Université PARIS-XII — Faculté de Sciences économiques et de gestion
Master *Sciences économiques*
Mention *Économie quantitative et expertise internationale*
Économétrie III — *F. LEGENDRE* — année universitaire 2005-06

Corrigé sommaire de l'interrogation écrite n°1

1) Le modèle se met sous la forme matricielle suivante quand les données sont rangées sur le critère individu/période :

$$(1) \quad \underline{y} = \mathrm{X}\underline{a} + \underline{u}$$

avec

$$\underline{y} = \begin{pmatrix} 2 \\ 4 \\ 3 \\ 3 \end{pmatrix} \quad \mathrm{X} = \begin{pmatrix} 8 & 1 & 0 \\ 2 & 0 & 1 \\ 4 & 1 & 0 \\ 10 & 0 & 1 \end{pmatrix} \quad \underline{a} = \begin{pmatrix} \alpha \\ \beta_1 \\ \beta_2 \end{pmatrix} \quad \underline{u} = \begin{pmatrix} u_{11} \\ u_{12} \\ u_{21} \\ u_{22} \end{pmatrix}$$

2) La première colonne de la matrice X est la première variable explicative ; c'est donc le vecteur $\underline{x}$. Les deux autres colonnes de la matrice peuvent s'écrire comme suit

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \\ 1\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \underline{e}_2 \otimes \mathrm{I}_2$$

J'identifie ainsi le vecteur $\underline{c}$ au vecteur $\underline{e}_2$ et la matrice D à la matrice I_2.

3) Soit P la matrice de la projection orthogonale sur le sous-espace vectoriel engendré par les colonnes de la matrice Z. On sait que cette matrice s'écrit :

$$\mathrm{P} = \mathrm{Z}(\mathrm{Z}'\mathrm{Z})^{-1}\mathrm{Z}'$$

Au cas particulier où la matrice Z est $\underline{e}_\mathrm{N} \otimes \mathrm{I}_\mathrm{T}$, la matrice de la projection orthogonale est alors

$$\mathrm{P} = (\underline{e}_\mathrm{N} \otimes \mathrm{I}_\mathrm{T})[(\underline{e}_\mathrm{N} \otimes \mathrm{I}_\mathrm{T})'(\underline{e}_\mathrm{N} \otimes \mathrm{I}_\mathrm{T})]^{-1}(\underline{e}_\mathrm{N} \otimes \mathrm{I}_\mathrm{T})'$$

soit

$$(\underline{e}_\mathrm{N} \otimes \mathrm{I}_\mathrm{T})[\frac{1}{\mathrm{N}}\mathrm{I}_\mathrm{T}](\underline{e}_\mathrm{N} \otimes \mathrm{I}_\mathrm{T})' = \frac{1}{\mathrm{N}}(\underline{e}_\mathrm{N}\underline{e}'_\mathrm{N} \otimes \mathrm{I}_\mathrm{T})$$

Dans la fiche de travail n°5, un calcul comparable est détaillé.

En prenant, comme dans cet exercice, $\mathrm{N} = 2$ et $\mathrm{T} = 2$, je trouve la forme suivante.

$$\frac{1}{2}\left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}(1,1) \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] = \frac{1}{2}\left[\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] = \begin{pmatrix} 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \end{pmatrix}$$

Cette matrice permet d'obtenir les moyennes temporelles : z_{it} devient $\overline{z}_t$.

4) La matrice Q, égale à $\mathrm{I}_\mathrm{NT} - \mathrm{P}$, est symétrique si $\mathrm{Q}' = \mathrm{Q}$. On a

$$\mathrm{Q}' = (\mathrm{I}_\mathrm{NT} - \mathrm{P})' = \mathrm{I}'_\mathrm{NT} - \mathrm{P}' = \mathrm{I}_\mathrm{NT} - \mathrm{P} = \mathrm{Q}$$

parce que la matrice P est symétrique (puisqu'elle est associée à une projection orthogonale). La matrice Q est idempotente si $\mathrm{Q}^2 = \mathrm{Q}$. On a

$$\mathrm{Q}^2 = \mathrm{QQ} = (\mathrm{I}_\mathrm{NT} - \mathrm{P})(\mathrm{I}_\mathrm{NT} - \mathrm{P}) = \mathrm{I}_\mathrm{NT} - 2\mathrm{P} + \mathrm{P}^2 = \mathrm{I}_\mathrm{NT} - \mathrm{P} = \mathrm{Q}$$

parce que la matrice P est idempotente (puisqu'elle est associée à une projection orthogonale).

La matrice Q est la matrice de la projection orthogonale « ennemie » de P : $\underline{z} = \mathrm{P}\underline{z} + \mathrm{Q}\underline{z}$ et $\mathrm{P}\underline{z} \perp \mathrm{Q}\underline{z}$. Cette matrice permet d'obtenir les écarts à la moyenne temporelle : z_{it} devient $z_{it} - \overline{z}_t$.

5) La propriété que l'on cherche à montrer correspond en fait au théorème de FRISCH–WAUGH. Vous pouvez consulter la fiche de travail n°6. On part du modèle (1)

$$(1) \quad \underline{y} = \alpha\underline{x} + \beta_1\underline{t}_1 + \beta_2\underline{t}_2 + \underline{u}$$

où $\underline{t}_1$ est la première colonne de la matrice $\underline{e}_2 \otimes \mathrm{I}_2$ et $\underline{t}_2$ la seconde colonne de cette même matrice. $\underline{t}_1$ est donc la variable explicative associée au coefficient β_1 et de même pour $\underline{t}_2$.

Soit maintenant $p(\cdot)$ la projection orthogonale sur le sous espace vectoriel engendré par $\underline{t}_1$ et $\underline{t}_2$. En partant du modèle (1) ajusté et en lui appliquant cette projection, on obtient :

$$p(\underline{y}) = \widehat{\alpha}p(\underline{x}) + \widehat{\beta}_1 p(\underline{t}_1) + \widehat{\beta}_2 p(\underline{t}_2) + p(\underline{\widehat{u}})$$

Comme $p(\underline{t}_1) = \underline{t}_1$, $p(\underline{t}_2) = \underline{t}_2$ et $p(\underline{\widehat{u}}) = \underline{0}$, on obtient par différence

$$[\underline{y} - p(\underline{y})] = \widehat{\alpha}[\underline{x} - p(\underline{x})] + \underline{\widehat{u}}$$

En remarquant que $[\underline{x}-p(\underline{x})] \perp \underline{\widehat{u}}$, on montre que l'ajustement des MCO de la variable à expliquer $\underline{y}-p(\underline{y})$ sur la variable explicative $\underline{x}-p(\underline{x})$ donne le coefficient estimé $\widehat{\alpha}$. Enfin, il faut montrer que $\mathrm{Q}\underline{z} = \underline{z} - p(\underline{z})$ mais cela est évident par définition de la matrice Q : $\mathrm{Q}\underline{z} = (\mathrm{I}_{\mathrm{NT}} - \mathrm{P})\underline{z} = \underline{z} - \mathrm{P}\underline{z} = \underline{z} - p(\underline{z})$.

6) L'application de la matrice Q au vecteur $\underline{z}$ permet d'obtenir les écarts à la moyenne temporelle : z_{it} devient $z_{it} - \overline{z}_t$. On a $\overline{y}_1 = 2{,}5$; $\overline{y}_2 = 3{,}5$; $\overline{x}_1 = 6$ et $\overline{x}_2 = 6$. On en tire :

$$\mathrm{Q}\underline{y} = \begin{pmatrix} 2 \\ 4 \\ 3 \\ 3 \end{pmatrix} - \begin{pmatrix} 2{,}5 \\ 3{,}5 \\ 2{,}5 \\ 3{,}5 \end{pmatrix} = \begin{pmatrix} -0{,}5 \\ 0{,}5 \\ 0{,}5 \\ -0{,}5 \end{pmatrix} \qquad \mathrm{Q}\underline{x} = \begin{pmatrix} 8 \\ 2 \\ 4 \\ 10 \end{pmatrix} - \begin{pmatrix} 6 \\ 6 \\ 6 \\ 6 \end{pmatrix} = \begin{pmatrix} 2 \\ -4 \\ -2 \\ 4 \end{pmatrix}$$

On en déduit que

$$\widehat{\alpha} = \frac{-2\times 0{,}5 - 4\times 0{,}5 - 2\times 0{,}5 - 4\times 0{,}5}{2^2 + (-4)^2 + (-2)^2 + 4^2} = \frac{-6}{40} = -0{,}15$$

7) L'objet de ce fragment de programme SAS est de créer une table, nommée `table`, qui comporte 4 observations et qui contient les six variables suivantes :

`i`	variable de classification qui repère l'individu
`t`	variable de classification qui repère l'année
`y`	variable à expliquer, continue à deux indices
`x`	variable explicative, continue à deux indices
`t1`	variable indicatrice qui repère la première année
`t2`	variable indicatrice qui repère la seconde année

Les variables indicatrices qui repèrent les années sont engendrées à partir d'une expression logique qui est égale soit à « oui » soit à « non ». Cela permet d'obtenir exactement ce que l'on veut, c'est-à-dire des variables dont les valeurs sont soit 1 soit 0.

8) L'objet de ce fragment de programme SAS est d'estimer le modèle (1). Les variables explicatives `t1` et `t2` correspondent aux coefficients à estimer β_1 et β_2. Il faut coder `NOINT` pour ne pas permettre à SAS d'inclure la constante comme variable explicative supplémentaire.

9) L'objet de ce fragment de programme SAS est d'estimer le modèle suivant.

$$[\underline{y}-p(\underline{y})] = \gamma[\underline{x}-p(\underline{x})] + \underline{v}$$

La variable $\underline{y}-p(\underline{y})$ est le résidu de l'ajustement par les MCO de la variable $\underline{y}$ sur les variables explicatives $\underline{t}_1$ et $\underline{t}_2$. Pour l'obtenir, il faut donc coder, au sein d'une `PROC reg`, une instruction `MODEL` puis une instruction `OUTPUT` afin de récupérer le résidu de l'ajustement. Ce résidu est stocké dans la table `yTable` ; son nom est `qy`. De même, la variable $\underline{x}-p(\underline{x})$ est le résidu de l'ajustement de la variable $\underline{x}$ sur ces mêmes variabes explicatives. Ce résidu est stocké dans la table `xTable`, son nom est `qx`.

Ensuite, les deux tables `yTable` et `xTable` sont fusionnées dans une étape `DATA`. Il n'est pas nécessaire de coder une instruction `BY` : la fusion est faite observation par observation puisque les deux tables sont exactement relatives aux mêmes observations.

Enfin, la dernière `PROC reg` a pour objet d'estimer le modèle.