathbf{M} - \mathbf{M^2} \quad = \mathbf{M} - \mathbf{M} = \mathbf{0} \tag{12}$$

Ponadto:

$$\mathbf{PX} = \mathbf{X}, \tag{13}$$

co tłumaczy się tym, że rzut **X** na podprzestrzeń rozpiętą przez **X** daje **X**.

Naturalną konsekwencją (2) i (7) jest stwierdzenie, że MNK dzieli **y** na dwie ortogonalne części:

$$\mathbf{y} = \mathbf{Py} + \mathbf{My} = rzut + reszty \tag{14}$$

2 Twierdzenie Frischa-Waugha

Załóżmy, że dokonujemy regresji **y** względem dwóch zbiorów zmiennych — $\mathbf{X_1}$ i $\mathbf{X_2}$:

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon = \mathbf{X_1}\beta_\mathbf{1} + \mathbf{X_2}\beta_\mathbf{2} + \varepsilon. \tag{15}$$

Jaka jest algebraiczna postać estymatora $\hat{\beta}_\mathbf{2}$ parametru $\beta_\mathbf{2}$? Równania normalne dla (15) można zapisać w następujący sposób:

$$\begin{array}{c} \text{i} \\ \text{ii} \end{array} \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} X_1'y \\ X_2'y \end{bmatrix} \tag{16}$$

$$\begin{aligned}\mathbf{X_1'X_1}\hat{\beta}_\mathbf{1} + \mathbf{X_1'X_2}\hat{\beta}_\mathbf{2} &= \mathbf{X_1'y} \\ \mathbf{X_1'X_1}\hat{\beta}_\mathbf{1} &= \mathbf{X_1'y} - \mathbf{X_1'X_2}\hat{\beta}_\mathbf{2} \\ \hat{\beta}_\mathbf{1} &= \left(\mathbf{X_1'X_1}\right)^{-1}\mathbf{X_1'}\left(\mathbf{y} - \mathbf{X_2}\hat{\beta}_\mathbf{2}\right)\end{aligned} \tag{17}$$

Wynik ten pokazuje, że $\hat{\beta}_1$ można otrzymać dokonując regresji reszt z regresji y względem X_2 względem X_1. Warto zwrócić uwagę na pewien ciekawy fakt. Jeśli $X_1'X_2 = 0$ (czyli X_1 i X_2 są ortogonalne), wtedy $\hat{\beta}_1$ się redukuje do:

$$\hat{\beta}_\mathbf{1} = \left(\mathbf{X_1'X_1}\right)^{-1}\mathbf{X_1y} \tag{18}$$

W takim przypadku $\hat{\beta}_1$ jest wektorem parametrów regresji y względem X_1. Prowadzi to do następującego twierdzenia:

Twierdzenie 2.1 *(Orthogonal Partitioned Regression) W regresji wielorakiej* $\mathbf{y}$ *względem dwóch zbiorów zmiennych —* $\mathbf{X_1}$ *i* $\mathbf{X_2}$*, jeśli te dwa zbiory są ortogonalne, to „osobne" wektory współczynników można otrzymać w wyniku „osobnych" regresji* $\mathbf{y}$ *względem* $\mathbf{X_1}$ *i* $\mathbf{y}$ *względem* $\mathbf{X_2}$*.*

Podstawienie (17) do (15) daje następujący wynik:

$$(\mathbf{X_2'X_1})\left(\mathbf{X_1'X_1}\right)^{-1}\mathbf{X_1'}\left(\mathbf{y}-\mathbf{X_2}\hat{\beta}_\mathbf{2}\right)+\mathbf{X_2'X_2'}\hat{\beta}_\mathbf{2}=\mathbf{X_2'y}$$

$$\begin{aligned}
\mathbf{X_2'}\left[\mathbf{I}-\mathbf{X_1}\left(\mathbf{X_1'X_1}\right)^{-1}\mathbf{X_1'}\right]\mathbf{X_2}\hat{\beta}_\mathbf{2} &= \mathbf{X_2'}\left[\mathbf{I}-\mathbf{X_1}\left(\mathbf{X_1'X_1}\right)^{-1}\mathbf{X_1'}\right]\mathbf{y} \\
\mathbf{X_2'M_1X_2}\hat{\beta}_\mathbf{2} &= \mathbf{X_2'M_1y} \\
\hat{\beta}_\mathbf{2} &= \left(\mathbf{X_2'M_1X_2}\right)^{-1}\mathbf{X_2'M_1y} \qquad (19)
\end{aligned}$$

Korzystając z faktu, że $\mathbf{M_1}$ jest idempotentna i symetryczna oraz podstawiając $\mathbf{X_2^*}=\mathbf{M_1X_2}$ i $\mathbf{y}^*=\mathbf{M_1y}$ otrzymujemy:

$$\hat{\beta}_\mathbf{2}=\left(\mathbf{X_2^{*\prime}X_2^*}\right)^{-1}\mathbf{X_2^*y^*} \qquad (20)$$

Wynik ten ma zasadnicze znaczenie w analizie regresji.

Twierdzenie 2.2 *(Frisch-Waugh [1933]) W regresji MNK wektora* $\mathbf{y}$ *względem dwóch zbiorów zmiennych* $\mathbf{X_1}$ *i* $\mathbf{X_2}$ *podwektor* $\hat{\beta}_\mathbf{2}$ *jest zbiorem współczynników otrzymanych z regresji reszt z regresji* $\mathbf{y}$ *względem* $\mathbf{X_1}$ *względem zbioru reszt pochodzących z regresji każdej kolumny* $\mathbf{X_2}$ *względem* $\mathbf{X_1}$*.*

3 Przykład liczbowy

Poniżej przykład liczbowy pokazujący, że powyższe twierdzenie naprawdę „działa".

W przykładzie wykorzystano dane z podręcznika Ramathana dołączone do pakietu Gretl[1] dotyczące kosztów naprawy pojazdów w serwisach Toyoty.

Regresja zmiennej *cost* względem pozostałych daje następujące wyniki:

$$\widehat{\text{cost}} = \underset{(0{,}229)}{26,1888} + \underset{(10{,}094)}{28,0163}\,\text{age} - \underset{(-7{,}475)}{\mathbf{154,635}}\,\text{miles}$$

$$T=57 \quad \bar{R}^2=0,9492 \quad F(2,54)=523,8 \quad \hat{\sigma}=259,55$$

(statystyka t w nawiasach)

Następnie oszacujemy parametr przy zmiennej *miles* (wytłuszczony) zgodnie z metodą z twierdzenia Frischa-Waugha. Pierwszy krok polega na oszacowaniu równania regresji *cost* względem stałej i zmiennej *age*. Wynik obliczeń

[1] GNU Regression, Econometrics and Time Series Library

przedstawia się następująco:

$$\widehat{\text{cost}} = \underset{(-6{,}010)}{-625,935} + \underset{(22{,}281)}{7,34348}\,\text{age}$$

$$T = 57 \quad \bar{R}^2 = 0,8985 \quad F(1,55) = 496,46 \quad \hat{\sigma} = 366,85$$

$$\text{(statystyka } t \text{ w nawiasach)}$$

Reszty z tego równania zapisujemy jako zmienną $uhat2$.

W kolejnym kroku szacujemy regresję zmiennej $miles$ względem stałej i zmiennej age:

$$\widehat{\text{miles}} = \underset{(8{,}781)}{4,21719} + \underset{(87{,}961)}{0,133688}\,\text{age}$$

$$T = 57 \quad \bar{R}^2 = 0,9928 \quad F(1,55) = 7737,2 \quad \hat{\sigma} = 1,6917$$

$$\text{(statystyka } t \text{ w nawiasach)}$$

Reszty z tego równania zapisujemy jako zmienną $uhat3$.

Ostatnim etapem tego ćwiczenia jest oszacowanie równania regresji $uhat3$ względem $uhat2$:

$$\widehat{\text{uhat2}} = \underset{(-7{,}612)}{-\mathbf{154,635}}\,\text{uhat3}$$

$$T = 57 \quad \bar{R}^2 = 0,5085 \quad F(1,56) = 57,938 \quad \hat{\sigma} = 254,88$$

$$\text{(statystyka } t \text{ w nawiasach)}$$

Na zakończenie cieszymy się, że nam wyszło. Wyróżniona wartość współczynnika jest równa tej z pierwszego modelu.

Poniżej zapis sesji Gretla:

```
open /usr/local/share/gretl/data/data3-7.gdt
ols cost const age miles
ols cost const age
(* commands pertaining to model 2 *)
genr uhat2 = $uhat
ols miles const age
(* commands pertaining to model 3 *)
genr uhat3 = $uhat
ols uhat2 uhat3
```

Literatura

[1] W.H. Greene: *Econometric Analysis*, Prentice Hall, Englewood Cliffs 2000.