Économétrie 1, 01/2019, Correction

Exo 1.

1. Les résidus estimés sont $\hat{\varepsilon} = Y - \hat{Y} = Y - (\hat{\mu} + \hat{\beta} X)$.

On obtient donc la moyenne des résidus estimés

$$\bar{\hat{\varepsilon}} = \bar{Y} - (\hat{\mu} + \hat{\beta}\bar{X}) = \bar{Y} - \hat{\beta}\bar{X} - \hat{\mu} = \hat{\mu} - \hat{\mu} = 0.$$

Donc la somme des 5 résidus vaut zéro. Ce résultat n'est pas surprenant, on l'a toujours pour le modèle $Y = \mu + \beta X + \varepsilon$.

2. Cette fois ci le modèle est $Y = \beta X + \varepsilon$. La formule d'estimation $\hat{\theta} = (X'X)^{-1} X'Y$ est toujours valable. On obtient $\hat{\beta} = \hat{\theta} = \frac{\sum X_i Y_i}{\sum X_i^2}$.

Les résidus $\hat{\varepsilon} = Y - \hat{Y} = Y - \hat{\beta} X = Y - \frac{\sum X_i Y_i}{\sum X_i^2} X$. La moyenne des résidus $\bar{\hat{\varepsilon}} = \bar{Y} - \frac{\sum X_i Y_i}{\sum X_i^2} \cdot \bar{X} \neq 0$. Donc la somme des résidus ne vaut pas zéro.

Exo 2.

1.

<none>	reg1 = lm(Y ~ X1 + X2 + X3)	modèle 1
X1	reg2 = lm(Y ~ X2 + X3)	modèle 2
X2	reg3 = lm(Y ~ X1 + X3)	modèle 3
X3	reg4 = lm(Y ~ X1 + X2)	modèle 4

RSS = sum(residuals(reg i)^2), i = 1, 2, 3, 4

$= \sum (Y_i - \hat{Y}_i)^2$

: Somme des carrées des résidus du modèle i

8.6174 : $RSS(reg3) - RSS(reg1) = 35.199 - 26.582$

26.582 : $RSS(reg1)$

35.199 : $RSS(reg3)$

17.159 : $-2\log L + k *$ nombre des paramètres $= n\ln\left(\frac{RSS}{n}\right) + 2 *$ nombre des paramètres

2. On garde le modèle 2 $Y \sim X_2 + X_3$ pour la prochaine étape.

Exo3.

1. $Y_i \sim \mathcal{N}(\beta_1 X_i + \beta_2 X_i^2, \sigma^2)$

Non, parceque les espérances de Y_i sont différentes.

2. <u>MMC</u> : $X = \begin{pmatrix} X_1 & X_1^2 \\ \vdots & \vdots \\ X_n & X_n^2 \end{pmatrix}$, $X'X = \begin{pmatrix} \sum X_i^2 & \sum X_i^3 \\ \sum X_i^3 & \sum X_i^4 \end{pmatrix}$

$$\det(X'X) = \sum X_i^2 \sum X_i^4 - \left(\sum X_i^3\right)^2$$

$$(X'X)^{-1} = \frac{1}{\det}\begin{pmatrix} \sum X_i^4 & -\sum X_i^3 \\ -\sum X_i^3 & \sum X_i^2 \end{pmatrix}$$

$$X'Y = \begin{pmatrix} X_1 & \cdots & X_n \\ X_1^2 & \cdots & X_n^2 \end{pmatrix}\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} \sum X_i Y_i \\ \sum X_i^2 Y_i \end{pmatrix}$$

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \hat{\theta} = (X'X)^{-1}X'Y = \frac{1}{\det}\begin{pmatrix} \sum X_i^4 \sum X_i Y_i - \sum X_i^3 \sum X_i^2 Y_i \\ \sum X_i^2 \sum X_i^2 Y_i - \sum X_i^3 \sum X_i Y_i \end{pmatrix}$$

$$SCR = \|Y - (\hat{\beta}_1 X + \hat{\beta}_2 X^2)\|^2 = \|Y - P_{[X]}Y\|^2 \sim \sigma^2 \cdot \chi^2(n-2)$$

$$\hat{\sigma}^2 = \frac{SCR}{n-2}$$

$\underline{MV}$: La densité de Y_i est

$$f(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\beta_1 X_i - \beta_2 X_i^2)^2}{2\sigma^2}}.$$

La vraisemblance est

$$L(\beta_1, \beta_2, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum (Y_i - \beta_1 X_i - \beta_2 X_i^2)^2}.$$

On obtient

$$H(\beta_1, \beta_2, \sigma^2) = -2\log L = n\log(2\pi) + n\log(\sigma^2) + \frac{1}{\sigma^2}\sum (Y_i - \beta_1 X_i - \beta_2 X_i^2)^2,$$

$$\frac{\partial H}{\partial \beta_1} = \frac{1}{\sigma^2}\sum 2(Y_i - \beta_1 X_i - \beta_2 X_i^2)(-X_i),$$

$$\frac{\partial H}{\partial \beta_2} = \frac{1}{\sigma^2}\sum 2(Y_i - \beta_1 X_i - \beta_2 X_i^2)(-X_i^2),$$

$$\frac{\partial H}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{1}{\sigma^4}\sum (Y_i - \beta_1 X_i - \beta_2 X_i^2)^2,$$

$$\frac{\partial H^2}{\partial \beta_1^2} = \frac{1}{\sigma^2}\sum 2X_i^2, \quad \frac{\partial H^2}{\partial \beta_2^2} = \frac{1}{\sigma^2}\sum 2X_i^4, \quad \frac{\partial H^2}{\partial (\sigma^2)^2} = -\frac{n}{\sigma^4} + \frac{2}{\sigma^6}\sum (Y_i - \beta_1 X_i - \beta_2 X_i^2)^2,$$

$$\frac{\partial H^2}{\partial \beta_1 \partial \beta_2} = \frac{\partial H^2}{\partial \beta_2 \partial \beta_1} = \frac{1}{\sigma^2}\sum 2X_i^3,$$

$$\frac{\partial H^2}{\partial \beta_1 \partial \sigma^2} = \frac{\partial H^2}{\partial \sigma^2 \partial \beta_1} = \frac{2}{\sigma^4}\sum (Y_i - \beta_1 X_i - \beta_2 X_i^2) \cdot X_i,$$

$$\frac{\partial H^2}{\partial \beta_2 \partial \sigma^2} = \frac{\partial H^2}{\partial \sigma^2 \partial \beta_2} = \frac{2}{\sigma^4}\sum (Y_i - \beta_1 X_i - \beta_2 X_i^2) \cdot X_i^2.$$

La matrice Hesienne est
$$M = \frac{2}{\sigma^2}\begin{pmatrix} \sum X_i^2 & \sum X_i^3 & \frac{1}{\sigma^2}\sum X_i(Y_i - \beta_1 X_i - \beta_2 X_i^2) \\ \sum X_i^3 & \sum X_i^4 & \frac{1}{\sigma^2}\sum X_i^2(Y_i - \beta_1 X_i - \beta_2 X_i^2) \\ \frac{1}{\sigma^2}\sum X_i(Y_i - \beta_1 X_i - \beta_2 X_i^2) & \sim & -\frac{n}{2\sigma^2} + \frac{1}{\sigma^4}\sum (Y_i - \beta_1 X_i - \beta_2 X_i^2)^2 \end{pmatrix}$$

qui est définie positive.

Donc les estimateurs du maximum de vraisemblance sont

$$\hat{\beta}_1 = \frac{\sum X_i^4 \sum X_i Y_i - \sum X_i^3 \sum X_i^2 Y_i}{\sum X_i^2 \sum X_i^4 - \left(\sum X_i^3\right)^2},$$

$$\hat{\beta}_2 = \frac{\sum X_i^2 \sum X_i^2 Y_i - \sum X_i^3 \sum X_i Y_i}{\sum X_i^2 \sum X_i^4 - \left(\sum X_i^3\right)^2},$$

$$\hat{\sigma}^2 = \frac{\sum (Y_i - \hat{\beta}_1 X_i - \hat{\beta}_2 X_i^2)^2}{n}.$$

3. <u>MMC</u> : Oui

<u>MV</u> : Oui pour $\hat{\beta}_1$ et $\hat{\beta}_2$, non pour $\hat{\sigma}^2$

4. <u>MMC</u> : $$Var\begin{pmatrix}\hat{\beta}_1 \\ \hat{\beta}_2\end{pmatrix} = \sigma^2 (X'X)^{-1} = \frac{\sigma^2}{det}\begin{pmatrix}\sum X_i^4 & -\sum X_i^3 \\ -\sum X_i^3 & \sum X_i^2\end{pmatrix}$$

$$Var(\hat{\beta}_1) = \frac{\sigma^2 \sum X_i^4}{det}, \quad Var(\hat{\beta}_2) = \frac{\sigma^2 \sum X_i^2}{det}.$$

$$Var(\hat{\sigma}^2) = Var\left(\frac{SCR}{n-2}\right) = \frac{1}{(n-2)^2}\cdot \sigma^4 \cdot 2(n-2) = \frac{2}{n-2}\sigma^4$$

<u>MV</u> : $$Var(\hat{\beta}_1) = \frac{Var \sum \left(X_i \sum X_i^4 - X_i^2 \sum X_i^3\right) Y_i}{det^2}$$

$$= \frac{\sum \left(X_i \sum X_i^4 - X_i^2 \sum X_i^3\right)^2 \cdot \sigma^2}{det^2}$$

$$= \frac{\left(\sum X_i^2 \left(\sum X_i^4\right)^2 - \left(\sum X_i^3\right)^2 \sum X_i^4\right)\sigma^2}{det^2}$$

$$= \frac{\sum X_i^4}{det}\cdot \sigma^2$$

$$Var(\hat{\beta}_2) = \frac{Var \sum (X_i^2 \sum X_i^2 - X_i \sum X_i^3) Y_i}{det^2}$$

$$= \frac{\sum (X_i^2 \sum X_i^2 - X_i \sum X_i^3)^2 \cdot \sigma^2}{det^2}$$

$$= \frac{\sum X_i^4 (\sum X_i^2)^2 - \sum X_i^2 (\sum X_i^3)^2}{det^2} \cdot \sigma^2$$

$$= \frac{\sum X_i^2}{det} \cdot \sigma^2$$

$$Var(\hat{\sigma^2}) = Var\left(\frac{SCR}{n}\right) = \frac{1}{n^2} \cdot \sigma^4 \cdot 2(n-2) = \frac{2(n-2)}{n^2} \sigma^4$$

5. D'après l'exercice 1 du TD8 on a pour le modèle linéaire non gaussien

$\hat{\sigma^2} = \frac{SCR}{n-K}$, $E(\hat{\sigma^2}) = \sigma^2$ où $\hat{\sigma}^2$ est l'estimateur de σ^2 par MMC.

La normalité des ε_i ne sont pas utilisée pour calculer $E(\hat{\beta}_1)$ et $E(\hat{\beta}_2)$.

Donc les $\hat{\beta}_1$ et $\hat{\beta}_2$ sont toujours sans biais, c.a.d.

$E(\hat{\beta}_1) = \beta_1$, $E(\hat{\beta}_2) = \beta_2$.

<u>MMC</u> : $E(\hat{\sigma^2}) = E\left(\frac{SCR}{n-2}\right) = \frac{E(SCR)}{n-2} = \frac{(n-2)\sigma^2}{n-2} = \sigma^2$.

<u>MV</u> : $E(\hat{\sigma^2}) = E\left(\frac{SCR}{n}\right) = \frac{(n-2)\sigma^2}{n}$.