Correction de l'épreuve de Séries Temporelles Multivariées de mai 2019

Gilbert Colletaz

16 mai 2019

1 Exercices

Exercice 1

Dans tout ce qui suit on va travailler au seuil de risque de 5%.

1. Les deux premières tables donnent les résultats des tests τ, τ_μ et τ_t de Dickey-Fuller. Les taux de chômage étant toujours positif ou nuls, on ne considère pas la statistique τ qui implique une modélisation inadaptée sous H1. Sous cette condition, pour l'une ou l'autre des statistiques restantes, on ne peut pas rejeter la présence d'une racine unitaire dans chacune des deux séries.

   La première moitié du tableau 3 présente le résultat du test de cointégration de Johansen lorque la constante est à l'extérieur de l'espace de cointégration. Dans ce cas on ne rejette pas l'hypothèse d'un rang de cointégration nul, i.e. que les deux taux seraient I(1) non cointégrés.
   La seconde moitié refait ce test en mettant la constante dans le vecteur cointégrant. Avec cette restriction, on rejette l'hypothèse d'un rang de cointégration nul pour ne pas rejeter celle d'un rang égal à l'unité : ici les taux seraient toujours I(1) mais cointégrés.
   Au regard de la question posée, l'absence de cointégration de variable I(1) ne plaide pas en faveur de la convergence à long terme des économies. En revanche, leur cointégration témoignerait de l'existence d'une relation d'équilibre de long terme entre les taux de chômage des deux économies. Selon la place de la constante, on aboutit donc à des résultats contradictoires en ce qui concerne la convergence.

   Dans la table 4 on a le résultat du test de restriction sur la constante qui permet de rejeter à 5% l'hypothèse nulle d'une constante dans l'espace de cointégration. On préférerait donc retenir la présence d'un trend linéaire déterministe dans au moins une des deux variables ce qui peut paraître curieux dans une modélisation à long terme d'un taux de chômage. Surtout, cela conduirait à retenir d'après la table 3 une configuration défavorable à la cointégration et donc à la convergence à long terme.

2. Les commandes de la question 2 réclament l'estimation d'un VECM sur les deux taux de chômage avec donc une relation de cointégration, la constante étant mise dans l'espace de cointégration. Il s'agit donc d'une modélisation conforme aux résultats de la seconde moitié de la table 3, modélisation qui avait été rejetée par ceux de la table 4, et qui sont favorables à la thèse de la convergence.
   Avec l'option p=1 dans la commande `model` ..., on retient une écriture autorégressive en niveau avec un seul retard : il n'y a donc pas d'augmentation dans le VECM correspondant : la table 5 présente alors seulement l'estimation de la matrice π soit, pour un vecteur de variables explicatives mis dans l'ordre $(1, germany_{t-1}, france_{t-1})$, i.e. avec la constante en tête :
$$\pi = \begin{pmatrix} 0.17498 & 0.03068 & -0.03400 \\ 0.16645 & 0.02918 & -0.0324 \end{pmatrix}$$
   Comme précisé dans la commande `cointeg`..., $r = 1$, il existe donc deux matrices $\underset{(2\times1)}{\alpha}$ et $\underset{(3\times1)}{\beta}$ telles $\pi = \alpha\beta^\top$.
   D'où :
$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \begin{pmatrix} \beta_0 & \beta_g & \beta_f \end{pmatrix} = \pi = \begin{pmatrix} 0.17498 & 0.03068 & -0.03400 \\ 0.16645 & 0.02918 & -0.0324 \end{pmatrix}$$
   On sait aussi d'après la table 6 que $\alpha_1 = -0.03400$ et $\alpha_2 = -0.03234$. Par ailleurs, l'option `normalize=france` dans la commande `cointeg` force $\beta_f = 1$. Dans ces conditions, la première ligne du système précédent donne :
   . $-0.03400\beta_0 = 0.17498 \Rightarrow \beta_0 = -5.14647$
   . $-0.03400\beta_g = 0.03068 \Rightarrow \beta_g = -0.9023$

Pour vérifier les calculs il suffit de prendre la deuxième ligne du système :

. $-0.03234\beta_0 = 0.16645 \Rightarrow \beta_0 = -5.14647$
. $-0.03234\beta_g = 0.02918 \Rightarrow \beta_g = -0.9023$

La relation d'équilibre estimée aurait donc comme écriture : $-5.14647 - 0.9023 \times germany_t + france_t = 0$

3. (a) On aurait évidemment : $france_t = 5.14647 + 0.9023 \times germany_t + u_t$, où u_t est une aléatoire stationnaire centrée.
   (b) En égalisant les deux taux, il vient : $germany = 5.14647/(1. - 0.9023) = 52.67$! Ainsi, à moins de profondes réformes structurelles ou d'une crise d'ampleur inédite, la France ne rattrapera jamais le taux de chômage observé en Allemagne.
4. L'exogénéité ici renvoie à la question de savoir si une variable s'ajuste ou ne s'ajuste pas à un déséquilibre de long terme, c'est à dire si le loading factor α qui lui correspond dans l'écriture du VECM est nul ou pas. Comme le VECM ne met en jeu que des variables I(0), on peut employer le test usuel de significativité d'un coefficient qu'est le student. Ce test apparaît dans la table 6 : si on regarde l'équation afférente au taux allemand, on a une statistique de student égale à 5 et on rejette la nullité de α_1, ce taux de chômage allemand ne serait donc pas exogène faible et réagirait à un déséquilibre de long terme.
5. La chargée d'études modélise le couple formé par les deux taux de chômage par un VAR en différence première. On rappelle que cette modélisation correspond à la conclusion tirée du test de Johansen en l'absence de restriction sur la constante, absence de restriction qui n'est pas invalidée par la table 3. C'est donc la modélisation qui devrait être privilégiée au vu de ces deux tables 2 et 3 : on est en présence de variables I(1) non cointégrées.
   (a) Pour tester la causalité au sens de Granger, il suffit de tester successivement la nullité des coefficients du chômage allemand retardé dans l'équation du taux de chômage français et réciproquement. L'ordre du VAR étant égal à l'unité, seule la nullité d'un coefficient doit être considérée et on s'en sort donc par un student. Ces statistiques sont présentes dans la table 7.
      — H0 : le taux de chômage allemand n'est pas un prédicteur avancé du taux français : on a un student de 1.06 avec une $p-value$ de 0.2934 : aux seuils de risque usuels on ne rejette pas H0.
      — H0 : le taux de chômage français n'est pas un prédicteur avancé du taux allemand : on a un student de -0.33 avec une $p-value$ de 0.7401 : aux seuils de risque usuels on ne rejette pas H0.
   (b) Pour Geweke, on sait qu'il faut considérer les variances résiduelles issues des autorégressions jointes, du VAR complet et des autorégressions marginales. Les deux premières sont données dans la matrice de var-cov des innovations de la table 7. Les dernières sont tirées des deux dernières commandes exécutées et se trouvent dans les tables 8, pour le taux allemand, et 9, pour le taux français.
   On sait aussi que la partie inexpliquée d'une variable est plus faible dans l'autorégression marginale que dans l'autorégression jointe et donc la variance résiduelle de la première doit être inférieure ou égale à celle de la seconde. On peut observer que cela n'est pas vrai pour l'Allemagne (0.00195 dans la jointe, 000191 dans la marginale). La seule cause possible de ce résultat est que ce sont les estimateurs OLS des variances qui prennent en compte les degrés de liberté des sommes des carrés des résidus (soit 59-3 dans la table 7, puisqu'on perd une observation en raison du passage aux différences premières et que 3 coefficients sont estimés dans chaque régression jointe, et 59-2 dans les tables 8 et 9). Pour repasser aux estimateurs du maximum de vraisemblance qui évitent l'incohérence précédente, il suffit de faire , pour la table 6 :

$$\frac{59-3}{59} \times \begin{pmatrix} 0.00195 & 0.00011 \\ 0.00011 & 0.00803 \end{pmatrix} \rightarrow \begin{pmatrix} 0.00185 & 0.00010 \\ 0.00010 & 0.00762 \end{pmatrix}$$

   et respectivement pour les tables 8 et 9

$$\frac{59-3}{59} \times 0.00191 \rightarrow 0.00185$$
$$\frac{59-3}{59} \times 0.00805 \rightarrow 0.00778$$

   Avec ces chiffres, il vient :
   . $C_{germany \rightarrow france} = log(\frac{0.00778}{0.00762}) = 0.02$
   . $C_{france \rightarrow germany} = log(\frac{0.00185}{0.00185}) = 0.00$
   . $C_{france \leftrightarrow germany} = log\left(\frac{0.00185 \times 0.00762}{\begin{vmatrix} 0.00185 & 0.0010 \\ 0.00010 & 0.00762 \end{vmatrix}}\right) = 1.00$

. $C_{france,germany} = log\left(\frac{0.00185\times 0.00778}{\begin{vmatrix} 0.00185 & 0.0010 \\ 0.00010 & 0.00762 \end{vmatrix}}\right) = 1.02 = 0.02 + 0.00 + 1.00$

Pour les tests de significativité :

. sous H0 :le chômage allemand n'est pas un prédicteur avancé du français, $59*0.02 = 1.18$ serait la réalisation d'un $\chi^2(1) \rightarrow$ non rejet.

. sous H0 :le chômage français n'est pas un prédicteur avancé du allemand, $59*0.00 = 0.00$ serait la réalisation d'un $\chi^2(1) \rightarrow$ non rejet.

. sous H0 :pas de dépendance linéaire instantanée entre les deux taux, $59 * 1.00 = 59.00$ serait la réalisation d'un $\chi^2(1) \rightarrow$ rejet.

. sous H0 :pas de dépendance linéaire entre les deux taux, $59 * 1.02 = 60.18$ serait la réalisation d'un $\chi^2(3) \rightarrow$ rejet.

Au total il existe bien une dépendance linéaire entre les taux de chômage et celle-ci n'est qu'instantanée.

6. En lisant le code on reconnaît évidemment la procédure de test de cointégration en deux étapes d'Engle-Granger. La proc reg effectue la régression de la première étape et sauvegarde les résidus. La proc arima récupère ces résidus et réclame le calcul des statistiques de Dickey-Fuller sachant que τ_μ, dans ce contexte, devient la statistique d'Engle-Granger.
   La valeur critique est construite d'après la table de McKinnon (N=2, no trend) comme : $-3.3377 - \frac{5.967}{60} - \frac{8.98}{60^2} = -3.44$. Que l'on prenne la version DF ou ADF du test, on est conduit à ne pas rejeter la présence d'une racine unitaire dans les résidus et donc l'absence de cointégration des deux taux de chômage. En conséquence, ces résultats sont plutôt favorables à la modélisation VAR sur différences premières qu'à la modélisation de type VECM.

Exercice 2

Il s'agit de l'exercice 2 de l'épreuve de mai 2018. Pour sa correction, reportez-vous à celle de cette épreuve.