Devoir en temps libre n$^{\text{o}}$ 5

Exercice 1 Modèle linéaire : cas des erreurs centrées et non corrélées.

Nous considérons une variable d'intérêt Y et k variables exogènes $(X_1, \ldots, X_k)$ que nous avons pu observer sur un échantillon de taille n. Pour l'observation i, nous avons donc le vecteur ligne $(Y_i, X_{1,i}, \ldots, X_{k,i})$. Considérons un modèle linéaire :

$$\boldsymbol{Y} = X\boldsymbol{\theta} + \boldsymbol{\epsilon}, \tag{1}$$

où X est la matrice de taille (n, k) des données exogènes où chaque ligne est constituée par un individu et chaque colonne est associée à l'une des variables X_k que nous avons observées, θ est un vecteur de k paramètres inconnus et $\boldsymbol{\epsilon}$ un vecteur aléatoire de taille n.

Nous faisons uniquement les deux hypothèses suivantes sur la loi du vecteur $\boldsymbol{\epsilon}$:

(H1) $\mathbb{E}[\boldsymbol{\epsilon}] = 0$.

(H2) $\mathbb{V}\text{ar}[\boldsymbol{\epsilon}] = \sigma^2 I_n$, où σ est inconnu et est de ce fait un paramètre additionnel du modèle.

(H3) $X'X \in \mathcal{GL}_k(\mathbb{R})$.

Nous notons $[X]$, le sous-espace vectoriel de $\mathbb{R}^n$ engendré par les colonnes de la matrice X. Pour F un sous-espace vectoriel de $\mathbb{R}^n$, nous notons P_F le projecteur orthogonal sur F.

1. Déterminer $\widehat{\boldsymbol{\theta}}$ l'estimateur des moindres carrés du vecteur de paramètres $\boldsymbol{\theta}$ du modèle (1). Préciser le biais et la variance de $\widehat{\boldsymbol{\theta}}$ ainsi qu'une expression du projecteur $P_{[X]}$ à l'aide de la matrice X.

2. Nous nous intéressons aux propriétés du carré moyen résiduel, $\frac{1}{n-k}\left\|\boldsymbol{Y} - X\widehat{\boldsymbol{\theta}}\right\|^2$, en tant qu'estimateur $\widehat{\sigma^2}$ du paramètre σ^2.

   a. Montrer que $(n-k)\widehat{\sigma^2} = \text{Tr}\left(\boldsymbol{\epsilon}' P_{[X]\perp}\boldsymbol{\epsilon}\right)$.

   b. En déduire que $(n-k)\mathbb{E}\left[\widehat{\sigma^2}\right] = \sigma^2\text{Tr}\left(P_{[X]\perp}\mathbb{E}\left[\boldsymbol{\epsilon}'\boldsymbol{\epsilon}\right]\right)$.

   c. Conclure.

3. $\widehat{\boldsymbol{\theta}}$ et $\widehat{\sigma^2}$ sont-ils corrélés ?

4. L'objectif est de montrer que l'estimateur des moindres carrés $\widehat{\boldsymbol{\theta}}$ est l'unique estimateur linéaire sans biais optimal parmi tous les estimateurs linéaires sans biais (Théorème de Gauss-Markov). L'optimalité signifie que pour tout autre estimateur linéaire sans biais $\widetilde{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$:

$$\mathbb{V}\text{ar}\left[\widetilde{\boldsymbol{\theta}}\right] - \mathbb{V}\text{ar}\left[\widehat{\boldsymbol{\theta}}\right] \in \mathcal{S}_n^+. \tag{2}$$

   a. Soit $\widetilde{\boldsymbol{\theta}}$ un estimateur linéaire sans biais de $\boldsymbol{\theta}$. Justifier qu'il existe une matrice M de taille (k, n) telle que $\widetilde{\boldsymbol{\theta}} = MY$. Montrer que $MX = I_k$.

b. Trouver une matrice T de taille (k, n) telle que $\widehat{\boldsymbol{\theta}} = TP_{[X]}Y$.

c. Montrer que $\widetilde{\boldsymbol{\theta}} = \widehat{\boldsymbol{\theta}} + MP_{[X]^\perp}Y$ et que les deux termes du membre de droite de l'équation précédente ne sont pas corrélés. En déduire le théorème de Gauss-Markov.

Exercice 2 Moindres carrés généralisés

Nous considérons une variable d'intérêt Y et k variables exogènes $(X_1, \ldots, X_k)$ que nous avons pu observer sur un échantillon de taille n. Pour l'observation i, nous avons donc le vecteur ligne $(Y_i, X_{1,i}, \ldots, X_{k,i})$. Considérons un modèle linéaire :

$$\boldsymbol{Y} = X\boldsymbol{\theta} + \boldsymbol{\epsilon}, \tag{3}$$

où X est la matrice de taille (n, k) des données exogènes où chaque ligne est constituée par un individu et chaque colonne est associée à l'une des variables X_k que nous avons observées, θ est un vecteur de k paramètres inconnus et $\boldsymbol{\epsilon}$ un vecteur aléatoire de taille n.

Nous faisons uniquement les deux hypothèses suivantes sur la loi du vecteur $\boldsymbol{\epsilon}$:

(H1) $\mathbb{E}[\boldsymbol{\epsilon}] = 0$.

(H2) $\mathbb{V}\mathrm{ar}[\boldsymbol{\epsilon}] = \sigma^2\Sigma$, où Σ est une matrice connue et σ est inconnu ce qui en fait un paramètre additionnel du modèle.

(H3) $X'X \in \mathcal{GL}_k(\mathbb{R})$.

1. Déterminer l'estimateur $\widehat{\boldsymbol{\theta}}$ des moindres carrés ordinaires du vexteur de paramètre $\boldsymbol{\theta}$ du modèle 3, puis calculer son biais ainsi que sa variance.

2. L'espace vectoriel $\mathbb{R}^n$ est désormais muni de la métrique définie par la distance suivante :

$$\forall(\boldsymbol{y_1}, \boldsymbol{y_2}) \in \mathbb{R}^n \times \mathbb{R}^n, \quad \|\boldsymbol{y_1} - \boldsymbol{y_2}\|_\Sigma^2 = (\boldsymbol{y_1} - \boldsymbol{y_2})'\Sigma^{-1}(\boldsymbol{y_1} - \boldsymbol{y_2}). \tag{4}$$

Vérifier que l'équation (5) suivante permet bien de définir un unique estimateur des moindres carrés généralisés $\widehat{\boldsymbol{\theta}}_{\boldsymbol{G}}$.

$$\widehat{\boldsymbol{\theta}}_{\boldsymbol{G}} = \mathrm{argmin}_{\boldsymbol{\theta} \in \mathbb{R}^k} \left(\|\boldsymbol{Y} - X\boldsymbol{\theta}\|_\Sigma^2 \right). \tag{5}$$

Donner l'expression de $\widehat{\boldsymbol{\theta}}_{\boldsymbol{G}}$, son espérance et sa variance.

3. Retrouver le résultat de la question **2.** en utilisant les changements de variable $\widetilde{\boldsymbol{Y}} = \Sigma^{-1/2}\boldsymbol{Y}$, $\widetilde{X} = \Sigma^{-1/2}X$ et $\widetilde{\boldsymbol{\epsilon}} = \Sigma^{-1/2}\boldsymbol{\epsilon}$, avec $\Sigma^{-1/2}$ est l'unique matrice définie postive dont le carré est égal à Σ^{-1}, puis en appliquant le théorème de Gauss-Markov, démontré à l'exercice 1, au modèle (6) ci-dessous.

$$\widetilde{\boldsymbol{Y}} = \widetilde{X}\boldsymbol{\theta} + \widetilde{\boldsymbol{\epsilon}}, \tag{6}$$

4. En déduire une généralisation du théorème de Gauss-Markov au cas de l'estimation par moindres carrés généralisés ainsi que l'optimalité de l'estimateur $\widehat{\boldsymbol{\theta}}_{\boldsymbol{G}}$ parmi la classe des estimateurs linéaires sans biais de $\boldsymbol{\theta}$.

5. Comparer la variance de $\widehat{\boldsymbol{\theta}}$ et $\widehat{\boldsymbol{\theta}}_{\boldsymbol{G}}$. Quelle est l'inégalité matricielle associée ?

6. Trouver un estimateur sans biais de σ^2.