Add dyn_fval_dynamic_{residual_g1, g2, residual_g1_g2, g3}

1874

Exercice 1

DGP $y_t = \mu_1 + \mu_2 \mathbb{1}_{\{t>\kappa\}} + \varepsilon_t$

avec $\varepsilon_t \sim iid(0,\sigma_\varepsilon^2)$, $T > \kappa > 1$ un entier donné et $t = 1, \dots, T$

(1) On estime le modèle

$$y_t = c + u_t$$

Calculons l'estimateur des MCO de c

$$\hat{c} = \arg\min_{\{c\}} \sum_{t=1}^{T} (y_t - c)^2$$

en posant le programme, puis résolvant la CNO, on trouve

$$\hat{c} = \frac{1}{T}\sum_{t=1}^{T} y_t \equiv \bar{y}$$

(2) Calculons l'espérance de l'estimateur des MCO :

$$\mathbb{E}[\hat{c}] = \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[\mu_1 + \mu_2 \mathbb{1}_{\{t>\kappa\}} + \varepsilon_t\right]$$ en substituant le DGP

$$\Leftrightarrow \mathbb{E}[\hat{c}] = \frac{1}{T}\sum_{t=1}^{T} \mu_1 + \frac{1}{T}\sum_{t=\kappa+1}^{T} \mu_2$$ car $\mathbb{E}[\varepsilon_t] = 0 \ \forall t$

$$\Leftrightarrow \mathbb{E}[\hat{c}] = \mu_1 + \frac{T-\kappa}{T}\mu_2 \neq \mu_1 \text{ ou } \mu_2$$

(3) Quand T tend vers l'infini l'espérance tend vers $\mu_1 + \mu_2$.

$$\mathbb{V}[\hat{c}] = \mathbb{V}\left[\frac{1}{T}\sum_t \left(\mu_1 + \mu_2 \mathbb{1}_{\{t>\kappa\}} + \varepsilon_t\right)\right]$$

$$= \mathbb{V}\left[\frac{1}{T}\sum \varepsilon_t\right]$$ car la partie déterministe est de variance nulle

$$= \frac{1}{T^2}\mathbb{V}\left[\sum \varepsilon_t\right]$$

$$= \frac{1}{T^2}\sum_{t=1}^{T}\mathbb{V}[\varepsilon_t]$$ car ε_t est iid

$$= \frac{\sigma_\varepsilon^2}{T} \xrightarrow[T\to\infty]{} 0$$

Ainsi lorsque T tend vers l'infini $\hat{c}$ tend en probabilité vers $\mu_1 + \mu_2$

Exercice 2

DGP $y_t = \mu_1 + \mu_2 \mathbb{1}_{\{t > \kappa T\}} + \varepsilon_t$

avec $\varepsilon_t \sim iid(0,\sigma_\varepsilon^2)$, $\kappa \in]0,1[$ donné. On estime toujours $y_t = c + u_t$. Comment les propriétés de $\hat{c}$ sont-elles changées ?

On a toujours $\hat{c}_T = \frac{1}{T}\sum_{t=1}^{T} y_t$. Calculons l'espérance de l'estimateur des MCO :

$$\mathbb{E}[\hat{c}_T] = \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[\mu_1 + \mu_2 \mathbb{1}_{\{t>[\kappa T]\}} + \varepsilon_t\right]$$

$$\Leftrightarrow \mathbb{E}[\hat{c}_T] = \mu_1 + \frac{1}{T}\sum_{t=[\kappa T]+1}^{T} \mu_2 \qquad \text{car } \mathbb{E}[\varepsilon_t] = 0$$

$$\Leftrightarrow \mathbb{E}[\hat{c}_T] = \mu_1 + \frac{T-[\kappa T]}{T}\mu_2$$

$$\Rightarrow \lim_{T\to\infty} \mathbb{E}[\hat{c}_T] = \mu_1 + (1-\kappa)\mu_2$$

le comportement de la variance reste identique puisque le seul changement par rapport à l'exercice 1 concerne la partie déterministe. Ainsi

$$\hat{c}_T \xrightarrow[T\to\infty]{proba} \mu_1 + (1-\kappa)\mu_2$$

Exercice 3 Soit le modèle empirique

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_K x_{Ki} + u_i$$

(1) Écrivons ce modèle matriciellement. On pose $Y = (y_1, y_2, \dots, y_N)'$ un vecteur $N\times 1$

$$x_i = (1, x_{1i}, x_{2i}, \dots, x_{Ki}) \quad i = 1,\dots,N \text{ des vecteurs } 1\times(K+1)$$

$$\mathbb{X} = (x_1', x_2', \dots, x_N')' \text{ une matrice } N\times(K+1)$$

et $u = (u_1, u_2, \dots, u_N)$ un vecteur $N\times 1$

Le modèle peut alors s'écrire de façon équivalente sous la forme

$$\underset{N\times 1}{Y} = \underbrace{\underset{N\times(K+1)}{X}\underset{(K+1)\times 1}{\beta}}_{N\times 1} + \underset{N\times 1}{u} \qquad (*)$$

où $\beta = (\beta_0, \beta_1, \dots, \beta_K)$ le vecteur $(K+1)\times 1$

Rmq Il est important de toujours bien vérifier les dimensions des matrices et vecteurs et que les membres de gauche et de droite de (*) aient bien la même dimension.

(2) Calculons l'estimateur des MCO. Celui-ci est obtenu la somme des carré des

résidus

$$\hat{\beta}_N = \underset{\{\beta\}}{\arg\min}\ \underset{1\times N}{u'}\ \underset{N\times 1}{u}$$

Rmq l'objectif st bien un scalaire

$$\Leftrightarrow \hat{\beta}_N = \underset{\{\beta\}}{\arg\min}\ (Y - X\beta)'(Y - X\beta)$$

ne dépend pas de β

$$= \underset{\{\beta\}}{\arg\min}\ Y'Y - Y'X\beta - \beta'X'Y + \beta'X'X\beta$$

$$= \underset{\{\beta\}}{\arg\min}\ \beta'X'X\beta - Y'X\beta - \beta'X'Y$$

$$= \underset{\{\beta\}}{\arg\min}\ \underline{\beta'X'X\beta - 2\beta'X'Y}$$

car $Y'X\beta = (Y'X\beta)' = \beta'X'Y$, un scalaire st toujours égal à sa transposée!

Comment écrire la CNO? Nous devons dériver l'objectif (scalaire),

$$\underset{1\times 1}{S(\beta)} = \underset{1\times(K+1)}{\beta'}\ \underset{(K+1)\times(K+1)}{X'X}\ \underset{(K+1)\times 1}{\beta} - 2\ \underset{1\times(K+1)}{\beta'}\ \underset{(K+1)\times N}{X'}\ \underset{N\times 1}{Y}$$

par rapport à β...

On a :

$$dS = d\beta'X'X\beta + \beta'X'Xd\beta - 2d\beta'X'Y$$

variation infinitésimale de l'objectif

Toujours en exploitant le fait qu'un scalaire est invariant par transposition, on a encore

$$dS = d\beta'2X'X\beta - d\beta'2X'Y$$

$$\Leftrightarrow dS = \left(d\beta'X'X\beta - d\beta'X'Y\right) \times 2,$$

$$\Rightarrow \frac{dS}{d\beta'} = 2\left(X'X\beta - X'Y\right)$$

La CNO qui nous permet d'identifier l'estimateur des MCO s'écrit donc :

$$X'X\hat{\beta} - X'Y = 0$$

$$\Leftrightarrow \quad \hat{\beta} = (X'X)^{-1}X'Y$$

en supposant que $X'X$ st une matrice de plein rang.

Explicitons le contenu de cette expression matricielle, en faisant apparaître les observations.

$$X'Y = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_{11} & x_{12} & x_{13} & \cdots & x_{1N} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2N} \\ \vdots & & & & \vdots \\ x_{j1} & x_{j2} & x_{j3} & \cdots & x_{jN} \\ \vdots & & & & \vdots \\ x_{K1} & x_{K2} & x_{K3} & & x_{KN} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{pmatrix}$$

$$(=) \quad X'Y = \begin{pmatrix} y_1 + y_2 + \cdots + y_N \\ x_{11}y_1 + x_{12}y_2 + x_{13}y_3 + \cdots + x_{1N}y_N \\ x_{21}y_1 + x_{22}y_2 + x_{23}y_3 + \cdots + x_{2N}y_N \\ \vdots \\ x_{j1}y_1 + x_{j2}y_2 + x_{j3}y_3 + \cdots + x_{jN}y_N \\ \vdots \\ x_{K1}y_1 + x_{K2}y_2 + x_{K3}y_3 + \cdots + x_{KN}y_N \end{pmatrix}$$

$$(=) \quad X'Y = \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_{1i}y_i \\ \sum_{i=1}^{N} x_{2i}y_i \\ \vdots \\ \sum_{i=1}^{N} x_{ji}y_i \\ \vdots \\ \sum_{i=1}^{N} x_{Ki}y_i \end{pmatrix} \quad [1]$$

Pour $X'X$ c'est la même idée.

$$X'X = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{12} & \cdots & x_{1N} \\ \vdots & & & \\ x_{j1} & x_{j2} \cdots x_{jl} & \cdots & x_{jN} \\ \vdots & & & \\ x_{K1} & x_{K2} & \cdots & x_{KN} \end{pmatrix} \begin{pmatrix} 1 & x_{11} & \cdots x_{j1} \cdots & x_{K1} \\ 1 & x_{12} & x_{j2} & x_{K2} \\ & & \vdots & \vdots \\ & & x_{jl} & \\ & & \vdots & \\ 1 & x_{1N} \cdots & x_{jN} & x_{KN} \end{pmatrix}$$

$\sum_{i=1}^{N} 1$

$$\underset{(K+1)\times(K+1)}{X'X} = \begin{pmatrix} N & \sum_{i=1}^{N} x_{1i} & \cdots & \sum_{i=1}^{N} x_{ji} & \cdots & \sum_{i=1}^{N} x_{Ki} \\ \sum_{i=1}^{N} x_{1i} & \sum_{i=1}^{N} x_{1i}^2 & \cdots & \sum_{i=1}^{N} x_{1i}x_{ji} & & \sum_{i=1}^{N} x_{1i}x_{Ki} \\ \vdots & & & & & \\ \sum_{i=1}^{N} x_{ji} & \sum_{i=1}^{N} x_{ji}x_{1i} & \cdots & \sum_{i=1}^{N} x_{ji}^2 & \cdots & \sum_{i=1}^{N} x_{ji}x_{Ki} \\ \vdots & & & & & \vdots \\ \sum_{i=1}^{N} x_{Ki} & \sum_{i=1}^{N} x_{Ki}x_{1i} & \cdots & \sum_{i=1}^{N} x_{Ki}x_{ji} & \cdots & \sum_{i=1}^{N} x_{Ki}^2 \end{pmatrix} \quad [2]$$

Cette matrice est bien symétrique.

On obtient l'estimateur $\hat{\beta}$ en regroupant les expressions explicites obtenues en [1] et [2].

(3) Donnons une jolie expression de $\hat{\beta}$ dans le cas $K=1$ (une constante et une variable explicative). En appliquant ce que nous avons obtenu plus haut :

$$\hat{\beta} = \begin{pmatrix} N & \sum_{i=1}^{N} x_{1i} \\ \sum_{i=1}^{N} x_{1i} & \sum_{i=1}^{N} x_{1i}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_i y_i \end{pmatrix}$$

Il nous faut inverser une matrice 2×2 $(X'X)$, on a :

$$|X'X| = N\sum_{i=1}^{N} x_{1i}^2 - \left(\sum_{i=1}^{N} x_{1i}\right)^2$$

On peut montrer que ce déterminant est strictement positif dès lors que N est plus grand que 1 (et que les observations sont non nulles). Il suffit de montrer que

$$\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(x_{1i}-x_{1j})^2 = N\sum_{i=1}^{N} x_{1i}^2 - \left(\sum_{i=1}^{N} x_{1i}\right)^2$$

Comme une somme de carrés est toujours positive, cela implique que le déterminant doit être positif.

Preuve

$$\sum_{i=1}^{N}\sum_{j=1}^{N}(x_{1i}-x_{1j})^2 = \sum_{i=1}^{N}\sum_{j=1}^{N}\left(x_{1i}^2+x_{1j}^2-2x_{1i}x_{1j}\right)$$

$$= \sum_{j=1}^{N}\left(\sum_{i=1}^{N} x_i^2\right) + \sum_{i=1}^{N}\left(\sum_{j=1}^{N} x_j\right)^2 - 2\sum_{i=1}^{N}\sum_{j=1}^{N} x_i x_j$$

$$= 2N\sum_{i=1}^{N} x_{1i}^2 - 2\left(\sum_{i=1}^{N} x_{1i}^2 + 2\sum_{i=1}^{N}\sum_{j>i} x_{1i}x_{1j}\right)$$

$$= 2N\sum_{i=1}^{N} x_{1i}^2 - 2\left(\sum_{i=1}^{N} x_{1i}\right)^2$$

cqfd $\square$

Ainsi

$$(X'X)^{-1} = \frac{1}{N\sum_{i=1}^{N} x_{1i} - \left(\sum_{i=1}^{N} x_{1i}\right)^2} \underbrace{\begin{pmatrix} \sum_{i=1}^{N} x_{1i}^2 & -\sum_{i=1}^{N} x_{1i} \\ -\sum_{i=1}^{N} x_{1i} & N \end{pmatrix}}$$

la transposée (la matrice est symétrique) de la matrice des co-facteurs.

Il nous reste à postmultiplier par $X'Y$:

$$\hat{\beta} = \frac{1}{N\sum_{i=1}^{N} x_{1i}^2 - \left(\sum_{i=1}^{N} x_{1i}\right)^2} \begin{pmatrix} \sum_{i=1}^{N} x_{1i}^2 & -\sum_{i=1}^{N} x_{1i} \\ -\sum_{i=1}^{N} x_{1i} & N \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_{1i} y_i \end{pmatrix}$$

$$\Leftrightarrow \hat{\beta} = \frac{1}{N\sum_{i=1}^{N} x_{1i}^2 - \left(\sum_{i=1}^{N} x_{1i}\right)^2} \begin{pmatrix} \sum_{i=1}^{N} y_i \sum_{i=1}^{N} x_{1i}^2 - \sum_{i=1}^{N} x_{1i} \sum_{i=1}^{N} x_{1i} y_i \\ N\sum_{i=1}^{N} x_{1i} y_i - \sum_{i=1}^{N} x_{1i} \sum_{i=1}^{N} y_i \end{pmatrix}$$

$$\Rightarrow \hat{\beta} = \frac{1}{N\sum_{i=1}^{N} x_{1i}^2 - \left(N\bar{x}_1\right)^2} \begin{pmatrix} N\bar{y}\sum_{i=1}^{N} x_{1i}^2 - N\bar{x}_1 \sum_{i=1}^{N} x_{1i} y_i \\ N\sum_{i=1}^{N} x_{1i} y_i - N\bar{x}_1 N\bar{y} \end{pmatrix}$$

<u>Remarque 1</u> $\sum_{i=1}^{N} (x_{1i}-\bar{x}_1)^2 = \sum_{i=1}^{N} x_{1i}^2 + N\bar{x}_1^2 - 2\bar{x}_1 \sum_{i=1}^{N} x_{1i}$

$$\Rightarrow \sum_{i=1}^{N} (x_{1i}-\bar{x}_1)^2 = \sum_{i=1}^{N} x_{1i}^2 - N\bar{x}_1^2$$

En simplifiant par N, il vient donc :

$$\hat{\beta} = \frac{1}{\sum_{i=1}^{N} (x_{1i}-\bar{x}_1)^2} \begin{pmatrix} \bar{y}\sum_{i=1}^{N} x_{1i}^2 - \bar{x}_1 \sum_{i=1}^{N} x_{1i} y_i \\ \sum_{i=1}^{N} x_{1i} y_i - N\bar{x}_1\bar{y} \end{pmatrix}$$

<u>Remarque 2</u>

$$\sum_{i=1}^{N} (x_{1i}-\bar{x}_1)(y_i-\bar{y}) = \sum_{i=1}^{N} x_{1i} y_i - \bar{y}\sum_{i=1}^{N} x_{1i} - \bar{x}_1 \sum_{i=1}^{N} y_i + N\bar{x}_1\bar{y}$$

$$= \sum_{i=1}^{N} x_{1i} y_i - 2N\bar{y}\bar{x}_1 + N\bar{x}_1\bar{y}$$

$$= \sum_{i=1}^{N} x_{1i} y_i - N\bar{x}_1\bar{y}$$

On a donc :

$$\hat{\beta} = \frac{1}{\sum_{i=1}^{N} (x_{1i}-\bar{x})^2} \begin{pmatrix} \bar{y}\left(\sum_{i=1}^{N} (x_{1i}-\bar{x}_1)^2 + N\bar{x}_1^2\right) - \bar{x}_1\left(\sum_{i=1}^{N} (x_{1i}-\bar{x}_1)(y_i-\bar{y}) + N\bar{x}_1\bar{y}\right) \\ \sum_{i=1}^{N} (x_{1i}-\bar{x}_1)(y_i-\bar{y}) \end{pmatrix}$$

$$\Leftrightarrow \hat{\beta} = \begin{pmatrix} \bar{y} - \bar{x}_1 \dfrac{\sum_{i=1}^{N}(x_{1i}-\bar{x}_1)(y_i-\bar{y})}{\sum_{i=1}^{N}(x_{1i}-\bar{x}_1)^2} \\ \dfrac{\sum_{i=1}^{N}(x_{1i}-\bar{x}_1)(y_i-\bar{y})}{\sum_{i=1}^{N}(x_{1i}-\bar{x}_1)^2} \end{pmatrix}$$

estimateur de l'ordonnée à l'origine

estimateur de la pente

On a finalement

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1,i} + \hat{u}_i \qquad (*)$$

avec $\hat{\beta}_1 = \dfrac{\sum_{i=1}^{N}(x_{1i}-\bar{x}_1)(y_i-\bar{y})}{\sum_{i=1}^{N}(x_{1i}-\bar{x}_1)^2}$

et $\hat{\beta}_0 = \bar{y} - \hat{\beta}\bar{x}_1$

(4) On déduit directement que la somme des résidus estimés est nulle. En effet, en sommant (*) sur $i=1,\dots,N$, il vient :

$$\sum_{i=1}^{N} y_i = N\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{N} x_{1i} + \sum_{i=1}^{N} \hat{u}_i$$

$$\Leftrightarrow \sum_{i=1}^{N} \hat{u}_i = \sum_{i=1}^{N} y_i - N\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{N} x_{1i}$$

$$\Leftrightarrow \sum_{i=1}^{N} \hat{u}_i = N\underbrace{(\bar{y} - \hat{\beta}_0 - \hat{\beta}_1\bar{x}_1)}_{=0 \text{ par définition de } \hat{\beta}_0}$$

~~Exercice 4~~ Soit le modèle empirique

$$y_t = a + bt + u_t \qquad t=1,\dots,T$$

avec $u_t \sim iid(0,\sigma_\varepsilon^2)$. On obtient les estimateurs des MCO de a et b en suivant les mêmes étapes que dans l'exercice 3.

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} T & \sum_{t=1}^{T} t \\ \sum_{t=1}^{T} t & \sum_{t=1}^{T} t^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=1}^{T} y_t \\ \sum_{t=1}^{T} t y_t \end{pmatrix}$$

On sait que

$$\sum_{t=1}^{T} t = \frac{T(T+1)}{2} \quad \text{et} \quad \sum_{t=1}^{T} t^2 = \frac{T(2T+1)(T+1)}{6}$$

et donc en factorisant T :

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} 1 & \frac{T+1}{2} \\ \frac{T+1}{2} & \frac{(2T+1)(T+1)}{6} \end{pmatrix}^{-1} \begin{pmatrix} \bar{y} \\ \frac{1}{T}\sum_{t=1}^{T} t y_t \end{pmatrix}$$

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \frac{12}{(T+1)(T-1)} \begin{pmatrix} \frac{(2T+1)(T+1)}{6} & -\frac{T+1}{2} \\ -\frac{T+1}{2} & 1 \end{pmatrix} \begin{pmatrix} \bar{y} \\ \frac{1}{T}\sum_{t=1}^{T} t y_t \end{pmatrix}$$

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \frac{12}{(T+1)(T-1)} \begin{pmatrix} \frac{(2T+1)(T+1)}{6}\bar{y} - \frac{T+1}{2T}\sum_{t=1}^{T} t y_t \\ \frac{1}{T}\sum_{t=1}^{T} t y_t - \frac{T+1}{2}\bar{y} \end{pmatrix}$$

Exercice 4 L'estimateur de la constante (voir l'exercice 1) est

$$\hat{C}_T = \frac{1}{T}\sum_{t=1}^{T} y_t \equiv \bar{y}_T$$

Pour étudier les propriétés de cet estimateur ($\hat{C}_T$ est-il un estimateur sans biais de μ_0 ?) on substitue le DGP

$$y_t = \mu_0 + \mu_1 t + \varepsilon_t$$

dans l'expression de l'estimateur :

$$\hat{C}_T = \frac{1}{T}\sum_{t=1}^{T}(\mu_0 + \mu_1 t + \varepsilon_t)$$

$$= \mu_0 + \mu_1 \frac{1}{T}\sum t + \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t$$

$$= \mu_0 + \mu_1 \frac{T+1}{2} + \bar{\varepsilon}$$

Ainsi

$$\mathbb{E}[\hat{C}_T] = \mu_0 + \mu_1 \frac{T+1}{2} \neq \mu_0$$

$\hat{C}_T$ n'est pas un estimateur sans biais du paramètre μ_0 (l'ordonnée à l'origine). Le biais diverge (linéairement) quand la taille de l'échantillon (T) tend vers l'infini. Ce résultat tient à la non stationnarité de la variable analysée (→

parler du biais de variable omise en général et de l'hypothèse d'orthogonalité entre la variable omise et les variables non omises).

<u>Exercice 5</u> (1) Nous souhaitons montrer que :

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11}^{-1}(I + A_{12}BA_{21}A_{11}^{-1}) & -A_{11}^{-1}A_{12}B \\ -BA_{21}A_{11}^{-1} & B \end{pmatrix}$$

avec $B = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}$

où on suppose que A_{11} et $A_{22} - A_{21}A_{11}^{-1}A_{12}$ sont inversibles. Il y a deux façons (au moins) de le faire. En montrant que

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

ou en cherchant les matrices B_{11}, B_{12}, B_{21} et B_{22} telles que

$$\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

(2) Si $A_{12} = 0$ et $A_{21} = 0$ alors on dit que la matrice est <u>bloc diagonale</u> et on a directement (en appliquant le résultat obtenu plus haut) :

$$\begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22}^{-1} \end{pmatrix}$$

C'est une généralisation du résultat bien connu pour les matrices diagonales.

<u>Exercice 6</u> (1) Si les variables dans X_1 sont orthogonales aux variables dans X_2 alors par définition on doit avoir $X_1'X_2$ nul et $X_2'X_1$ nul (pour que la covariance soit nulle). On a alors :

$$X'X = \begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{pmatrix} = \begin{pmatrix} X_1'X_1 & 0 \\ 0 & X_2'X_2 \end{pmatrix}$$

On peut alors appliquer le résultat sur

l'inversion des matrices bloc diagonales (question 2 de l'exercice 5) pour calculer $\hat{\beta}_1$ et $\hat{\beta}_2$

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = (X'X)^{-1} X'Y$$

$$\Leftrightarrow \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} (X_1'X_1)^{-1} & 0 \\ 0 & (X_2'X_2)^{-1} \end{pmatrix} \begin{pmatrix} X_1 Y \\ X_2 Y \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} \hat{\beta}_1 = (X_1'X_1)^{-1} X_1' Y \\ \hat{\beta}_2 = (X_2'X_2)^{-1} X_2 Y \end{cases}$$

Autrement dit si on a deux groupes de variables orthogonales, on peut faire deux estimations séparées (pour chaque groupe de variables X_1 et X_2), ce sera équivalent à l'estimation de l'équation de toutes les variables. En creux, cela veut dire que si les variables dans X_2 (par exemple) sont omises, l'estimateur de β_1 dans

$$Y = X_1 \beta_1 + u$$

sera toujours un estimateur sans biais. On peut le vérifier

$$\hat{\beta}_1 = (X_1'X_1)^{-1} X_1' Y$$

le DGP

$$\Rightarrow \hat{\beta}_1 = (X_1'X_1)^{-1} X_1' (X_1\beta_1 + X_2\beta_2 + \varepsilon)$$

$$\Leftrightarrow \hat{\beta}_1 = \underbrace{(X_1'X_1)^{-1} X_1'X_1}_{=I} \beta_1 + (X_1'X_1)^{-1} \underbrace{X_1'X_2}_{=0} \beta_2 + (X_1'X_1)^{-1} X_1' \varepsilon$$

$$\Leftrightarrow \hat{\beta}_1 = \beta_1 + (X_1'X_1)^{-1} X_1' \varepsilon$$

$$\Rightarrow \mathbb{E}[\hat{\beta}_1] = \beta_1 + \mathbb{E}[(X_1'X_1)^{-1} X_1' \varepsilon]$$

Si X_1 est déterministe (identique d'un échantillon à l'autre) ou si X_1 est orthogonal à ε, alors l'estimateur est sans biais (le dernier terme est nul) :

$$\mathbb{E}[\hat{\beta}_1] = \beta_1$$

(2) Si X_1 et X_2 ne sont pas des matrices orthogonales, on ne peut plus utiliser le cas

particulier des matrices bloc diagonales pour inverser $X'X$. Calculons l'estimateur de β_1 en considérant le modèle complet. On a

$$\begin{pmatrix} \hat\beta_1 \\ \hat\beta_2 \end{pmatrix} = \begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X \end{pmatrix}^{-1} \begin{pmatrix} X_1Y \\ X_2Y \end{pmatrix}$$

$$= \begin{pmatrix} (X_1'X_1)^{-1}\left[I + X_1'X_2 \overbrace{(X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}X_1'X_2)^{-1}}^{B} X_2'X_1(X_1'X_1)^{-1}\right] & -(X_1'X_1)^{-1}X_1'X_2 B \\ -BX_2'X_1(X_1'X_1)^{-1} & B \end{pmatrix} \begin{pmatrix} X_1'Y \\ X_2'Y \end{pmatrix}$$

soit en considérant le premier bloc de lignes :

$$\hat\beta_1 = \underbrace{(X_1'X_1)^{-1}X_1'Y}_{\substack{\text{l'estimateur de } \beta_1 \\ \text{dans le cas où les} \\ \text{variables } X_2 \text{ sont} \\ \text{omises}}} + \underbrace{(X_1'X_1)^{-1}X_1'X_2 B X_2'X_1(X_1'X_1)^{-1}X_1'Y - (X_1'X_1)^{-1}X_1'X_2 B X_2'Y}_{\text{a priori non nul...}}$$

↳ Omettre des variables change l'estimateur de β_1.

Comment ? Pour le voir plus simplement, notons que

$$\begin{pmatrix} \hat\beta_1 \\ \hat\beta_2 \end{pmatrix} = \begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{pmatrix}^{-1} \begin{pmatrix} X_1'Y \\ X_2'Y \end{pmatrix}$$

$$\iff \begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{pmatrix} \begin{pmatrix} \hat\beta_1 \\ \hat\beta_2 \end{pmatrix} = \begin{pmatrix} X_1'Y \\ X_2'Y \end{pmatrix}$$

$$\iff \begin{cases} (X_1'X_1)\hat\beta_1 + X_1'X_2\hat\beta_2 = X_1'Y \\ (X_2'X_1)\hat\beta_1 + (X_2'X_2)\hat\beta_2 = X_2'Y \end{cases}$$

$$\iff \begin{cases} \hat\beta_1 = (X_1'X_1)^{-1}X_1'Y - (X_1'X_1)^{-1}X_1'X_2\hat\beta_2 & [1] \\ \hat\beta_2 = (X_2'X_2)^{-1}X_2'Y - (X_2'X_2)^{-1}X_2'X_1\hat\beta_1 & [2] \end{cases}$$

$$\hookrightarrow \hat\beta_1 = \overbrace{(X_1'X_1)^{-1}X_1'Y}^{\text{estimateur de } \beta \text{ dans le modèle avec variables omises}} - (X_1'X_1)^{-1}X_1'X_2\hat\beta_2$$

Omettre X_2 ne change rien, par rapport à l'estimateur du modèle complet, si et seulement si $X_1'X_2 = 0$

(les variables omises sont orthogonales aux variables non omises) ou si $\hat{\beta}_2 = 0$ (c'est-à-dire si Y n'est pas affecté par les variables omises, qui ne servent à rien).

(3) Les résidus estimés, dans le modèle complet, sont définis par :

$$\hat{u} = Y - X\hat{\beta}$$

soit en substituant l'expression de l'estimateur des MCO :

$$\hat{u} = Y - X(X'X)^{-1}X'Y$$

$$\Leftrightarrow \hat{u} = (I - X(X'X)^{-1}X')Y$$

$$\Leftrightarrow \hat{u} = M.Y$$

(4) Montrons que la matrice M est idempotente :

$$M.M = (I - X(X'X)^{-1}X')(I - X(X'X)^{-1}X')$$

$$\Leftrightarrow M.M = I - X(X'X)^{-1}X' - X(X'X)^{-1}X' + X(X'X)^{-1}\underbrace{X'X(X'X)^{-1}}_{I}X'$$

$$\Leftrightarrow MM = I - 2X(X'X)^{-1}X' + X(X'X)^{-1}X'$$

$$\Leftrightarrow M.M = I - X(X'X)^{-1}X' \equiv M$$ cqfd

(5) Montrons que M est orthogonale à X, c'est-à-dire que $M.X = 0$

$$M.X = (I - X(X'X)^{-1}X')X$$

$$\Leftrightarrow M.X = X - X\underbrace{(X'X)^{-1}X'X}_{I}$$

$$\Leftrightarrow M.X = X - X$$

$$\Leftrightarrow M.X = 0$$ cqfd

(6) La prédiction sur la variable expliquée est

$$\hat{Y} = X\hat{\beta}$$

$$\Leftrightarrow \hat{Y} = X(X'X)^{-1}X'Y$$

$$\Leftrightarrow \hat{Y} = \underbrace{(I - M)}_{H}Y$$

la prédiction et linéaire en Y. On montre facilement que H et aussi idempotente

$$H.H = (I-M)(I-M)$$
$$\iff H.H = I - 2M + M.M$$
$$\iff H.H = I - 2M + M$$
$$\iff H.H = I - M \equiv M$$

(7) En substituant [1] dans [2], on obtient une équation nous permettant d'identifier $\hat{\beta}_2$

$$\hat{\beta}_2 = (X_2'X_2)^{-1}X_2'Y - (X_2'X_2)^{-1}X_2'X_1\underbrace{(X_1'X_1)^{-1}X_1'Y}_{\hat{\beta}_1} + (X_2'X_2)^{-1}X_2'X_1(X_1'X_1)^{-1}X_1'X_2\hat{\beta}_2$$

$$\iff (X_2'X_2)\hat{\beta}_2 = X_2'Y - X_2'X_1(X_1'X_1)^{-1}X_1'Y + X_2'X_1(X_1'X_1)^{-1}X_1'X_2\hat{\beta}_2$$

$$\iff X_2'\underbrace{(I - X_1(X_1'X_1)^{-1}X_1)}_{M_1}X_2\hat{\beta}_2 = X_2'\underbrace{(I - X_1(X_1'X_1)^{-1}X_1')}_{M_1}Y$$

$$\iff (X_2'M_1X_2)\hat{\beta}_2 = X_2'M_1Y$$

$$\iff \hat{\beta}_2 = (X_2'M_1X_2)^{-1}X_2'M_1Y$$

où la matrice de projection M_1

$$M_1 = I - X_1(X_1'X_1)^{-1}X_1$$

retourne les résidus de la régression de Y sur les variables X_1. Comme cette matrice est idempotente et symétrique, on a

$$\hat{\beta}_2 = (X_2'M_1M_1X_2)^{-1}X_2'M_1M_1Y$$
$$\iff \hat{\beta}_2 = (X_2'M_1'M_1X_2)^{-1}X_2'M_1'M_1Y$$
$$\iff \hat{\beta}_2 = ((M_1X_2)'M_1X_2)^{-1}(M_1X_2)'M_1Y$$

$$\iff \hat{\beta}_2 = (\tilde{X}_2'\tilde{X}_2)^{-1}\tilde{X}_2'\tilde{Y}$$

où $\tilde{X}_2$ sont les résidus des regressions des variables X_2 sur X_1 et $\tilde{Y}$ le résidu de la régression de Y sur X_1.

On peut donc estimer β_2 en trois étapes :
(i) filtrer de l'effet de X_1 sur X_2 en calculant

les résidus des régressions de X_2 avec X_1 $\tilde{U}_2$

(ii) Calculer les résidus de la régression de Y avec X_1 : $\hat{U}_1$

(iii) Estimer le modèle (par les MCO)

$$\hat{U}_1 = \beta_2 \tilde{U}_2 + V$$

L'estimation de β_2 dans ce modèle est identique à celle que nous obtiendrions dans le modèle complet (les propriétés statistiques de l'estimateur aussi). C'est le théorème de Frisch_Waugh. L'intérêt de cette approche est que nous obtenons une estimation en inversant des plus petites matrices (c'était important il y a cinquante ans, mais aujourd'hui ce n'est plus un avantage décisif).