Ecole Normale Supérieure
FIMFA – Statistique 2011-2012

Correction de l'exercice 3, Feuille 6.

Q.1. $\hat{\theta}_1 = (W'W)^{-1}W'Y = (W'W)^{-1}W'(W\theta_1 + Z\theta 2 + V)$. D'où,

$$\mathbb{E}[\hat{\theta}_1] = (W'W)^{-1}W'(W\theta_1 + Z\theta 2) = \theta_1 + (W'W)^{-1}W'Z\theta_2.$$

Cet estimateur est donc biaisé.

Q.2.(a) Dans le cas du modèle d'omission de variables, on a considéré le modèle $Y = W\theta_1 + Z\theta_2 + V$ avec la contrainte $\theta_2 = 0$. Cette contrainte peut se réécrire sous la forme

$$R\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix} = r$$

où $R = (\mathbf{0}_{p_2\times p_1}, \mathbf{I}_{p_2})$ et $r = \mathbf{0}_{p_2}$.

Q.2.(b) • Tout d'abord, $\hat{\beta} = (X'X)^{-1}X'Y$ (X est supposée de rang plein).

• Ensuite, l'estimateur du maximum de vraisemblance sous la contrainte est le vecteur β (sous réserve d'existence et d'unicité) qui maximise la vraisemblance

$$\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i - (X\beta)_i)^2\right)$$

parmi les vecteurs β tels que $R\beta = r$ (bien qu'ici les deux paramètres β et σ^2 sont inconnus, ce maximiseur ne dépend pas de σ^2). Evidemment, ce problème est équivalent à trouver le vecteur β, parmi ceux vérifiant la contrainte, qui minimise $\sum_{i=1}^n (y_i - (X\beta)_i)^2$. Pour cela, on introduit le lagrangien

$$L(\beta_1, ..., \beta_p, \lambda_1, ..., \lambda_d) = \sum_{i=1}^n (y_i - (X\beta)_i)^2 - \sum_{j=1}^d \lambda_j((R\beta_j) - r_j),$$

où p est la dimension de β et d celle de r.

On suppose dans ce qui suit que R est de rang $d \leq p$. En particulier R est injective et $R(X'X)^{-1}R'$ est inversible.

Le lagrangien est une fonction convexe et les contraintes sont linéaires. On sait alors que s'il existe un unique couple (β^*, λ^*) tel que

$$\frac{\partial L}{\partial \beta_i}(\beta^*, \lambda^*) = 0 \text{ et } \frac{\partial L}{\partial \lambda_j}(\beta^*, \lambda^*) = 0, \quad \forall 1 \leq i \leq p, 1 \leq j \leq d,$$

alors β^* sera le minimum recherché. Dans notre cas, ceci revient à chercher (β^*, λ^*) tels que

$$\begin{cases} 2X'Y - 2X'X\beta^* + R'\lambda^* = \mathbf{0}_p \\ R\beta^* - r = \mathbf{0}_d \end{cases} \tag{1}$$

En multipliant la première ligne par $R(X'X)^{-1}$ et en y remplaçant ensuite $R\beta^*$ par r, on obtient,

$$\lambda^* = -2\left(R(X'X)^{-1}R'\right)^{-1}\left(R(X'X)^{-1}X'Y) - r\right).$$

Puis en remplaçant λ^* par cette valeur dans la première équation de (1) et en la multipliant par $(X'X)^{-1}$, on obtient,

$$\beta^* = \hat{\beta} + (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}(r - R\hat{\beta}).$$

Il est facile de voir que ce couple (β^*, λ^*) vérifie bien (1). On a donc existence et unicité. On en déduit que l'estimateur sous contrainte cherché est $\hat{\beta}^c = \beta^*$.

Q.2.(c) $\mathbb{E}[\hat{\beta}^c] = \beta + (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}(r - R\beta)$.

Q.2.(d) Le résultat cherché se démontre facilement en écrivant $\hat{\beta}^c = A\hat{\beta} + C$, avec A et C des matrices déterministes adéquates, puis en utilisant $\mathrm{Var}(A\hat{\beta} + C) = A\mathrm{Var}(\hat{\beta})A'$ et $\mathrm{Var}(\hat{\beta}) = \sigma_E^2(X'X)^{-1}$. On en déduit que

$$\mathrm{Var}(\hat{\beta}^c) = \mathrm{Var}(\hat{\beta}) - S,$$

où S est une matrice symétrique semi-définie positive. Par conséquent, pour tout vecteur z de dimension p, $z'\mathrm{Var}(\hat{\beta}^c)z \leq z'\mathrm{Var}(\hat{\beta})z$.

Q.2.(e) On en déduit que l'estimateur construit à partir du faux modèle a une variance plus petite que celui construit à partir du vrai modèle. Cependant il est biaisé.