Université Paris I, Panthéon - Sorbonne

Première Année Master M.A.E.F. 2025 – 2026

Econométrie II

Contrôle continu n°1, mars 2026

Examen de 1h30. Tout document ou calculatrice est interdit.

On suppose que l'on a observé $Y_1, \dots, Y_n$ que l'on explique par deux groupes de variables exogènes:

- $(X_i^{(j)})_{1 \leq i \leq n}$ pour $1 \leq j \leq j_1$ et on définira la matrice $X_1 = (X_i^{(j)})_{1 \leq i \leq n,\, 1 \leq j \leq j_1}$;
- $(X_i^{(j)})_{1 \leq i \leq n}$ pour $j_1 + 1 \leq j \leq j_1 + j_2$ et on définira la matrice $X_2 = (X_i^{(j)})_{1 \leq i \leq n,\, j_1+1 \leq j \leq j_1+j_2}$;

avec $j_1, j_2 \in \mathbf{N}^*$ tels que $j_1 + j_2 \leq n$ et les $X_i^{(j)}$ étant des réels observés. Soit $X = (X_i^{(j)})_{1 \leq i \leq n,\, 1 \leq j \leq j_1+j_2}$, on supposera que le rang de X est $j_1 + j_2$. On définit $[X] = \{X\, \beta,\ \beta \in \mathbf{R}^{j_1+j_2}\}$ et de la même manière $[X_1]$ et $[X_2]$. On suppose que:

$$Y = X_1\, \theta^{(1)} + X_2\, \theta^{(2)} + \varepsilon,$$

où ε est un vecteur gaussien centré de matrice de variance-covariance $\sigma_\varepsilon^2\, I_n$, I_n étant la matrice identité de taille n, $Y = {}^t(Y_1, \dots, Y_n)$ et $\theta^{(1)} = {}^t(\theta_1^{(1)}, \dots, \theta_{j_1}^{(1)})$, $\theta^{(2)} = {}^t(\theta_1^{(2)}, \dots, \theta_{j_2}^{(2)})$.

1. Montrer que l'on peut écrire que $Y = X\, \theta + \varepsilon$. En déduire, en fonction de X et de Y, l'estimateur par moindres carrés $\widehat{\theta}$ de θ, puis les expressions des estimateurs $\widehat{\theta}^{(1)}$ et $\widehat{\theta}^{(2)}$ de $\theta^{(1)}$ et $\theta^{(2)}$ dont on précisera les lois de probabilités.

2. Démontrer que le rang de X_1 est j_1. On note $P_{[X_1]}$ la matrice de projection orthogonale sur $[X_1]$ dans $\mathbf{R}^n$. Montrer que si $U_2 = (I_n - P_{[X_1]})\, Y$ alors $U_2 = X_2'\, \theta^{(2)} + \varepsilon^{(2)}$ où l'on précisera X_2' et la loi de $\varepsilon^{(2)}$.

3. Montrer que si $Z_2 \in \mathbf{R}^{j_2}$ est tel que $X_2'\, Z_2 = 0$ alors il existe $Z_1 \in \mathbf{R}^{j_1}$ tel que $X_2\, Z_2 = X_1\, Z_1$. En déduire que le rang de X_2' est j_2. En déduire qu'un autre estimateur de $\theta^{(2)}$ est $\widetilde{\theta}^{(2)}$ tel que:
$$\widetilde{\theta}^{(2)} = \left({}^tX_2'\, X_2'\right)^{-1}\, {}^tX_2'\, U_2.$$

4. Démontrer que $\widetilde{\theta}^{(2)}$ est non biaisé et déterminer sa loi en fonction de σ_ε^2, $\theta^{(2)}$, $P_{[X_1]}$ et X_2.

5. On suppose dans cette question et uniquement dans cette question que $[X_1]$ et $[X_2]$ sont deux sous-espaces vectoriels orthogonaux. Montrer que $\widehat{\theta}^{(2)} = \widetilde{\theta}^{(2)}$.

6. On ne suppose donc plus désormais que $[X_1]$ et $[X_2]$ sont deux sous-espaces vectoriels orthogonaux. Montrer que ${}^tX\, X\, \widehat{\theta} = {}^tX\, Y$ et en déduire que $\widehat{\theta}^{(1)}$ et $\widehat{\theta}^{(2)}$ vérifient
$$\begin{cases} {}^tX_1\, X_1\, \widehat{\theta}^{(1)} + {}^tX_1\, X_2\, \widehat{\theta}^{(2)} = {}^tX_1\, Y \\ {}^tX_2\, X_1\, \widehat{\theta}^{(1)} + {}^tX_2\, X_2\, \widehat{\theta}^{(2)} = {}^tX_2\, Y \end{cases}.$$
En déduire que $\widetilde{\theta}^{(2)} = \widehat{\theta}^{(2)}$.

7. On considère les résidus $\widehat{\varepsilon}^{(2)} = Y - X_2\, \widehat{\theta}^{(2)}$. Montrer que $\widehat{\theta}^{(1)} = \left({}^tX_1\, X_1\right)^{-1}\, {}^tX_1\, \widehat{\varepsilon}^{(2)}$.

8. Soit $\widetilde{\sigma}^2 = \dfrac{1}{n - j_1 - j_2}\, \left\| Y - X_1\, \widetilde{\theta}^{(1)} - X_2\, \widetilde{\theta}^{(2)} \right\|^2$. Déterminer la loi de $\widetilde{\sigma}^2$.

9. On s'intéresse au problème de test: $H_0 :\ \theta^{(2)} = 0$ contre $H_1 :\ \theta^{(2)} \neq 0$. En utilisant $\widetilde{\theta}^{(2)}$ et $\widetilde{\sigma}^2$, proposer une statistique de test pour ce problème de test et donner sa loi sous H_0.

Proof. 1. $Y=(Y_i)$, $\varepsilon=(\varepsilon_i)$, $\theta={}^t\big({}^t\theta^{(1)},{}^t\theta^{(2)}\big)$ et $X=\big(X_1,X_2\big)$, matrice de taille (n,j_1+j_2).
Comme le rang de X est j_1+j_2, on en déduit que ${}^tX\,X$ est inversible et $\widehat\theta=\big({}^tX\,X\big)^{-1}{}^tX\,Y$.

On en déduit que $\widehat\theta^{(1)}=J_1\,\widehat\theta$ avec $J_1=\begin{pmatrix}1&0&\cdots&0&\cdots&0\\0&1&\cdots&0&\cdots&0\\\vdots&\vdots&\ddots&&&\vdots\\0&0&\cdots&1&\cdots&0\end{pmatrix}$ matrice de taille (j_1,j_1+j_2)

et $\widehat\theta^{(2)}=J_2\,\widehat\theta$ avec $J_2=\begin{pmatrix}0&\cdots&0&1&0&\cdots&0\\0&\cdots&0&0&1&\cdots&0\\\vdots&&\vdots&\vdots&\vdots&\ddots&\\0&\cdots&0&0&0&\cdots&1\end{pmatrix}$ matrice de taille (j_2,j_1+j_2).

Sous l'hypothèse ε vecteur gaussien centré de matrice de variance-covariance $\sigma^2\,I_n$, alors $\widehat\theta\overset{\mathcal L}{\sim}\mathcal N_{j_1+j_2}\big(\theta\,,\,\sigma^2_\varepsilon\,\big({}^tX\,X\big)^{-1}\big)$ d'après le cours, et $\widehat\theta^{(1)}$ et $\widehat\theta^{(2)}$ sont aussi des vecteurs gaussiens (matrices réelles multipliées par un vecteur gaussien), et on a:

$$\begin{cases}\widehat\theta^{(1)}\overset{\mathcal L}{\sim}\mathcal N_{j_1}\big(J_1\,\theta\,,\,\sigma^2_\varepsilon\,J_1\,\big({}^tX\,X\big)^{-1}\,{}^tJ_1\big)=\mathcal N_{j_1}\big(\theta^{(1)}\,,\,\sigma^2_\varepsilon\,J_1\,\big({}^tX\,X\big)^{-1}\,{}^tJ_1\big)\\ \widehat\theta^{(2)}\overset{\mathcal L}{\sim}\mathcal N_{j_2}\big(J_2\,\theta\,,\,\sigma^2_\varepsilon\,J_2\,\big({}^tX\,X\big)^{-1}\,{}^tJ_2\big)=\mathcal N_{j_2}\big(\theta^{(2)}\,,\,\sigma^2_\varepsilon\,J_2\,\big({}^tX\,X\big)^{-1}\,{}^tJ_2\big)\end{cases}$$

2. Le rang de X est plein, donc X s'écrit comme j_1+j_2 vecteurs colonne de $\mathbf R^n$ qui forme une famille libre. Donc toute sous famille également: les j_1 premiers vecteurs sont libres et donc X_1 est de rang j_1.
On a $Y=X_1\,\theta^{(1)}+X_2\,\theta^{(2)}+\varepsilon$ donc comme $(I_n-P_{[X_1]})=P_{[X_1]^\perp}$ et $X_1\,\theta^{(1)}\in[X_1]$, $(I_n-P_{[X_1]})\,Y=(I_n-P_{[X_1]})\,X_2\,\theta^{(2)}+(I_n-P_{[X_1]})\,\varepsilon$, soit
$$U_2=X_2'\,\theta^{(2)}+\varepsilon^{(2)}\quad\text{avec}\quad X_2'=(I_n-P_{[X_1]})\,X_2\text{ et }\varepsilon^{(2)}=(I_n-P_{[X_1]})\,\varepsilon.$$
Comme $(I_n-P_{[X_1]})$ matrice de réels et ε vecteur gaussien centré, la loi de $\varepsilon^{(2)}$ est $\mathcal N_n\big(0\,,\,\sigma^2_\varepsilon\,P_{[X_1]^\perp}\big)$.

3. Si $X_2'\,Z_2=0$ alors $P_{[X_1]^\perp}\,X_2\,Z_2=0$ donc $X_2\,Z_2\in[X_1]$: on en déduit l'existence de Z_1 tel que $X_2\,Z_2=X_1\,Z_1$.
De cette équation on en déduit que $X\begin{pmatrix}Z_1\\-Z_2\end{pmatrix}=0$. Comme le rang de X est j_1+j_2, l'équation $X\,\beta=0$ n'admet que 0 comme solution, d'où $Z_2=0$. Ainsi le noyau de X_2' est le vecteur nul, donc le rang de X_2' est j_2.
On a $U_2=X_2'\,\theta^{(2)}+\varepsilon^{(2)}$ avec X_2' de rang plein: on peut donc effectuer une régression linéaire de U_2 par rapport à X_2' et on a bien
$$\widetilde\theta^{(2)}=\big({}^tX_2'\,X_2'\big)^{-1}{}^tX_2'\,U_2.$$

4. En développant, on a
$$\widetilde\theta^{(2)}=\big({}^tX_2'\,X_2'\big)^{-1}{}^tX_2'\,U_2=\big({}^tX_2\,P_{[X_1]^\perp}\,X_2\big)^{-1}{}^tX_2\,P_{[X_1]^\perp}\,P_{[X_1]^\perp}\big(X_1\,\theta^{(1)}+X_2\,\theta^{(2)}+\varepsilon\big)=\theta^{(2)}+\big({}^tX_2\,P_{[X_1]^\perp}\,X_2\big)^{-1}{}^tX_2\,P_{[X_1]^\perp}\,\varepsilon.$$
D'où $\mathbb E\big[\widetilde\theta^{(2)}\big]=\theta^{(2)}$: l'estimateur n'est pas biaisé.
De la formule précédente, il est clair que la loi de $\widetilde\theta^{(2)}$ est gaussienne. On a ainsi $\widetilde\theta^{(2)}\overset{\mathcal L}{\sim}\mathcal N_{j_2}\big(\theta^{(2)}\,,\,\sigma^2_\varepsilon\,\big({}^tX_2\,P_{[X_1]^\perp}\,X_2\big)^{-1}\big)$.

5. On a alors ${}^tX_1\,X_2=0$ et $X_2\in[X_1]^\perp$. Donc $P_{[X_1]^\perp}X_2=X_2$ et ainsi $\widetilde\theta^{(2)}=\big({}^tX_2\,X_2\big)^{-1}{}^tX_2\,Y$. Mais on également ${}^tX\,X=\begin{pmatrix}{}^tX_1\,X_1&0\\0&{}^tX_2\,X_2\end{pmatrix}$, d'où $\big({}^tX\,X\big)^{-1}=\begin{pmatrix}({}^tX_1\,X_1)^{-1}&0\\0&({}^tX_2\,X_2)^{-1}\end{pmatrix}$ et ${}^tX\,Y=\begin{pmatrix}{}^tX_1\,Y\\{}^tX_2\,Y\end{pmatrix}$, ce qui entraine que $\widehat\theta=\begin{pmatrix}({}^tX_1\,X_1)^{-1}{}^tX_1\,Y\\({}^tX_2\,X_2)^{-1}{}^tX_2\,Y\end{pmatrix}=\begin{pmatrix}\widehat\theta^{(1)}\\\widehat\theta^{(2)}\end{pmatrix}$, d'où $\widehat\theta^{(2)}=\widetilde\theta^{(2)}$.

6. On sait que $X\widehat\theta\in[X]$ et $Y-X\widehat\theta\in[X]^\perp$, donc le produit scalaire entre $Y-X\widehat\theta$ et tous les vecteurs de X, et ainsi ${}^tX\,\big(Y-X\widehat\theta\big)=0$, d'où ${}^tX\,X\widehat\theta={}^tX\,Y$.
On a ${}^tX\,X=\begin{pmatrix}{}^tX_1\,X_1&{}^tX_1\,X_2\\{}^tX_2\,X_1&{}^tX_2\,X_2\end{pmatrix}$, $\widehat\theta=\begin{pmatrix}\widehat\theta^{(1)}\\\widehat\theta^{(2)}\end{pmatrix}$ et ${}^tX\,Y=\begin{pmatrix}{}^tX_1\,Y\\{}^tX_2\,Y\end{pmatrix}$, d'où ${}^tX\,X\widehat\theta={}^tX\,Y$ implique:
$$\begin{pmatrix}{}^tX_1\,X_1&{}^tX_1\,X_2\\{}^tX_2\,X_1&{}^tX_2\,X_2\end{pmatrix}\begin{pmatrix}\widehat\theta^{(1)}\\\widehat\theta^{(2)}\end{pmatrix}=\begin{pmatrix}{}^tX_1\,Y\\{}^tX_2\,Y\end{pmatrix},\quad\text{d'où}\quad\begin{cases}{}^tX_1\,X_1\,\widehat\theta^{(1)}+{}^tX_1\,X_2\,\widehat\theta^{(2)}={}^tX_1\,Y\\{}^tX_2\,X_1\,\widehat\theta^{(1)}+{}^tX_2\,X_2\,\widehat\theta^{(2)}={}^tX_2\,Y\end{cases}.$$
Dans l'équation 1, on trouve $\widehat\theta^{(1)}=\big({}^tX_1\,X_1\big)^{-1}{}^tX_1\,\big(Y-X_2\,\widehat\theta^{(2)}\big)$ et en substituant dans l'équation 2,
$${}^tX_2\,X_1\,\big({}^tX_1\,X_1\big)^{-1}{}^tX_1\,\big(Y-X_2\,\widehat\theta^{(2)}\big)+{}^tX_2\,X_2\,\widehat\theta^{(2)}={}^tX_2\,Y\quad\Longrightarrow\quad{}^tX_2\,P_{[X_1]}\,X_2\,\widehat\theta^{(2)}={}^tX_2\,\big(I_n-P_{[X_1]}\big)\,Y$$
d'où $\widetilde\theta^{(2)}=\widehat\theta^{(2)}$.

7. Il suffit de reprendre l'expression $\widehat\theta^{(1)}=\big({}^tX_1\,X_1\big)^{-1}{}^tX_1\,\big(Y-X_2\,\widehat\theta^{(2)}\big)$.

8. Il est clair que $\widetilde\sigma^2=\dfrac{1}{n-j_1-j_2}\,\big\|Y-X\,\widehat\theta\big\|^2$ d'après la question 4., donc on retrouve le résultat du cours
$$\widetilde\sigma^2\overset{\mathcal L}{\sim}\frac{\sigma^2_\varepsilon}{n-j_1-j_2}\,\chi^2(n-j_1-j_2).$$

9. Sous H_0, on a
$$\widetilde{\theta}^{(2)} \overset{\mathcal{L}}{\sim} \mathcal{N}\big(0\,,\,\sigma_\varepsilon^2\,({}^tX_2'X_2')^{-1}\big).$$
On considère la statistique
$$\widehat{F} = \frac{1}{j_2}\,\frac{1}{\widetilde{\sigma}^2}\,{}^t\widetilde{\theta}^{(2)}({}^tX_2'X_2')\widetilde{\theta}^{(2)}.$$
Alors, sous H_0, $\widehat{F} \overset{\mathcal{L}}{\sim} \mathcal{F}(j_2, n-j_1-j_2)$ car on a bien un rapport de χ^2 normalisé et on sait d'après le cours que $\widehat{\theta}$, donc $\widetilde{\theta}^{(2)}$, est indépendant de $\widetilde{\sigma}^2$.

$\square$