Exercice 013 - Solutions

GSF-6053

Hiver 2025

Énoncé

Supposons que le modèle de régression présente une hétéroscédasticité telle que :

$$Var(\epsilon_i) = \sigma^2 X_i' X_i$$

où X_i est le vecteur des variables explicatives pour l'observation i.

Données :

Considérez les données suivantes pour $n = 4$ observations et $k = 2$ variables explicatives :

$$X = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix}, \quad y = \begin{pmatrix} 1 \\ 2 \\ 2.5 \\ 4 \end{pmatrix}$$

Questions :

1. **a)** Estimez les coefficients β par MCO et calculez les résidus $\hat{\epsilon}$.
2. **b)** Estimez la variance $Var(\epsilon_i)$ pour chaque observation à partir des résidus obtenus.
3. **c)** Formulez la matrice de poids W pour la méthode FGLS en utilisant les estimations de variance obtenues.
4. **d)** Estimez les coefficients β par FGLS en appliquant les poids obtenus.
5. **e)** Comparez les estimations obtenues par MCO et FGLS. Discutez de l'efficacité des deux méthodes dans ce contexte hétéroscédastique.

Solution

Question 2.a : Estimation des coefficients β par MCO et calcul des résidus $\hat{\epsilon}$

L'estimateur OLS est donné par :

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'y$$

Calcul de $X'X$ et $X'y$:

$$\begin{aligned} X'X &= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 10 \\ 10 & 30 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} X'y &= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 2.5 \\ 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 + 2 + 2.5 + 4 \\ 1 \times 1 + 2 \times 2 + 3 \times 2.5 + 4 \times 4 \end{pmatrix} \\ &= \begin{pmatrix} 9.5 \\ 28.5 \end{pmatrix} \end{aligned}$$

Calcul de $(X'X)^{-1}$:

$$\begin{aligned} (X'X)^{-1} &= \frac{1}{(4)(30) - (10)^2} \begin{pmatrix} 30 & -10 \\ -10 & 4 \end{pmatrix} \\ &= \frac{1}{120 - 100} \begin{pmatrix} 30 & -10 \\ -10 & 4 \end{pmatrix} \\ &= \frac{1}{20} \begin{pmatrix} 30 & -10 \\ -10 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1.5 & -0.5 \\ -0.5 & 0.2 \end{pmatrix} \end{aligned}$$

$$\begin{aligned}\hat{\beta}_{OLS} &= \begin{pmatrix} 1.5 & -0.5 \\ -0.5 & 0.2 \end{pmatrix} \begin{pmatrix} 9.5 \\ 28.5 \end{pmatrix} \\ &= \begin{pmatrix} 1.5 \times 9.5 + (-0.5) \times 28.5 \\ -0.5 \times 9.5 + 0.2 \times 28.5 \end{pmatrix} \\ &= \begin{pmatrix} 14.25 - 14.25 \\ -4.75 + 5.7 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0.95 \end{pmatrix}\end{aligned}$$

Calcul des résidus $\hat{\epsilon}$:

$$\hat{\epsilon} = y - X\hat{\beta}_{OLS}$$

$$\begin{aligned}X\hat{\beta}_{OLS} &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 0.95 \end{pmatrix} \\ &= \begin{pmatrix} 0 + 0.95 \\ 0 + 1.9 \\ 0 + 2.85 \\ 0 + 3.8 \end{pmatrix} \\ &= \begin{pmatrix} 0.95 \\ 1.9 \\ 2.85 \\ 3.8 \end{pmatrix}\end{aligned}$$

$$\hat{\epsilon} = \begin{pmatrix} 1 \\ 2 \\ 2.5 \\ 4 \end{pmatrix} - \begin{pmatrix} 0.95 \\ 1.9 \\ 2.85 \\ 3.8 \end{pmatrix} = \begin{pmatrix} 0.05 \\ 0.10 \\ -0.35 \\ 0.20 \end{pmatrix}$$

Question 2.b : Estimation de la variance $Var(\epsilon_i)$ pour chaque observation

$$Var(\epsilon_i) = \sigma^2 X_i' X_i$$

Estimation de σ^2 :

$$
\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{n-k}\sum_{i=1}^{n}\hat{\epsilon}_i^2 \\
&= \frac{1}{4-2}\left(0.05^2 + 0.10^2 + (-0.35)^2 + 0.20^2\right) \\
&= \frac{1}{2}(0.0025 + 0.01 + 0.1225 + 0.04) \\
&= \frac{1}{2} \times 0.175 \\
&= 0.0875
\end{aligned}
$$

Calcul de $Var(\epsilon_i)$:

$$Var(\epsilon_i) = 0.0875 \times X_i' X_i$$

Pour chaque observation i :

$$
\begin{aligned}
Var(\epsilon_1) &= 0.0875 \times (1 + 1^2) = 0.175 \\
Var(\epsilon_2) &= 0.0875 \times (1 + 2^2) = 0.4375 \\
Var(\epsilon_3) &= 0.0875 \times (1 + 3^2) = 0.875 \\
Var(\epsilon_4) &= 0.0875 \times (1 + 4^2) = 1.4875
\end{aligned}
$$

Question 2.c : Formulation de la matrice de poids W pour la méthode FGLS

La matrice de poids W est l'inverse de la matrice de variance-covariance des erreurs. Puisque $Var(\epsilon_i)$ est hétéroscédastique et non corrélée, W est diagonale avec des éléments inverses des variances estimées :

$$W = \begin{pmatrix} \frac{1}{0.175} & 0 & 0 & 0 \\ 0 & \frac{1}{0.4375} & 0 & 0 \\ 0 & 0 & \frac{1}{0.875} & 0 \\ 0 & 0 & 0 & \frac{1}{1.4875} \end{pmatrix} = \begin{pmatrix} 5.7143 & 0 & 0 & 0 \\ 0 & 2.2857 & 0 & 0 \\ 0 & 0 & 1.1429 & 0 \\ 0 & 0 & 0 & 0.6727 \end{pmatrix}$$

Étapes détaillées :

$$
\begin{gathered}
Var(\epsilon_i) = \sigma^2 X_i' X_i \\
\hat{\sigma}^2 = 0.0875 \\
Var(\epsilon_i) = 0.0875 \times X_i' X_i
\end{gathered}
$$

Pour chaque observation i :

$$\begin{aligned}
Var(\epsilon_1) &= 0.0875 \times (1 + 1^2) = 0.175 \\
Var(\epsilon_2) &= 0.0875 \times (1 + 2^2) = 0.4375 \\
Var(\epsilon_3) &= 0.0875 \times (1 + 3^2) = 0.875 \\
Var(\epsilon_4) &= 0.0875 \times (1 + 4^2) = 1.4875
\end{aligned}$$

Ainsi, la matrice de poids W est :

$$W = \begin{pmatrix} \frac{1}{0.175} & 0 & 0 & 0 \\ 0 & \frac{1}{0.4375} & 0 & 0 \\ 0 & 0 & \frac{1}{0.875} & 0 \\ 0 & 0 & 0 & \frac{1}{1.4875} \end{pmatrix} = \begin{pmatrix} 5.7143 & 0 & 0 & 0 \\ 0 & 2.2857 & 0 & 0 \\ 0 & 0 & 1.1429 & 0 \\ 0 & 0 & 0 & 0.6727 \end{pmatrix}$$

Question 2.d : Estimation des coefficients β par FGLS

L'estimateur FGLS est donné par :

$$\hat{\beta}_{FGLS} = \left(X'WX\right)^{-1} X'Wy$$

Calcul de $X'WX$ et $X'Wy$:

$$X' = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \quad W = \begin{pmatrix} 5.7143 & 0 & 0 & 0 \\ 0 & 2.2857 & 0 & 0 \\ 0 & 0 & 1.1429 & 0 \\ 0 & 0 & 0 & 0.6727 \end{pmatrix}$$

$$\begin{aligned}
X'W &= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 5.7143 & 0 & 0 & 0 \\ 0 & 2.2857 & 0 & 0 \\ 0 & 0 & 1.1429 & 0 \\ 0 & 0 & 0 & 0.6727 \end{pmatrix} \\
&= \begin{pmatrix} 5.7143 & 2.2857 & 1.1429 & 0.6727 \\ 5.7143 & 4.5714 & 3.4287 & 2.6908 \end{pmatrix}
\end{aligned}$$

$$X'WX = \begin{pmatrix} 5.7143 & 2.2857 & 1.1429 & 0.6727 \\ 5.7143 & 4.5714 & 3.4287 & 2.6908 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix}$$
$$= \begin{pmatrix} 5.7143 + 2.2857 + 1.1429 + 0.6727 & 5.7143 + 4.5714 + 3.4287 + 2.6908 \\ 5.7143 + 4.5714 + 3.4287 + 2.6908 & 5.7143 + 9.1428 + 10.2861 + 10.7632 \end{pmatrix}$$
$$= \begin{pmatrix} 9.8147 & 16.4042 \\ 16.4042 & 35.9064 \end{pmatrix}$$

Calcul de $X'Wy$:

$$X'Wy = \begin{pmatrix} 5.7143 & 2.2857 & 1.1429 & 0.6727 \\ 5.7143 & 4.5714 & 3.4287 & 2.6908 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 2.5 \\ 4 \end{pmatrix}$$
$$= \begin{pmatrix} 5.7143 \times 1 + 2.2857 \times 2 + 1.1429 \times 2.5 + 0.6727 \times 4 \\ 5.7143 \times 1 + 4.5714 \times 2 + 3.4287 \times 2.5 + 2.6908 \times 4 \end{pmatrix}$$
$$= \begin{pmatrix} 5.7143 + 4.5714 + 2.8573 + 2.6908 \\ 5.7143 + 9.1428 + 8.5718 + 10.7632 \end{pmatrix}$$
$$= \begin{pmatrix} 15.8338 \\ 34.1911 \end{pmatrix}$$

Calcul de $(X'WX)^{-1}$:

$$\begin{aligned} \det(X'WX) &= (9.8147)(35.9064) - (16.4042)^2 \\ &= 352.1434 - 268.9846 \\ &= 83.1588 \end{aligned}$$

$$\begin{aligned} (X'WX)^{-1} &= \frac{1}{83.1588} \begin{pmatrix} 35.9064 & -16.4042 \\ -16.4042 & 9.8147 \end{pmatrix} \\ &\approx \begin{pmatrix} 0.4320 & -0.1973 \\ -0.1973 & 0.1180 \end{pmatrix} \end{aligned}$$

$$
\begin{aligned}
\hat{\beta}_{FGLS} &= \begin{pmatrix} 0.4320 & -0.1973 \\ -0.1973 & 0.1180 \end{pmatrix} \begin{pmatrix} 15.8338 \\ 34.1911 \end{pmatrix} \\
&= \begin{pmatrix} 0.4320 \times 15.8338 + (-0.1973) \times 34.1911 \\ -0.1973 \times 15.8338 + 0.1180 \times 34.1911 \end{pmatrix} \\
&= \begin{pmatrix} 6.8356 - 6.7415 \\ -3.1214 + 4.0354 \end{pmatrix} \\
&= \begin{pmatrix} 0.0941 \\ 0.9140 \end{pmatrix}
\end{aligned}
$$

Question 2.e : Comparaison des estimations obtenues par MCO et FGLS

Résultats obtenus :

$$
\hat{\beta}_{OLS} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \hat{\beta}_{FGLS} = \begin{pmatrix} 0.0941 \\ 0.9140 \end{pmatrix}
$$

Discussion :

— **Précision des Estimations :** L'estimateur FGLS ajuste les poids en fonction de l'hétéroscédasticité présente, ce qui permet une meilleure précision des coefficients par rapport à l'estimateur OLS qui suppose une variance constante des erreurs.

— **Variance des Estimateurs :** FGLS réduit la variance des estimateurs en pondérant les observations selon leur variance spécifique, ce qui rend les estimations plus fiables.

— **Biais :** Dans cet exercice, les deux estimateurs sont sans biais. Cependant, FGLS offre une meilleure efficacité en termes de variance.

— **Adaptation à l'Hétéroscédasticité :** FGLS prend en compte l'hétéroscédasticité des erreurs en ajustant les poids, ce qui améliore la précision des estimations. OLS, en revanche, peut être inefficace lorsque l'hétéroscédasticité est présente.

Conclusion :

Dans ce contexte hétéroscédastique, l'estimateur FGLS offre des avantages significatifs en termes d'efficacité par rapport à l'estimateur OLS. En ajustant les poids en fonction de la variance spécifique de chaque observation, FGLS réduit la variance des estimateurs $\hat{\beta}$, conduisant à des estimations plus précises et fiables.