Université Paris I, Panthéon - Sorbonne

Première Année Master M.A.E.F. 2023 – 2024

Econométrie II

Contrôle continu n°1, mars 2024

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

Soit le modèle linéaire

$$Y = X\,\theta + \varepsilon$$

tel que:

- $Y = {}^t(Y_1, \ldots, Y_n)$ est un vecteur aléatoire observé;
- X est une matrice connue de réels, de taille (n, p) où $p \geq 2$;
- $\theta = {}^t(\theta_1, \ldots, \theta_p) \in \mathbf{R}^p$ est un vecteur de paramètres réels inconnus;
- $\varepsilon = {}^t(\varepsilon_1, \ldots, \varepsilon_n)$ est un vecteur aléatoire non observé, tel que les ε_i forme une famille de v.a.i.i.d. centrées de variance $\sigma^2 > 0$ inconnue.

Dans la suite, on notera $\|U\|^2 = {}^tU\,U$ pour $U \in \mathbf{R}^n$ et $[X] = \{X\,\theta,\, \theta \in \mathbf{R}^p\}$.

1. On suppose jusqu'à la fin de l'exercice que le rang de X est r avec $1 \leq r < p$. Donner 2 cadres réels distincts pour lesquels cette situation peut advenir **(1pt)**.
2. On définit $\ker(X) = \{\theta \in \mathbf{R}^p,\, X\,\theta = 0\}$. Démontrer que $\ker(X)$ est un sous-espace vectoriel et préciser sa dimension **(1.5pts)**.
3. Démontrer que la fonction $Y_0 \in [X] \mapsto \big\|Y - Y_0\big\|$ admet un unique minimum atteint en $Y_0 = \widehat{Y}$ que l'on précisera **(2pts)**.
4. On définit $\widehat{\theta} \in \arg\min\limits_{\theta \in \mathbf{R}^p} \big\{\big\|Y - X\,\theta\big\|\big\}$. Montrer que $\widehat{\theta}$ vérifie l'équation ${}^tX\,X\,\widehat{\theta} = {}^tX\,Y$ **(1.5pts)** et trouver l'ensemble des solutions de cette équation **(1pt)**.
5. On considère H une matrice de taille $(p - r, p)$ de rang $p - r$ telle que $\ker(H) \cap \ker(X) = \{0_p\}$ avec $\ker(H) = \{\theta \in \mathbf{R}^p,\, H\,\theta = 0\}$. On définit alors $\widetilde{\theta}$ tel que:
$$ {}^tX\,X\,\widetilde{\theta} = {}^tX\,Y \qquad \text{et} \qquad H\,\widetilde{\theta} = 0_{p-r}.$$
A titre d'exemple, et uniquement dans cette question, considérer $X = (x_{ij})_{1 \leq i \leq n,\, 1 \leq j \leq 2}$ avec $x_{i1} = 1$ et $x_{i2} = 2$ pour tout $1 \leq i \leq n$, choisir une matrice H et déterminer $\widetilde{\theta} = {}^t(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_p)$ **(2pts)**.
6. Dans le cas général, montrer que $X' = \begin{pmatrix} X \\ H \end{pmatrix}$ est une matrice de rang p **(1pt)**.
7. Demontrer que ${}^tX'\,X'\,\widetilde{\theta} = {}^tX\,Y$ **(1pt)**, en déduire que $\widetilde{\theta}$ est unique et donner son expression **(1pt)**.
8. Démontrer que $\widehat{Y} = X\,\widetilde{\theta}$ ne dépend pas de H **(0.5pts)**. En déduire que $X\,\big({}^tX'\,X'\big)^{-1}\,{}^tX$ ne dépend pas de H **(0.5pts)**.
9. Déterminer la loi de $\widetilde{\theta}$ lorsque les ε_i suivent une loi gaussienne (cadre gaussien) **(1.5pts)**.
10. Proposer, en fonction de Y, X' et X, un estimateur $\widetilde{\sigma}^2$ non biaisé de la variance σ^2 (prouver qu'il est non biaisé) **(2pts)**. Dépend-il du choix de H **(0.5pts)**? Est-il convergent si n tend vers l'infini (et p et r fixés) **(2pts)**? Dans le cadre gaussien, montrer que $\widehat{Y}$ est indépendant de $\widetilde{\sigma}^2$ **(0.5pts)**.

11. On se place dans le cadre gaussien et on note $X\,\theta = ((X\,\theta)_j)_{1\leq j\leq n}$. On veut tester l'hypothèse H_0: $(X\,\theta)_n = 0$ contre l'hypothèse H_1: $(X\,\theta)_n < 0$. Déterminer une statistique de test **(1pt)** et donner, en justifiant, sa loi sous H_0 **(1.5pts)**. Si on fixe un risque de première espèce $0 < \alpha < 1$ à ce test, déterminer la zone de rejet du test **(1pt)**.

Proof.

1. Premier cas: $n < p$. Deuxième cas: des variables exogènes colinéaires.

2. $\ker(X)$ est un sous-espace de $\mathbf{R}^p$, qui est un e.v. Il suffit de montrer que $\{0\}_i n \ker(X)$ ce qui est clairement vrai, et si $\lambda_1 \in \mathbf{R}$, $\theta_1, \theta_2 \in \ker(X)$ alors $X\,(\lambda_1\theta_1+\theta_2) = \lambda_1 X\,\theta_1 + X\,\theta_2 = 0$, donc $\lambda_1\theta_1+\theta_2 \in \ker(X)$.
Il est clair que $\dim(\ker(X)) + \mathrm{Rang}(X) = p$, puisque si u_X est l'application linéaire $\theta \in \mathbf{R}^p \mapsto X\,\theta \in \mathbf{R}^n$, alors $\dim(\ker(X)) + \mathrm{Rang}(X) = \dim(\ker(u_X)) + \dim(\Im(u_X)) = \dim(\mathbf{R}^p) = p$. Donc $\dim(\ker(X)) = p - r$.

3. Comme $[X]$ est un sev de $\mathbf{R}^n$ de dimension finie, alors $P_{[X]}Y$ projection orthogonale de Y sur $[X]$ existe est unique. Or pour $Y_0 \in [X]$,
$$\|Y - Y_0\|^2 = \|(Y - P_{[X]}Y) + (P_{[X]}Y - Y_0)\|^2 = \|P_{[X]^\perp}Y + (P_{[X]}Y - Y_0)\|^2 = \|P_{[X]^\perp}Y\|^2 + \|P_{[X]}Y - Y_0\|^2$$
par Pythagore. Donc l'unique minimum de $\|Y - Y_0\|^2$, donc de $\|Y - Y_0\|$, est atteint pour $Y_0 = \widehat{Y} = P_{[X]}Y$.

4. Si $\widehat{\theta} \in \underset{\theta\in\mathbf{R}^p}{\arg\min}\big\{\|Y - X\,\theta\|\big\}$ alors d'après la question précédente $X\,\widehat{\theta} = P_{[X]}Y$. Soit $X\,\widehat{\theta} = Y - P_{[X]^\perp}Y$ puisque l'on travaille dans $\mathbf{R}^n$ qui est de dimension finie. Donc en composant par ${}^tX$, on obtient ${}^tX\,X\,\widehat{\theta} = {}^tX\,Y - {}^tX\,P_{[X]^\perp}Y$. Mais comme $P_{[X]^\perp}Y \in [X]^\perp$, alors ${}^tX\,P_{[X]^\perp}Y$ qui est composé des produits scalaires de vecteurs engendrant $[X]$ avec le vecteur $P_{[X]^\perp}Y$ qui est dans $[X]^\perp$ est donc nul. Ainsi ${}^tX\,X\,\widehat{\theta} = {}^tX\,Y$. Il existe au moins un $\theta_0 \in \mathbf{R}^p$ tel que $P_{[X]}Y = X\,\theta_0$, car $P_{[X]}Y \in [X]$. Mais si on prend $\theta \in \ker(X)$, alors $X\,(\theta_0+\theta) = X\,\theta_0 = P_{[X]}Y$, et $\theta \notin \ker(X)$ alors $X\,(\theta_0+\theta) \neq X\,\theta_0 = P_{[X]}Y$ et l'équation n'est pas vérifiée. On en déduit que l'ensemble des solutions est $\{\theta_0+\theta, \theta\in\ker(X)\}$, espace affine de dimension $p-r$.

5. Ici on a $p = 2$ et $r = 1$. Donc H doit être de taille $(1,2)$. On peut choisir $H = (1,0)$ qui est bien de rang $2-1 = 1$. Alors $\ker(H) = \{{}^t(x,y),\, x = 0\}$, et $\ker(X) = \{{}^t(x,y),\, x+2y = 0\}$, donc on a bien $\ker(H)\cap\ker(X) = \{0_p\}$.
Si on pose $\widetilde{\theta} = {}^t(x,y)$ alors ${}^tX\,X\,\widetilde{\theta} = n(x+2y)\,{}^t(1,2)$ et ${}^tX\,Y = \Big(\sum_{k=1}^n Y_k\Big)\,{}^t(1,2)$, ce qui revient à une seule ligne $(x+2y) = \overline{Y}_n$.
Par ailleurs, si on rajoute la condition $H\,\widetilde{\theta} = 0_{p-r}$, alors on a comme condition supplémentaire $x = 0$. On obtient ainsi une unique solution $\widetilde{\theta} = {}^t(0, \frac{1}{2}\,\overline{Y}_n)$.

6. Soit $\theta \in \mathbf{R}^p$. Alors $X'\,\theta = \begin{pmatrix} X\,\theta \\ H\,\theta \end{pmatrix}$. Supposons que $\theta \in \ker(X')$. Alors $X\,\theta = 0$ et $H\,\theta = 0$, donc $\theta \in \ker(H)\cap\ker(X)$. Mais comme $\ker(H)\cap\ker(X) = \{0_p\}$ alors $\ker(X') = \{0_p\}$ et ainsi $\mathrm{Rang}(X') = p$.

7. On sait que ${}^tX\,X\,\widetilde{\theta} = {}^tX\,Y$ et $H\,\widetilde{\theta} = 0_p$. On en déduit en supperposant les équations que $X'\,\widetilde{\theta} = \begin{pmatrix} X\,\widetilde{\theta} \\ H\,\widetilde{\theta} \end{pmatrix} = \begin{pmatrix} X\,\widetilde{\theta} \\ 0_p \end{pmatrix}$. Ainsi ${}^tX'\,X'\,\widetilde{\theta} = ({}^tX, {}^tH)\begin{pmatrix} X\,\widetilde{\theta} \\ 0_p \end{pmatrix} = {}^tX\,X\,\widetilde{\theta}$. D'où le résultat.
Par suite, comme le rang de X' est p alors le rang de ${}^tX'\,X'$ est p: la matrice ${}^tX'\,X'$ est inversible et on obtient $\widetilde{\theta} = \big({}^tX'\,X'\big)^{-1}\,{}^tX\,Y$.

8. On sait que $X\,\widetilde{\theta} = P_{[X]}Y$ qui ne dépend pas de H.
Or $X\,\widetilde{\theta} = X\,\big({}^tX'\,X'\big)^{-1}\,{}^tX\,Y = P_{[X]}Y$ qui ne dépend pas de H. Donc $X\,\big({}^tX'\,X'\big)^{-1}\,{}^tX$ ne dépend pas de H.

9. Comme $\widetilde{\theta} = \big({}^tX'\,X'\big)^{-1}\,{}^tX\,Y = \big({}^tX'\,X'\big)^{-1}\,{}^tX\,X\,\theta + \big({}^tX'\,X'\big)^{-1}\,{}^tX\,\varepsilon$, et comme $\mathbb{E}[\varepsilon] = 0_n$, on en déduit que $\mathbb{E}[\widetilde{\theta}] = \big({}^tX'\,X'\big)^{-1}\,{}^tX\,X\,\theta$.
Par ailleurs, $\mathrm{cov}(\widetilde{\theta}) = \mathrm{cov}\big(\big({}^tX'\,X'\big)^{-1}\,{}^tX\,\varepsilon\big) = \sigma^2\,\big({}^tX'\,X'\big)^{-1}\,{}^tX\,X\,\big({}^tX'\,X'\big)^{-1}$.
Dans le cadre gaussien, $\widetilde{\theta}$ est un vecteur gaussien comme transformation linéaire d'un vecteur gaussien, donc
$$\widetilde{\theta} \overset{\mathcal{L}}{\sim} \mathcal{N}_p\big(\big({}^tX'\,X'\big)^{-1}\,{}^tX\,X\,\theta\,,\, \sigma^2\,\big({}^tX'\,X'\big)^{-1}\,{}^tX\,X\,\big({}^tX'\,X'\big)^{-1}\big).$$

10. On va utiliser $\widetilde{\sigma}^2 = \frac{1}{n-r}\,\|Y - \widehat{Y}\|^2 = \frac{1}{n-r}\,\|Y - X\,{}^tX'\,X')^{-1}\,{}^tX\,Y\|^2$. Il est non biaisé car:
$$\mathbb{E}\big[\widetilde{\sigma}^2\big] = \frac{1}{n-r}\,\mathbb{E}\big[\|P_{[X]^\perp}\varepsilon\|^2\big] = \frac{1}{n-r}\,\mathbb{E}\big[{}^t\varepsilon\,P_{[X]^\perp}\varepsilon\big] = \frac{1}{n-r}\,\mathbb{E}\big[\mathrm{Trace}\big({}^t\varepsilon\,P_{[X]^\perp}\varepsilon\big)\big] = \frac{1}{n-r}\,\mathrm{Trace}\big(\mathbb{E}\big[P_{[X]^\perp}\varepsilon\,{}^t\varepsilon\big]\big) = \mathrm{Trace}\Big(\frac{\sigma^2}{n-r}\,P_{[X]^\perp}\Big) = \sigma^2.$$
Comme $\widehat{Y}$ ne dépend pas de H, alors $\widetilde{\sigma}^2$ ne dépend pas de H.
En écrivant $\widetilde{\sigma}^2 = \frac{1}{n-r}\,\big(\|\varepsilon\|^2 - \|P_{[X]}\varepsilon\|^2\big) \geq 0$ et $\frac{1}{n-r}\,\|\varepsilon\|^2 \underset{n\to+\infty}{\overset{\mathcal{P}}{\longrightarrow}} \sigma^2$ et $\mathbb{E}[\frac{1}{n-r}\,\|P_{[X]}\varepsilon\|^2] = \frac{r\,\sigma^2}{n-r} \underset{n\to\infty}{\longrightarrow} 0$, donc par l'inégalité de Markov $\frac{1}{n-r}\,\|P_{[X]}\varepsilon\|^2 \underset{n\to+\infty}{\overset{\mathcal{P}}{\longrightarrow}} 0$, on en déduit que $\widetilde{\sigma}^2 \underset{n\to+\infty}{\overset{\mathcal{P}}{\longrightarrow}} \sigma^2$.
Dans le cadre gaussien, $X\,\widetilde{\theta} = P_{[X]}Y = X\,\theta + P_{[X]}\varepsilon$ est indépendant de $\widetilde{\sigma}^2 = \frac{1}{n-r}\,\|P_{[X]^\perp}\varepsilon\|^2$ d'après le Théorème de Cochran car $[X]$ et $[X]^\perp$ étant des sev orthogonaux, $P_{[X]}\varepsilon$ et $P_{[X]^\perp}\varepsilon$ sont des vecteurs gaussiens de $\mathbf{R}^n$ indépendants.

11. Dans le cadre gaussien, avec $\widehat{Y} = P_{[X]}Y = X\,\theta + P_{[X]}\varepsilon$ qui est aussi un vecteur gaussien (application linéaire d'un vecteur gaussien) et on a $\widehat{Y} \overset{\mathcal{L}}{\sim} \mathcal{N}_n\big(X\,\theta\,,\, \sigma^2\,P_{[X]}\big)$. On en déduit que $\widehat{Y}_n \overset{\mathcal{L}}{\sim} \mathcal{N}\big((X\,\theta)_n\,,\, \sigma^2\,(P_{[X]})_{nn}\big)$. Ainsi $\widehat{Y}_n$ est un estimateur naturel de $(X\,\theta)_n$.
Comme σ^2 est inconnue, plutôt que d'utiliser $\frac{\widehat{Y}_n - (X\,\theta)_n}{\sigma\,\sqrt{(P_{[X]})_{nn}}}$ comme statistique de test, on va choisir comme statistique de test:
$$\widehat{T} = \frac{\widehat{Y}_n - (X\,\theta)_n}{\widetilde{\sigma}\,\sqrt{(P_{[X]})_{nn}}}.$$

Sous H_0, il est clair que $\frac{\widehat{Y}_n-(X\,\theta)_n}{\sigma\,\sqrt{(P_{[X]})_{nn}}} \overset{\mathcal{L}}{\sim} \mathcal{N}(0,1)$. Or d'après Cochran, $\widetilde{\sigma}^2 \overset{\mathcal{L}}{\sim} \frac{\sigma^2}{n-r}\,\chi^2(n-r)$. Et on a vu que $\widetilde{\sigma}^2$ est indépendant de $\widehat{Y}$, donc de $\widehat{Y}_n$. Aussi a-t-on

$$\widehat{T} \overset{\mathcal{L}}{\sim} t(n-r) \quad \text{sous l'hypothèse } H_0.$$

Si on fixe un risque de première espèce $0<\alpha<1$ à ce test, comme le test est unilatéral, la zone de rejet du test est:

$$\widehat{T} \in]-\infty\,,\,q_\alpha(t(n-r))[\qquad \text{où } q_\alpha(t(n-r)) \text{ est le quantile d'ordre } \alpha \text{ de la loi } t(n-r).$$

$\square$