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 $\widehat{\theta} \in \arg\min_{\theta \in \mathbf{R}^p} \big\{\big\|Y - X\,\theta\big\|\big\}$. 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 $X\left({}^tX'\,X'\right)^{-1}{}^tX$ 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)**.