



Deuxième année
2005-2005

Séries temporelles linéaires
Corrigé des travaux dirigés n°7

Guillaume Lacôte

☞ Bureau **E03**

✉ Guillaume.Lacote@ensae.fr

☞ <http://ensae.no-ip.com/SE206/>

Corrigé de l'exercice 1

- ☞ Q1 L'observation de la série fait apparaître une tendance déterministe linéaire : la série observée est "assez proche" d'une certaine droite $\bar{y}_t = a + bt$.
Si l'on admet, comme le suggère son graphe, qu'à cette composante déterministe près elle est en outre stationnaire du second ordre, alors le processus Y dont elle est la réalisation admet une représentation auto-régressive infinie de la forme

$$\Phi(L)(Y_t - a - bt) = \epsilon_t$$

Or, à supposer que $\Phi(\mathbb{X}) = \sum_{k=0}^{+\infty} \phi_k \mathbb{X}^k$ ainsi que sa dérivée sont de rayon supérieur à 1 on a

$$\begin{aligned} \Phi(L) \circ (bt)_t &= b \sum_{k=0}^{+\infty} \phi_k (t - k) \\ &= b \left(\sum_{k=0}^{+\infty} \phi_k \right) t - b \left(\sum_{k=0}^{+\infty} k \phi_k \right) \\ &= b\Phi(1)t - b\Phi'(1) \end{aligned}$$

Comme $\Phi(L) \circ (a) = \Phi(1) \times a$ il vient en notant $\mu = a - b\Phi'(1)$ et $\beta = b\Phi(1)$

$$\Phi(L)Y_t = \mu + \beta t + \epsilon_t$$

En pratique, on se contente de choisir Φ de degré fini mais assez grand pour que ϵ soit significativement un bruit blanc.

- ☞ Q2 Afin de tester la présence éventuelle d'une racine unitaire, on cherche à factoriser "de forme" $\Phi(L)$ par $1 - L$ (transformation de BEVERIDGE-NELSON, voir TD 4, exercice 1) : pour ce faire on effectue la division euclidienne $\Phi(\mathbb{X})$ par $1 - \mathbb{X}$ ce qui s'écrit, comme 1 est une racine de $\Phi(\mathbb{X}) - \Phi(1)$ ¹

$$\Phi^+(\mathbb{X}) = \frac{\Phi(\mathbb{X}) - \Phi(1)}{(1 - \mathbb{X})} \in \mathbb{R}[X], \quad d^\circ \Phi^+ \leq d^\circ \Phi - 1$$

Alors

$$\Phi(\mathbb{X}) = \Phi(1) + (1 - \mathbb{X})\Phi^+(\mathbb{X})$$

¹Cette opération n'est normalement définie que pour les polynômes, i.e. lorsque Φ est de degré fini. Si $d^\circ \Phi = +\infty$ on définit $\Phi_N(\mathbb{X}) = \sum_{k=0}^N \phi_k \mathbb{X}^k$ dont on tire $\Phi_N^+(\mathbb{X})$ pour $N \geq 1$, dont on montre qu'elle est convergente de rayon supérieur à 1.

et donc

$$\begin{aligned} \Phi(L)Y &= \Phi(1) \times Y + (\mathbf{1} - L) \cdot \Phi^+(L) \circ Y \\ &= \Phi(1) \times Y + \Phi^+(L) \circ \Delta Y \\ &= \Phi(1) \times Y + \underbrace{\Phi^+(0)}_{\Phi(0) - \Phi(1)} \times \Delta Y + \left(\underbrace{\Phi^+(L) - \Phi^+(0)}_{-\Phi^*(L)} \right) \circ \Delta Y \\ &= \Phi(1) \times Y + (1 - \Phi(1)) \times \Delta Y - \Phi^*(L) \circ \Delta Y \\ &= \Delta Y + \Phi(1)LY - \Phi^*(L) \circ \Delta Y \end{aligned}$$

et comme par ailleurs $\Phi(L)Y_t = \mu + \beta t + \epsilon_t$ il vient en définitive

$$\Delta Y_t = \mu + \beta t + \Phi^*(L)\Delta Y_t - \Phi(1)Y_{t-1} + \epsilon_t$$

qui est bien le modèle recherché en posant $\phi = -\Phi(1)$.

Notons que par construction $\Phi^*(0) = 0$, donc $\Phi^*(\mathbb{X})$ ne contient que des puissances strictement positives de \mathbb{X} , et donc $\Phi^*(L) \circ \Delta Y_t$ ne contient que des valeurs passées $\Delta Y_{t-1}, \Delta Y_{t-2}, \dots$ de ΔY_t .

Tester si le modèle est intégré, c'est tester si $\Phi(1) = 0$; sous cette forme cela revient juste à tester si le coefficient ϕ de Y_{t-1} est non-significativement non-nul.

Il aurait été possible de pratiquer une telle régression directement sur l'équation en Y_t au lieu de celle en ΔY_t ; mais comme Y n'est pas stationnaire, $Cov(Y_t, Y_{t+h})$ est significative pour tout h , donc la régression aurait dû comporter une infinité de variables explicatives $(Y_{t-h})_{h \geq 1} \dots$. A l'inverse si ΔY est stationnaire, un nombre fini d'explicatives peut suffire.

☞ Q3 Il est tout d'abord nécessaire de déterminer Φ tel que le modèle soit vérifié. Pour ce faire on choisit un degré arbitrairement grand pour Φ (ici 9)² et on vérifie que le modèle estimé par la régression des moindres carrés ordinaires³ est statistiquement valide; dans le cas présent, le R^2 ajusté (qui vaut 0,5129) assure que le modèle capture une part suffisante de la variance, et donc que suffisamment de variables explicatives $\Delta Y_{t-1}, \Delta Y_{t-2}, \dots$ ont été introduites.

Remarquons que lorsque ΔY est stationnaire, les coefficients de la régression sont donc asymptotiquement normaux, ce qui justifie l'emploi du test de Student pour tester leur significativité.⁴

En l'occurrence, on observe que les coefficients de la régression sur $\Delta Y_{t-4}, \dots, \Delta Y_{t-8}$ ne sont (individuellement) pas significativement non-nuls (colonne Prob > |t|).

Ainsi la modélisation proposée est valide mais le modèle estimé ici n'est pas significatif.

- ☞ Q4 (a)
- (b)
- (c)

²Car $8 = d^o \Phi^* = d^o \Phi^+ = d^o \Phi - 1$.

³Puisque ϵ est censé être un bruit blanc.

⁴En toute généralité il serait possible que ΔY ne soit pas stationnaire, auquel cas l'emploi du test de Student serait inadapté.

- (d)
- (e)
- (f)
- (g)

(h) On se propose donc de déterminer le sous-modèle minimal qui soit totalement significatif.

Remarque : Bien que l'on observe à ce stade que le coefficient en Y_{t-1} est significativement non-nul, **on ne peut pas** en déduire qu'il le sera aussi dans le sous-modèle où tous les coefficients sont significatifs.⁵ Il est nécessaire de déterminer ce sous-modèle et d'y tester la significativité du coefficient en Y_{t-1} .

- Compte-tenu de la régression précédente, on teste la nullité **simultanée** des coefficients de $\Delta Y_{t-4}, \dots, \Delta Y_{t-8}$ dans la régression. En l'occurrence, l'hypothèse $\mathcal{H}_0 : \phi_{\Delta Y_{t-4}} \dots = \phi_{\Delta Y_{t-8}} = 0$ est acceptée car la statistique de Fisher associée vaut **0,7841**.
- On effectue donc la régression de ΔY sur $\langle 1, t, Y_{t-1}, \Delta Y_{t-1}, \dots, \Delta Y_{t-4} \rangle$, et on observe que le modèle ainsi estimé est valide (ce qui est heureux). Il apparaît cependant que les coefficients de ΔY_{t-2} et de ΔY_{t-3} sont **individuellement** non-significativement non-nuls.
- On effectue donc le test de la nullité simultanée de ces coefficients en formulant l'hypothèse $\mathcal{H}_0 : \phi_{\Delta Y_{t-2}} = \dots = \phi_{\Delta Y_{t-3}} = 0$.

Remarque : Noter que tester successivement

- i. la nullité simultanée des coefficients de $\Delta Y_{t-4}, \dots, \Delta Y_{t-8}$ dans la régression sur $\langle 1, t, Y_{t-1}, \Delta Y_{t-1}, \dots, \Delta Y_{t-8} \rangle$
- ii. puis la nullité simultanée des coefficients de ΔY_{t-2} et ΔY_{t-3} dans la régression sur $\langle 1, t, Y_{t-1}, \Delta Y_{t-1}, \dots, \Delta Y_{t-3} \rangle$

est moins robuste statistiquement que de tester directement la nullité simultanée des coefficients de $\Delta Y_{t-2}, \dots, \Delta Y_{t-8}$ dans la régression sur $\langle 1, t, Y_{t-1}, \Delta Y_{t-1}, \dots, \Delta Y_{t-8} \rangle$.

En l'occurrence, l'hypothèse est acceptée car la statistique de Fisher associée vaut **0,6597**

⁵La statistique de Student du coefficient β dans la régression ordinaire

$$Y = \alpha X^1 + \beta X^2 + \gamma X^3 + U$$

est

$$t_{X^2} = \frac{\hat{\beta} - \beta_0}{\sqrt{\widehat{\sigma}_2^2}}$$

où $\widehat{\sigma}_2^2$ désigne le terme diagonal de la matrice de variance-covariance Σ de $\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$.

Or rien ne dit que cette matrice ne dépend pas de X^3 , même si γ est non-significativement non-nul : comme $(X^2' X^2)^{-1} X^2' Y$ et $\hat{\gamma} = (X^3' X^3)^{-1} X^3' Y$ pour peu que $Cov(X^2, X^3) \neq 0$, la matrice Σ dépendra de $\hat{\gamma}$ (au sens $Cov(\Sigma^{1:3}, \hat{\gamma}) \neq 0$).

En l'occurrence, on sait (voir TD 2, exercice 3) que ΔY est stationnaire que ΔY_{t-1} et ΔY_{t-2} sont certainement corrélés.

– On effectue donc en définitive la régression de ΔY sur $\langle 1, t, Y_{t-1}, \Delta Y_{t-1} \rangle$, et on observe que le modèle ainsi estimé est valide. On constate en outre qu’aucun coefficient n’est individuellement non-significatif.

On procède donc à l’identification du modèle AR (voir TD 5, exercice 1) en ΔY sur les explicatives t et LY .

On observe en l’occurrence que les auto-corrélations directes décroissent “exponentiellement” vers 0 (en l’occurrence, elles ne sont plus significativement non-nulles dès l’ordre $h \geq 3$); en outre le test de Porte-Manteau permet de conclure que le résidu estimé ϵ est non-significativement différent d’un bruit blanc.

☞ Q5 Le modèle estimé s’écrit

$$\Delta Y_t = \mu + \beta t + \phi Y_{t-1} + \phi_1 \Delta Y_{t-1} + \epsilon_t$$

où ϕ_1 est le coefficient en \mathbb{X} de $\Phi^*(\mathbb{X})$, ce qui se réécrit encore

$$(\mathbf{1} - \rho L)Y_t = \mu + \beta t + \phi_1 \Delta Y_{t-1} + \epsilon_t$$

en posant $\rho = 1 + \phi$.

- ☞ Q6 (a)
 (b)
 (c)
 (d)
 (e) Sous \mathcal{H}_0^1 le modèle se réécrit

$$Y_t = \mu + Y_{t-1} + \phi_1 \Delta Y_{t-1} + \epsilon_t$$

La statistique de test associée est celle de la table VI, selon laquelle le fractile à 5% vaut 6,34 (on retient la valeur pour 250 observations⁶); comparée à la valeur empirique de F-Value : 5,83 on accepte donc l’hypothèse \mathcal{H}_0^1 .

- ☞ Q7 On constate que $\mathcal{H}_0^2 \rightarrow \mathcal{H}_0^1$; comme \mathcal{H}_0^1 est acceptable on se propose de tester \mathcal{H}_0^2 .
 Sous \mathcal{H}_0^2 le modèle se réécrit

$$Y_t = Y_{t-1} + \phi_1 \Delta Y_{t-1} + \epsilon_t$$

La statistique de test associée est celle de la table V, selon laquelle le fractile à 5% vaut 4,75; comparée à la valeur empirique F-Value : 10,75 on rejette donc l’hypothèse \mathcal{H}_0^2 .

*
 * *

Corrigé de l’exercice 2

⁶Plutôt que celle pour 100 observations, car le test est ainsi plus restrictif : si on accepte \mathcal{H}_0 avec le fractile à 250 observations, on l’accepterait a fortiori avec celui à 100 observations.

- ☞ Q1 (a) Le modèle étant sous forme canonique, ϵ est l’innovation du processus vectoriel X ; par définition de l’innovation $\epsilon_t = X_t - \mathbb{E}\mathbb{L}(X_t | X_{t-1}, \dots)$ et donc en projetant pour $i \in \{1, \dots, t-1\}$

$$\epsilon_t^i = X_t^i - \mathbb{E}\mathbb{L}(X_t^i | X_{t-1}, \dots)$$

(ϵ^2 est l’innovation du processus $X^2 | X^1$)

$$\begin{aligned} \mathbb{E}\mathbb{L}(X_t^1 | X_t^2, X_{t-1}, \dots) &= \mathbb{E}\mathbb{L}(X_t^1 | (X_t^2, X_{t-1}^2, \dots), (X_{t-1}^1, \dots)) \\ &= \mathbb{E}\mathbb{L}(X_t^1 | (\epsilon_t^2, \epsilon_{t-1}^2, \dots), (X_{t-1}^1, \dots)) \\ &= \mathbb{E}\mathbb{L}(X_t^1 | \epsilon_t^2, (X_{t-1}^2, \dots, X_{t-1}^1, \dots)) \\ &= \mathbb{E}\mathbb{L}(X_t^1 | \epsilon_t^2, X_{t-1}, \dots) \end{aligned}$$

Or $\forall t \in \mathbb{Z}$, $\epsilon_t^2 \perp\!\!\!\perp \langle X_{t-1}, \dots \rangle$ donc $\langle \epsilon_t^2, X_{t-1}, \dots \rangle = \langle \epsilon_t^2 \rangle \oplus \langle X_{t-1}, \dots \rangle$ et donc

$$\mathbb{E}\mathbb{L}(X_t^1 | X_t^2, X_{t-1}, \dots) = \mathbb{E}\mathbb{L}(X_t^1 | \epsilon_t^2) \oplus \mathbb{E}\mathbb{L}(X_t^1 | X_{t-1}, \dots)$$

Or

$$X_t = \sum_{k=1}^{d^\circ \Phi} \Phi_k X_{t-k} + \epsilon_t$$

et $\epsilon_t \perp\!\!\!\perp \langle X_{t-1}, \dots \rangle$ donc en projetant

$$\mathbb{E}\mathbb{L}(X_t^1 | \epsilon_t^2) = \mathbb{E}\mathbb{L}(\epsilon_t^1 | \epsilon_t^2)$$

et donc finalement

$$\mathbb{E}\mathbb{L}(X_t^1 | X_t^2, X_{t-1}, \dots) = \mathbb{E}\mathbb{L}(X_t^1 | X_{t-1}, \dots) + \mathbb{E}\mathbb{L}(\epsilon_t^1 | \epsilon_t^2)$$

- (b) Par définition pour $t \in \mathbb{Z}$

$$\begin{aligned} \boxed{X_t^2 \not\sim X_t^1} &\Leftrightarrow \mathbb{E}\mathbb{L}(X_t^1 | X_t^2, X_{t-1}, \dots) = \mathbb{E}\mathbb{L}(X_t^1 | X_{t-1}, \dots) \\ &\Leftrightarrow \mathbb{E}\mathbb{L}(\epsilon_t^1 | \epsilon_t^2) = 0 \\ &\Leftrightarrow \text{Cov}(\epsilon_t^1, \epsilon_t^2) = (0) \\ &\Leftrightarrow \Sigma^{1,2} = (0) \\ &\Leftrightarrow \text{Cov}(\epsilon_t^1, \epsilon_t^2) = (0) \\ &\Leftrightarrow \mathbb{E}\mathbb{L}(\epsilon_t^2 | \epsilon_t^1) = 0 \\ &\Leftrightarrow \mathbb{E}\mathbb{L}(X_t^2 | X_t^1, X_{t-1}, \dots) = \mathbb{E}\mathbb{L}(X_t^2 | X_{t-1}, \dots) \\ &\Leftrightarrow \boxed{X_t^2 \not\sim X_t^1} \end{aligned}$$

- ☞ Q2 (a) On a pour $t \in \mathbb{Z}$

$$X_t = \sum_{k=1}^{d^\circ \Phi} \phi_k X_{t-k} + \epsilon_t$$

et donc en projetant

$$X_t^1 = \sum_{k=1}^{d^\circ \Phi} \phi_k^{1,1} X_{t-k}^1 + \sum_{k=1}^{d^\circ \Phi} \phi_k^{1,2} X_{t-k}^2 + \epsilon_t^1$$

Si $\Phi(\mathbb{X})^{1,2} = (0)$ alors

$$X_t^1 = \sum_{k=1}^{d^\circ \Phi} \phi_k^{1,1} X_{t-k}^1 + \epsilon_t^1$$

et donc

$$\mathbb{E}\mathbb{L}(X_t^1 | X_{t-1}, \dots) = \sum_{k=1}^{d^\circ \Phi} \phi_k^{1,1} X_{t-k}^1 = \boxed{\mathbb{E}\mathbb{L}(X_t^1 | X_{t-1}, \dots)}$$

c'est-à-dire $X^2 \not\approx X^1$.

(b) Comme ϵ est l'innovation de X , soit $(A_k)_{k \in \mathbb{N}} \in \mathcal{M}_n(\mathbb{R})^{\mathbb{N}}$ telle que

$$\forall t \in \mathbb{Z}, X_t = \sum_{k=0}^{+\infty} A_k \epsilon_{t-k}$$

Alors en projetant il vient ⁷

$$\forall t \in \mathbb{Z}, X_t^1 = \sum_{k=0}^{+\infty} A_k^{1,1} \epsilon_{t-k}^1 + \sum_{k=0}^{+\infty} A_k^{1,2} \epsilon_{t-k}^2$$

(c) Supposons alors que $X^2 \approx X^1$; alors pour tout $t \in \mathbb{Z}$ $\mathbb{E}\mathbb{L}(X_t^1 | X_{t-1}, \dots) = \mathbb{E}\mathbb{L}(X_t^1 | X_{t-1}^1, \dots)$ et donc $X_t^1 - \mathbb{E}\mathbb{L}(X_t^1 | X_{t-1}^1, \dots) = X_t^1 - \mathbb{E}\mathbb{L}(X_t^1 | X_{t-1}, \dots) = \epsilon_t^1 : \epsilon^1$ est donc l'innovation de X^1 .

(d) Si $X^2 \not\approx X^1$, alors X^1 admet une représentation de Wold sous la forme

$$X_t^1 = \sum_{k=0}^{+\infty} B_k \epsilon_{t-k}^1$$

où $B_k \in \mathcal{M}_m(\mathbb{R})$. On a par ailleurs pour tout $t \in \mathbb{Z}$

$$X_t^2 = \sum_{k=0}^{+\infty} A_k^{2,1} \epsilon_{t-k}^1 + \sum_{k=0}^{+\infty} A_k^{2,2} \epsilon_{t-k}^2$$

et donc

$$\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} \sum_{k=0}^{+\infty} B_k & (0) \\ \sum_{k=0}^{+\infty} A_k^{2,1} & \sum_{k=0}^{+\infty} A_k^{2,2} \end{pmatrix} \times \begin{pmatrix} \epsilon_{t-k}^1 \\ \epsilon_{t-k}^2 \end{pmatrix}$$

⁷Chacune des deux sommes converge car $\|A_k^{1,j}\|_2^2 \leq \|A_k\|_2^2$ et que $\sum_k A_k$ est normalement convergente.

soit ⁸

$$X = \sum_{k=0}^{+\infty} \begin{pmatrix} B_k & (0) \\ A_k^{2,1} & A_k^{2,2} \end{pmatrix} L^k \circ \epsilon$$

Alors

$$\epsilon = \Phi(L)X = \Phi(L) \left(\sum_{k=0}^{+\infty} \begin{pmatrix} B_k & (0) \\ A_k^{2,1} & A_k^{2,2} \end{pmatrix} L^k \right) \circ \epsilon$$

et donc la matrice $\left(\sum_{k=0}^{+\infty} \begin{pmatrix} B_k & (0) \\ A_k^{2,1} & A_k^{2,2} \end{pmatrix} \mathbb{X}^k \right)$ sur l'anneau $\mathcal{M}_n(\mathbb{R})[\mathbb{X}]$ est régulière et est une inverse de $\Phi(\mathbb{X})$, ce qui s'écrit encore par blocs

$$\begin{pmatrix} \Phi(\mathbb{X})^{1,1} & \Phi(\mathbb{X})^{1,2} \\ \Phi(\mathbb{X})^{2,1} & \Phi(\mathbb{X})^{2,2} \end{pmatrix} \times \begin{pmatrix} \left(\sum_{k=0}^{+\infty} B_k \mathbb{X}^k \right) & (0) \\ \left(\sum_{k=0}^{+\infty} A_k^{2,1} \mathbb{X}^k \right) & \left(\sum_{k=0}^{+\infty} A_k^{2,2} \mathbb{X}^k \right) \end{pmatrix} = \begin{pmatrix} Id_m & (0) \\ (0) & Id_{n-m} \end{pmatrix}$$

et en particulier en considérant le bloc (1, 2) il vient

$$(0) \times \Phi(\mathbb{X})^{1,1} + \left(\sum_{k=0}^{+\infty} A_k^{2,2} \mathbb{X}^k \right) \times \Phi(\mathbb{X})^{1,2} = (0)$$

Or $\begin{pmatrix} \sum_{k=0}^{+\infty} B_k \mathbb{X}^k & (0) \\ \sum_{k=0}^{+\infty} A_k^{2,1} \mathbb{X}^k & \sum_{k=0}^{+\infty} A_k^{2,2} \mathbb{X}^k \end{pmatrix}$ est inversible, donc $\left(\sum_{k=0}^{+\infty} A_k^{2,2} \mathbb{X}^k \right)$ également ⁹ donc

$\Phi(\mathbb{X})^{1,2} = (0) \in \mathcal{M}_n(\mathbb{R})[\mathbb{X}]$ et finalement

$$\boxed{\Phi(L)^{1,2} = (0)}$$

ce qui achève la preuve de la réciproque.

☞ Q1 On sait que

$$\Phi(L)X = \epsilon$$

donc pour tout $\alpha \in \mathbb{R}$ en multipliant par $\begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix}$ il vient

$$\left(\begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix} \times \Phi(L) \right) \circ X = \begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix} \times \epsilon$$

⁸Avec la convention que

$$ML^k = \begin{pmatrix} M^{1,1}L^k & \dots & M^{1,n}L^k \\ \vdots & & \vdots \\ M^{n,1}L^k & \dots & M^{n,n}L^k \end{pmatrix}$$

et en remarquant que chaque $\sum_k A_k^{i,j} \mathbb{X}^k$ est convergente car $|A_k^{i,j}| \leq \|A_k\|_2$ et que $\sum_k A_k$ est normalement convergente

⁹Comme la matrice $\begin{pmatrix} A & 0 \\ B & C \end{pmatrix}$ est triangulaire par blocs, si elle est inversible, alors C l'est également.

Notons donc $\eta = \begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix} \times \epsilon$; alors pour tout $t \in \mathbb{Z}$ et $l \in \mathbb{Z}$

$$\text{Cov}(\eta_t, \eta_{t-l}) = \begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix}' \text{Cov}(\epsilon_t, \epsilon_{t-l}) \begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix}$$

Posons en particulier $\alpha = \frac{\Sigma^{1,2}}{\Sigma^{1,1}}$: alors $\text{Cov}(\eta_t, \eta_{t-l}) = 0$ et donc η est un bruit blanc. Sa matrice de variance-covariance est alors

$$\begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix} \Sigma \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \Sigma^{1,1} & 0 \\ 0 & \Sigma^{2,2} - \frac{(\Sigma^{1,2})^2}{\Sigma^{1,1}} \end{pmatrix}$$

Notons enfin $\Psi(\mathbb{X})$ la matrice à coefficients dans $\mathbb{R}[\mathbb{X}] \begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix} \times \Phi(\mathbb{X})$; alors

$$\Psi(0) = \begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix} \times \Phi(0) = \begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix}$$

(car $\Phi(0) = Id$), donc $\Psi(0)$ est bien inversible et le modèle est correctement spécifié.

Q2 D'une façon générale, l'estimateur des moindres carrés ordinaires de la régression de $Y = Xb + U$ est, si X est de plein rang, $\hat{b}_{mcg} = (X'\Omega^{-1}X)^{-1} X' \times \Omega^{-1}Y$, où Ω désigne la matrice de variance-covariance de U .

Considérons alors les deux sous-blocs

$$\begin{pmatrix} Y^1 \\ Y^2 \end{pmatrix} = \begin{pmatrix} Xb^1 \\ Xb^2 \end{pmatrix} + \begin{pmatrix} U^1 \\ U^2 \end{pmatrix}$$

et supposons que $\mathbb{V} \begin{pmatrix} U^1 \\ U^2 \end{pmatrix} = \begin{pmatrix} \Omega^{1,1} & (0) \\ (0) & \Omega^{2,2} \end{pmatrix}$ soit bloc-diagonale.

Alors

$$\begin{aligned} \hat{b}_{mcg} &= (X'\Omega^{-1}X)^{-1} X' \times \Omega^{-1} \begin{pmatrix} Y^1 \\ Y^2 \end{pmatrix} \\ &= \left(X' \begin{pmatrix} \Omega^{1,1} & (0) \\ (0) & \Omega^{2,2} \end{pmatrix}^{-1} X \right)^{-1} \times X' \times \begin{pmatrix} \Omega^{1,1} & (0) \\ (0) & \Omega^{2,2} \end{pmatrix}^{-1} \begin{pmatrix} Y^1 \\ Y^2 \end{pmatrix} \\ &= \begin{pmatrix} \left(X'(\Omega^{1,1})^{-1} X \right)^{-1} X' \times (\Omega^{1,1})^{-1} Y^1 \\ \left(X'(\Omega^{2,2})^{-1} X \right)^{-1} X' \times (\Omega^{2,2})^{-1} Y^2 \end{pmatrix} \\ &= \begin{pmatrix} \hat{b}_{mcg}^1 \\ \hat{b}_{mcg}^2 \end{pmatrix} \end{aligned}$$

Dans le cas présent, la régression \mathcal{R} s'écrit

$$\begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = (X_{t-1}^1 \cdots X_{t-p}^1 \mid X_{t-1}^2 \cdots X_{t-p}^2) \times \begin{pmatrix} \psi_1^{1,1} & \psi_1^{2,1} \\ \vdots & \vdots \\ \psi_p^{1,1} & \psi_p^{2,1} \\ \psi_1^{1,2} & \psi_1^{2,2} \\ \vdots & \vdots \\ \psi_p^{1,2} & \psi_p^{2,2} \end{pmatrix} + \begin{pmatrix} \eta_t^1 \\ \eta_t^2 \end{pmatrix}$$

tandis que les sous-régressions \mathcal{R}^1 et \mathcal{R}^2 s'écrivent respectivement

$$(\mathcal{R}^1): X_t^1 = (X_{t-1}^1 \cdots X_{t-p}^1 \mid X_{t-1}^2 \cdots X_{t-p}^2) \times \begin{pmatrix} \psi_1^{1,1} \\ \vdots \\ \psi_p^{1,1} \\ \psi_1^{1,2} \\ \vdots \\ \psi_p^{1,2} \end{pmatrix} + \eta_t^1$$

et

$$(\mathcal{R}^2): X_t^2 = (X_t^1 \mid X_{t-1}^1 \cdots X_{t-p}^1 \mid X_{t-1}^2 \cdots X_{t-p}^2) \times \begin{pmatrix} \alpha \\ \psi_1^{2,1} \\ \vdots \\ \psi_p^{2,1} \\ \psi_1^{2,2} \\ \vdots \\ \psi_p^{2,2} \end{pmatrix} + \eta_t^2$$

comme $\mathbb{V}(\eta_t) = \begin{pmatrix} \Sigma^{1,1} & 0 \\ 0 & \Sigma^{2,2} - \frac{(\Sigma^{1,2})^2}{\Sigma^{1,1}} \end{pmatrix}$ est diagonale, l'estimation des coefficients de \mathcal{R} est équivalente à celle résultant des deux sous-régressions \mathcal{R}^1 et \mathcal{R}^2 .

Q3 L'hypothèse \mathcal{H}_0 se réécrit " $\Sigma^{1,2} = 0$ " soit encore " $\alpha = 0$ ".

Remarquons en outre que si ϵ est gaussien, alors il en va de même pour $\eta = \begin{pmatrix} \epsilon^1 \\ \epsilon^2 - \frac{\text{Cov}(\epsilon^2, \epsilon^1)}{\text{V}(\epsilon^1)} \epsilon^1 \end{pmatrix}$.
tester \mathcal{H}_0 revient donc à tester la nullité individuelle du coefficient estimé $\hat{\alpha}$, ce qui peut être fait au moyen de la statistique de Student habituelle.

Remarquons que sous \mathcal{H}_0 seule la régression \mathcal{R}^2 est modifiée; il suffit donc de pratiquer le test dans la seule régression \mathcal{R}^2 .

Q4 L'hypothèse \mathcal{H}_0^* se réécrit " $\forall k \in \mathbb{N}, \phi_k^{1,2} = 0$ " soit encore " $\forall k \in \mathbb{N}, \psi_k^{1,2} = 0$ ".

Le test de \mathcal{H}_0^* revient alors au test de la nullité simultanée des coefficients estimés $(\widehat{\psi}_1^{1,2}, \dots, \widehat{\psi}_p^{1,2})$ qui peut être conduit au moyen de la statistique de Fisher.¹⁰

¹⁰ A savoir $F = \frac{SCR_{\mathcal{H}_0} - SCR_{-\mathcal{H}_0}}{SCR_{-\mathcal{H}_0}} \cdot \frac{T-2p}{p}$ où T désigne le nombre d'observations, $2p$ le nombre total de variable explicative et p le nombre de variables dont on teste la nullité simultanée, et bien-sûr $SCR_{\mathcal{H}_t}$ désigne la somme des carrés des résidus estimés (i.e. la variance non-expliquée) sous l'hypothèse \mathcal{H} . Si \mathcal{H}_0 est vraie alors F suit une loi de Fisher $\mathcal{F}(p, T-2p)$.

Remarquons que cette fois sous \mathcal{H}_0^* seule la régression \mathcal{R}^1 est modifiée, et tester \mathcal{H}_0^* revient à tester la régression \mathcal{R}^1 sous \mathcal{H}_0^* par rapport à \mathcal{R}^1 sous $\neg\mathcal{H}_0^*$. Ainsi il suffit de pratiquer le test dans la seule régression \mathcal{R}^1 .

*
* *