TD Econométrie 1, éléments de corrigé

Exercice 1. Rappels sur les estimateurs des Moindres Carrés. Propriétés à distance finie et propriétés asymptotiques

1.1. Estimateurs des MCO/MCG et propriétés à distance finie

Nous considérons ici l'inférence sur le modèle linéaire ($\mathbf{x}_i$ contient un terme constant) :

(1) $y_i = \mathbf{a}_0'\mathbf{x}_i + u_i$ et $E\left[u_i\right] = 0$

avec :

(2) $E\left[u_i / \mathbf{x}_i\right] = E\left[u_i\right] = 0 \quad \Rightarrow \quad E\left[\mathbf{x}_i u_i\right] = \mathbf{0}_{P\times 1}$

et :

(3) $E\left[\mathbf{u}\mathbf{u}'\right] = \sigma^2 \mathbf{\Omega}$ avec $\mathbf{\Omega}$ inversible

Dans la suite, les notations suivantes seront utilisées :

$$\mathbf{y} \equiv \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}_{N\times 1}, \quad \mathbf{X} \equiv \begin{bmatrix} \mathbf{x}_1' \\ \vdots \\ \mathbf{x}_N' \end{bmatrix}_{N\times P} \quad \text{et} \quad \mathbf{u} \equiv \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix}_{N\times 1}.$$

1.1.1. Commenter rapidement les hypothèses posées ici

- (1) ne donne que la forme du modèle : il est linéaire dans les paramètres et les variables.
- (2) est l'hypothèse fondamentale du modèle linéaire ordinaire. Elle indique que les variables explicatives $\mathbf{x}_i$ sont exogènes par rapport à terme d'erreur u_i. Elle est à la base des propriétés de sans biais et de convergence des estimateurs des Moindres Carrés.
- (3) indique que les termes d'erreur peuvent être auto-corrélés et/ou hétéroscédastiques (dans une certaine mesure). Le modèle considéré est donc un modèle linéaire généralisé.

$\mathbf{\Omega}$ inversible indique simplement qu'il n'y a pas plusieurs fois la même observation ou qu'une observation n'est pas une combinaison linéaire d'autres observations dans les données utilisées.

1.1.2. Rappeler la forme des estimateurs des MCO et des MCG de a_0 et donner leurs principales propriétés à distance finie (avec N fixe)

- $\hat{\mathbf{a}}_N^{MCO} \equiv \underset{\mathbf{a}}{ArgMin} \sum_{i=1}^{N} (y_i - \mathbf{x}_i'\mathbf{a})^2 = \underset{\mathbf{a}}{ArgMin} \left(\mathbf{y} - \mathbf{X}\mathbf{a}\right)' \left(\mathbf{y} - \mathbf{X}\mathbf{a}\right)$

- $\hat{\mathbf{a}}_N^{MCO} = \left[\mathbf{X}'\mathbf{X}\right]^{-1} \mathbf{X}'\mathbf{y} = \left[\sum_{i=1}^{N} \mathbf{x}_i\mathbf{x}_i'\right]^{-1} \sum_{i=1}^{N} \mathbf{x}_i y_i = \left[\frac{1}{N}\sum_{i=1}^{N} \mathbf{x}_i\mathbf{x}_i'\right]^{-1} \frac{1}{N}\sum_{i=1}^{N} \mathbf{x}_i y_i$

- $\hat{\mathbf{a}}_N^{MCG} \equiv \underset{\mathbf{a}}{ArgMin} \left(\mathbf{y} - \mathbf{X}\mathbf{a}\right)' \mathbf{\Omega}^{-1} \left(\mathbf{y} - \mathbf{X}\mathbf{a}\right)$

- $\hat{\mathbf{a}}_N^{MCG} = \left[\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X}\right]^{-1} \mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{y}$

- Remarque : $\hat{\mathbf{a}}_N^{MCO}$ est un cas particulier de $\hat{\mathbf{a}}_N^{MCG}$, on a $\hat{\mathbf{a}}_N^{MCG} = \hat{\mathbf{a}}_N^{MCO}$ si $\mathbf{\Omega} = \mathbf{I}_N$, i.e. si les u_i ne sont pas auto-corrélés et sont homoscédastiques.

- Remarque : $\hat{\mathbf{a}}_N^{MCG}$ peut être calculé comme un estimateur des MCO de $\mathbf{a}_0$ dans le modèle sphéricisé. Sous forme empilée, le modèle initial (généralisé) est donné par :

$$\mathbf{y} = \mathbf{X}'\mathbf{a} + \mathbf{u}.$$

La matrice $\mathbf{\Omega}$ a la structure d'une matrice de variance-covariance, elle est donc carrée, symétrique, semie-définie positive. Avec $\mathbf{\Omega}$ inversible, on sait qu'elle est définie positive et donc sait que $\mathbf{\Omega}^{-1}$, $\mathbf{\Omega}^{1/2}$ et $\mathbf{\Omega}^{-1/2}$ existent (et sont symétriques). Le modèle sphéricisé pour $\mathbf{\Omega}$ est donné par :

$$\mathbf{\Omega}^{-1/2}\mathbf{y} = \mathbf{\Omega}^{-1/2}\mathbf{X}\mathbf{a} + \mathbf{\Omega}^{-1/2}\mathbf{u}.$$

On a alors

$$\begin{aligned} V\left[\mathbf{\Omega}^{-1/2}\mathbf{u}\right] &= E\left[\mathbf{\Omega}^{-1/2}\mathbf{u}(\mathbf{\Omega}^{-1/2}\mathbf{u})'\right] \\ &= E\left[\mathbf{\Omega}^{-1/2}\mathbf{u}\mathbf{u}'\mathbf{\Omega}^{-1/2}\right] \\ &= \mathbf{\Omega}^{-1/2} E\left[\mathbf{u}\mathbf{u}'\right]\mathbf{\Omega}^{-1/2} \\ &= \sigma^2\mathbf{\Omega}^{-1/2}\mathbf{\Omega}\mathbf{\Omega}^{-1/2} = \sigma^2\mathbf{I}_N \end{aligned}$$

Aussi le modèle sphéricisé (avec la « bonne » matrice $\mathbf{\Omega}^{-1/2}$) est un modèle linéaire ordinaire, i.e. avec des termes d'erreur non auto-corrélés et homoscédastiques : c'est là tout l'intérêt de la sphéricisation !

Propriétés à distance finie :

- Lorsque les $\mathbf{x}_i$ sont exogènes par rapport aux u_i, on peut considérer que les $\mathbf{x}_i$ sont fixes (comme $\mathbf{X}$), ce qu'on fait pour étudier les propriétés à distance finie (N fixe) des estimateurs des MC.

- Quelque soit $\mathbf{\Omega}$, $\hat{\mathbf{a}}_N^{MCO}$ et $\hat{\mathbf{a}}_N^{MCG}$ sont sans biais $E\left[\hat{\mathbf{a}}_N^{MCO}\right] = E\left[\hat{\mathbf{a}}_N^{MCG}\right] = \mathbf{a}_0$.

- $\hat{\mathbf{a}}_N^{MCG}$ est le meilleur estimateur linéaire sans biais de $\mathbf{a}_0$ dans le modèle linéaire généralisé

- $\hat{\mathbf{a}}_N^{MCO}$ est le meilleur estimateur linéaire sans biais de $\mathbf{a}_0$ si le modèle est un modèle linéaire ordinaire, i.e. si $\mathbf{\Omega} = \mathbf{I}_N$.

- On ne connaît pas la loi de $\hat{\mathbf{a}}_N^{MCO}$ et $\hat{\mathbf{a}}_N^{MCG}$ si la loi de $\mathbf{u}$ n'est pas connue (mais la propriété de sans biais ne suppose rien quant à la loi de $\mathbf{u}$)

- Si $\mathbf{u}$ suit une loi normale $\mathbf{u} \sim \mathbf{N}\left(\mathbf{0}_{N\times 1}; \sigma^2\mathbf{\Omega}\right)$, alors $\hat{\mathbf{a}}_N^{MCG}$ (et donc $\hat{\mathbf{a}}_N^{MCO}$ qui est un cas particulier de $\hat{\mathbf{a}}_N^{MCG}$ qu'on ne doit utiliser que lorsque $\mathbf{\Omega} = \mathbf{I}_N$) est normal :

$$\mathbf{a}_N^{MCG} \sim \mathbf{N}\left(\mathbf{a}_0; \sigma^2\left[\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X}\right]^{-1}\right) = \mathbf{N}\left(\mathbf{a}_0; \sigma^2 \frac{1}{N}\left[\frac{\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X}}{N}\right]^{-1}\right).$$

Remarque : les estimateurs des MCO et des MCG dans les modèles linéaires sont pratiquement les seuls (hormis les estimateurs de moyennes simples) estimateurs utilisés en économétrie pour lesquels les propriétés sont connues à distance finie.

1.2. Propriétés asymptotiques de l'estimateur des MCO

1.2.1. Propriétés asymptotiques de l'estimateur des MCO à partir de l'exogénéité des $\mathbf{x}_i$.

Nous considérons dans un premier temps le modèle linéaire à variables explicatives exogènes.

(1) $\quad y_i = \mathbf{a}_0'\mathbf{x}_i + u_i$ et $E\left[u_i\right] = 0$

avec :

(2) $\quad E\left[u_i/\mathbf{x}_i\right] = E\left[u_i\right] = 0 \quad \Rightarrow \quad E\left[\mathbf{x}_i u_i\right] = \mathbf{0}_{P\times 1}$

et les (y_i ,$\mathbf{x}_i$) sont i.i.d..

1.2.1a. Montrer qu'il est possible de construire l'estimateur des MCO à partir de la contre-partie empirique de l'hypothèse d'exogénéité des $\mathbf{x}_i$: $E\left[\mathbf{x}_i u_i\right] = \mathbf{0}_{P\times 1}$.

- On sait que : $E\left[\mathbf{x}_i u_i\right] = \mathbf{0}_{P\times 1}$, donc que :

$$E\left[\mathbf{x}_i(y_i - \mathbf{x}_i'\mathbf{a}_0)\right] = \mathbf{0}_{P\times 1}.$$

Construisons maintenant un estimateur $\hat{\mathbf{a}}_N$ de $\mathbf{a}_0$ comme la solution de la contre-partie empirique de ce système de P équations (à P inconnues) :

$$\hat{\mathbf{a}}_N \text{ solution en } \mathbf{a} \text{ de } \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i(y_i - \mathbf{x}_i'\mathbf{a}) = \mathbf{0}_{P\times 1} :$$

D'où :

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i(y_i - \mathbf{x}_i'\hat{\mathbf{a}}_N) = \mathbf{0}_{P\times 1}$$

et :

$$\frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i y_i - \mathbf{x}_i\mathbf{x}_i'\hat{\mathbf{a}}_N) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i y_i - \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\hat{\mathbf{a}}_N = \mathbf{0}_{P\times 1}.$$

Ceci donne :

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i y_i = \left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]\hat{\mathbf{a}}_N$$

d'où finalement :

$$\hat{\mathbf{a}}_N = \left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i y_i = \left[\mathbf{X}'\mathbf{X}\right]^{-1}\mathbf{X}'\mathbf{y} = \hat{\mathbf{a}}_N^{MCO}.$$

1.2.1b. Montrer à l'aide des propriétés données dans l'annexe que $\hat{\mathbf{a}}_N^{MCO}$ ***est un estimateur convergent en probabilité de*** $\mathbf{a}_0$

- $\hat{\mathbf{a}}_N^{MCO}$ est un estimateur convergent en probabilité de $\mathbf{a}_0$ si :

$$\hat{\mathbf{a}}_N^{MCO} \xrightarrow[N\to+\infty]{p} \mathbf{a}_0$$

- Faisons d'abord apparaître $\mathbf{a}_0$ dans l'expression de $\hat{\mathbf{a}}_N^{MCO}$ avec $y_i = \mathbf{a}_0'\mathbf{x}_i + u_i$:

$$\begin{aligned}\hat{\mathbf{a}}_N^{MCO} &= \left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i y_i \\ &= \left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i(\mathbf{x}_i'\mathbf{a}_0 + u_i)\end{aligned}$$

Un simple réarrangement des termes donne :

$$
\begin{aligned}
\hat{\mathbf{a}}_N^{MCO} &= \left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i(\mathbf{x}_i'\mathbf{a}_0+u_i) \\
&= \left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]\mathbf{a}_0+\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_iu_i \\
&= \mathbf{a}_0+\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_iu_i
\end{aligned}
$$

Aussi, pour montrer que $\hat{\mathbf{a}}_N^{MCO}\xrightarrow[N\to+\infty]{p}\mathbf{a}_0$, il suffit de montrer que :

$$\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_iu_i\xrightarrow[N\to+\infty]{p}\mathbf{0}_{P\times 1}.$$

- Si la variance des $\mathbf{x}_i\mathbf{x}_i'$ est finie (ce qu'on suppose ici) alors, puisque les (y_i ,$\mathbf{x}_i$) sont i.i.d., la loi faible des grands nombres (*prop. A12b*) s'applique pour les $\mathbf{x}_i\mathbf{x}_i'$ et on a :

(a) $$\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\xrightarrow[N\to+\infty]{p}E\left[\mathbf{x}_i\mathbf{x}_i'\right].$$

- Si $E\left[\mathbf{x}_i\mathbf{x}_i'\right]$ est inversible (ce qui est nécessaire pour l'identification de $\mathbf{a}_0$) alors ce qu'on vient de montrer (a) implique que la *prop. A3* s'applique pour la suite $\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'$, on a alors :

(b) $$\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\xrightarrow[N\to+\infty]{p}E\left[\mathbf{x}_i\mathbf{x}_i'\right]^{-1}.$$

- Si la variance des $\mathbf{x}_iu_i$ est finie (ce qu'on suppose ici) alors, puisque les (y_i ,$\mathbf{x}_i$) sont i.i.d., la loi faible des grands nombres (*prop. A12b*) s'applique pour les $\mathbf{x}_iu_i$ et on a :

(c) $$\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_iu_i\xrightarrow[N\to+\infty]{p}E\left[\mathbf{x}_iu_i\right]=\mathbf{0}_{P\times 1},$$

<u>la dernière égalité provenant de l'exogénéité de $\mathbf{x}_i$ par rapport à u_i</u> (tout repose sur cette propriété).

- Avec les propriétés de convergence (b) et (c), on peut appliquer la *prop. A7* pour montrer que le produit des suites $\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}$ et $\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_iu_i$ converge vers le produit de leurs limites :

$$\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_iu_i\xrightarrow[N\to+\infty]{p}E\left[\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\times\mathbf{0}_{P\times 1}=\mathbf{0}_{P\times 1},$$

d'où finalement : $\hat{\mathbf{a}}_N^{MCO} \xrightarrow[N\to+\infty]{p} \mathbf{a}_0$.

- *Remarque* : Pour l'essentiel, on a utilisé des lois des grands nombres en s'appuyant principalement sur l'hypothèse : les (y_i ,$\mathbf{x}_i$) sont i.i.d. et sur l'exogénéité de $\mathbf{x}_i$.

- *Interprétation* : $\hat{\mathbf{a}}_N^{MCO} \xrightarrow[N\to+\infty]{p} \mathbf{a}_0$ indique que si N est grand alors la valeur calculée pour $\hat{\mathbf{a}}_N^{MCO}$ à partir des données de n'échantillon est proche de $\mathbf{a}_0$.

1.2.1b. Montrer à l'aide des propriétés données dans l'annexe que $\hat{\mathbf{a}}_N^{MCO}$ est asymptotiquement normal

- $\hat{\mathbf{a}}_N^{MCO}$ est as. normal (ou converge en racine de N) vers $\mathbf{a}_0$ si :

$$\sqrt{N}\left(\hat{\mathbf{a}}_N^{MCO} - \mathbf{a}_0\right) \xrightarrow[N\to+\infty]{L} \mathbf{N}\left(\mathbf{0}_{P\times 1}, \mathbf{\Sigma}_0\right).$$

- D'après ce qu'on vient de voir, on sait que :

$$\begin{aligned}\sqrt{N}\left(\hat{\mathbf{a}}_N^{MCO} - \mathbf{a}_0\right) &= \sqrt{N}\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i u_i \\ &= \left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{x}_i u_i\end{aligned}.$$

Il nous faut donc montrer que le terme $\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{x}_i u_i$ converge en loi vers une loi normale. Pour montrer la convergence d'une suite de var vers une loi normale, généralement on utilise un TCL, ce qui semble logique étant donnée la forme du terme qu'on considère.

- On vient de montrer que :

(b) $$\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1} \xrightarrow[N\to+\infty]{p} E\left[\mathbf{x}_i\mathbf{x}_i'\right]^{-1}.$$

- On sait que $E\left[\mathbf{x}_i u_i\right] = \mathbf{0}_{P\times 1}$ et que les (y_i ,$\mathbf{x}_i$) sont i.i.d., si de plus on sait que $\mathbf{W}_0 = V\left[\mathbf{x}_i u_i\right]$ est bien définie alors, il est possible d'appliquer le théorème central limite de Lindeberg-Levy (*prop A13*) pour la suite $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{x}_i u_i$:

(d) $\quad \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{x}_i u_i \xrightarrow[N\to+\infty]{L} \mathbf{N}\left(E\left[\mathbf{x}_i u_i\right];V\left[\mathbf{x}_i u_i\right]\right)=\mathbf{N}\left(\mathbf{0}_{P\times 1};\mathbf{W}_0\right).$

- Avec (b) et (d), il est possible d'appliquer les propriétés des combinaisons de convergence (*prop A8*). Avec *prop A8iii*, on a :

$$\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{x}_i u_i \xrightarrow[N\to+\infty]{L} E\left[\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\times\mathbf{N}\left(\mathbf{0}_{P\times 1};\mathbf{W}_0\right)$$

ce qui est équivalent à :

$$\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{x}_i u_i \xrightarrow[N\to+\infty]{L} \left(\mathbf{0}_{P\times 1};E\left[\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\mathbf{W}_0 E\left[\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\right).$$

Aussi, on a bien :

$$\sqrt{N}\left(\hat{\mathbf{a}}_N^{MCO}-\mathbf{a}_0\right)\xrightarrow[N\to+\infty]{L}\mathbf{N}\left(\mathbf{0}_{P\times 1},\mathbf{\Sigma}_0\right) \text{ avec } \mathbf{\Sigma}_0=E\left[\mathbf{x}_i\mathbf{x}_i'\right]^{-1}\mathbf{W}_0 E\left[\mathbf{x}_i\mathbf{x}_i'\right]^{-1}.$$

Interprétation : si N est grand, alors on n'a pas besoin de connaître la loi des u_i pour montrer que l'estimateur des MCO est normal, on sait d'après sa convergence en racine de N que :

$$\hat{\mathbf{a}}_N^{MCO}\simeq\mathbf{N}\left(\mathbf{a}_0,\frac{\mathbf{\Sigma}_0}{N}\right).$$

Remarque : $\sqrt{N}\left(\hat{\mathbf{a}}_N^{MCO}-\mathbf{a}_0\right)\xrightarrow[N\to+\infty]{L}\mathbf{N}\left(\mathbf{0}_{P\times 1},\mathbf{\Sigma}_0\right)$ implique que $\hat{\mathbf{a}}_N^{MCO}\xrightarrow[N\to+\infty]{p}\mathbf{a}_0$. La convergence en racine de N est donc une propriété plus « forte » que la convergence en probabilité pour un estimateur. Il arrive qu'on puisse montrer qu'un estimateur est convergent en probabilité sans qu'on puisse montrer qu'il est convergent en racine de N.

Remarque : La normalité as. de l'estimateur des MCO a été démontrée sans hypothèse sur l'homoscédasticité des u_i.