

Chapitre 9, exercice 6

Instructions pour étudier les séries du répertoire CH09EX06

Le répertoire CH09EX06 comporte un exercice de base destiné à tous les apprenants et un exercice avancé réservé aux seuls apprenants de la version avancée du cours.

Exercice de base (Pour tous les utilisateurs du cours)

Préalable Le chapitre 9 du cours doit avoir été suivi jusqu'à la page 156.



Objectif Le but est d'étudier les autocorrélations et les autocorrélations partielles de séries temporelles générées au moyen de processus autorégressifs d'ordre p ou $AR(p)$.



Données Les données sont artificielles. Elles ont été générées au moyen de plusieurs processus autorégressifs d'ordre p ou $AR(p)$. La série BLANC, présentée au chapitre 8, générée au moyen d'un processus bruit blanc a été employée pour toutes les simulations. Plusieurs jeux de paramètres sont utilisés. Les séries sont de longueur 400.



Structure de l'exercice

L'exercice comporte une partie :

- Dans la partie 1, le but est d'étudier, de manière empirique, les autocorrélations et les autocorrélations partielles de séries générées au moyen de plusieurs processus autorégressifs d'ordre p ou $AR(p)$. Non seulement plusieurs valeurs des paramètres sont employées mais plusieurs longueurs de séries sont considérées.

Partie 1 Un processus autorégressif d'ordre p ou $AR(p)$ est défini à partir d'un processus bruit blanc. Notons $\{e_t\}$ un processus bruit blanc de moyenne 0 et de variance σ^2 , égale à 1, par exemple. Notons $\{y_t\}$ le processus $AR(p)$. Il est défini par la relation $y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p} = e_t$, pour des valeurs déterminées des coefficients $\phi_1, \phi_2, \dots, \phi_p$. Pour générer des séries artificielles selon un processus $AR(p)$, on génère une série par le processus $\{e_t\}$ puis on calcule les y_t par la formule ci-dessus. Ceci nécessite p valeurs initiales. On peut par exemple, prendre $y_t = e_t$, pour $t = 1, \dots, p$.



Remarques


1. Dans ce cas, on recommande d'abandonner un certain nombre des premières données, 30 ou 100, par exemple, pour réduire l'effet du démarrage.
2. Ici, nous avons employé le moteur de calcul de Time Series Expert pour générer les séries. L'algorithme employé permet de ne pas abandonner de données au début parce qu'il n'y a pas d'effet de démarrage.

L'utilisation de Time Series Expert for Windows (TSE) a été déjà présentée lors de la réalisation des exercices 1 à 3. En cas de problème avec les instructions sommaires qui suivent, prière de revoir les instructions détaillées fournies lors de la réalisation de l'exercice 1.

1.1 EXAMEN DES AUTOCORRELATIONS ET AUTOCORRÉLATIONS PARTIELLES D'UNE SÉRIE ARTIFICIELLE

Nous commençons par une série générée par le processus $AR(p)$, défini par $y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p} = e_t$, avec $p = 2$, $\phi_1 = -0,5$, $\phi_2 = 0,3$, c'est-à-dire $y_t + 0,5 y_{t-1} - 0,3 y_{t-2} = e_t$. La série s'appelle AR2_503.


- ⇒ Choisissez le répertoire de données approprié sur votre disque (pas sur le CD-ROM): menu File ⇒ Open. Choisissez DATA puis CHAP09 puis CH09EX06.
- ⇒ Chargez le problème déjà préparé : ARTIF. Vous devez alors voir dans le bas de l'écran que la variable dépendante est BLANC, que l'échantillon d'estimation est 1 – 400 et que les prévisions seront calculées jusqu'en 400.
- ⇒ Pour visualisez la série AR2_503 : menu Graphics ⇒ Series. Choisissez AR2_503. Cliquez OK.

- 
- ⇒ Pour visualiser les autocorrélations de la série: menu Graphics
⇒ Autocorrelations and partials. Choisissez AR2_503. Cliquez OK.
 - ⇒ Pour visualiser les autocorrélations partielles de la série, pressez Enter.

? Vérifiez quels sont les retards pour lesquels les autocorrélations et les autocorrélations partielles sont significatives au seuil de 5%.

1.1.1 Votre réponse

Nous allons maintenant réduire la longueur de la série, comme dans l'exercice 3. Par exemple :

- 
- ⇒ Pour visualiser les autocorrélations des 100 dernières données de la série AR2_503: menu Graphics ⇒ Autocorrelations and partials. Choisissez AR2_503. Ensuite sur la ligne First date, tapez 301. Cliquez OK.
 - ⇒ Pour visualiser les autocorrélations partielles de la série, pressez Enter.
 - ⇒ Expérimentez avec différentes longueurs de séries extraites de la série AR2_503.

? Comment pourrait-on résumer les résultats obtenus?

1.1.2 Votre réponse

1.2 EXAMEN DES AUTOCORRELATIONS ET AUTOCORRÉLATIONS PARTIELLES D'AUTRES SÉRIES ARTIFICIELLES

Nous traitons maintenant des séries générées à partir d'autres processus autorégressifs. Voici les noms des séries et les équations qui ont servi à les

générer.

$$\text{AR21_6} \quad y_t + y_{t-1} + 0,6y_{t-2} = e_t$$

$$\text{AR202_9} \quad y_t - 0,2y_{t-1} + 0,9y_{t-2} = e_t$$

$$\text{AR409} \quad y_t - 0,9y_{t-4} = e_t$$

$$\text{AR5_6_8} \quad y_t - 0,6y_{t-1} - 0,8y_{t-4} + 0,48y_{t-5} = e_t$$

Pour chacune de ces quatre séries :

- ⇒ Visualisez la série: menu Graphics ⇒ Series. Choisissez son nom. Cliquez OK.
- ⇒ Visualisez les autocorrélations de la série: menu Graphics ⇒ Autocorrelations and partials. Choisissez son nom. Cliquez OK.
- ⇒ Visualisez les autocorrélations d'extraits de la série, en procédant comme indiqué au paragraphe 1.2.
- ⇒ Pour visualiser les autocorrélations partielles de la série, pressez Enter.

Dans les conclusions relatives aux autocorrélations partielles d'un processus $\text{AR}(p)$, on dit dans le cours :

- Les séries générées par un processus $\text{AR}(p)$ présentent de l'autocorrélation partielle de retard 1 à p , tandis que les autres $p_p + 1, p_p + 2, \dots, p_p + 3, \dots$ ne sont pas (ou sont peu) significatives (c'est-à-dire sont dans la bande ou presque).
- La longueur T de la série et la valeur des coefficients $\theta_1, \theta_2, \dots$ ont un effet sur les résultats.
- On dit qu'il y a *truncation* des autocorrélations partielles *au delà* du retard p .



Essayez de justifier ces conclusions.



1.2.1 Votre réponse

On dit aussi :

- On peut en effet montrer (voir des indications dans la cadre du *cours avancé*) que les autocorrélations partielles de retard $k > p$ d'un processus AR(p) sont égales à 0.

Ceci est étudié dans le cadre du cours avancé, à titre facultatif, dans la partie A du présent exercice.

SYNTHESE

Dans la partie 1 de l'exercice 5, nous avons introduit les autocorrélations *partielles* et nous avons montré leur intérêt pour caractériser les processus autorégressifs d'ordre 1. Dans le présent exercice, nous avons considéré des séries générées par des processus autorégressifs d'ordre p supérieur à 1. Nous avons montré que les autocorrélations partielles ont un comportement assez tranché en ce sens que, pour les retard $k > p$, elles sont pour la plupart statistiquement non significatives. On dit donc que les autocorrélations partielles d'un processus autorégressif d'ordre p sont tronquées au delà de p . Les autocorrélations partielles permettent donc de caractériser les processus autorégressifs. Nous avons déjà vu dans l'exercice 5, partie 2, que les autocorrélations partielles n'ont pas d'intérêt pour des processus moyenne mobile et ne permettent pas de les caractériser.

En conséquence, chaque fois que les autocorrélations présenteront une allure tronquée au delà d'un certain retard q , il faudra penser à un processus moyenne mobile d'ordre q . En revanche, quand ce seront les autocorrélations partielles qui présenteront une allure tronquée au delà d'un certain retard p , il faudra penser à un processus autorégressif d'ordre q .



Exercice avancé

(Pour les utilisateurs de la version avancée du cours)

Préalable



Le chapitre 9 du cours de base et avancé doit avoir été suivi jusqu'à la page 168.

Objectif



Le but est de développer la théorie relative aux autocorrélations d'un processus $AR(p)$ de manière légèrement plus ample que dans le cours, *à titre facultatif*.

Données

Néant.

Structure de l'exercice

L'exercice comporte une partie :

- Dans la partie A, le but est de procéder, *à titre facultatif*, à une démonstration mathématique qui explique le résultat des expériences vues dans la partie 1. La démonstration repose sur des concepts relativement simples. Il est conseillé d'attendre la fin de la partie du cours consacrée aux processus autorégressifs avant d'examiner cette partie de l'exercice.

Partie A Dans la partie 1, nous avons considéré les autocorrélations de processus $AR(p)$ de manière purement empirique, en nous basant sur des graphiques obtenus pour quelques séries artificielles. Les expériences ont permis de faire varier la valeur des coefficients et la longueur de la série. Ici, nous procédons à une démonstration mathématique qui explique le résultat des expériences. La démonstration repose sur les mêmes concepts que dans l'exercice 3.

A.a LA CONDITION DE STATIONNARITÉ

La condition de stationnarité, qui était très simple dans le cas du processus $AR(1)$, est bien plus compliquée dans le cas général. Considérons d'abord un processus $AR(2)$

$$y_t = -0,5y_{t-1} + 0,3y_{t-2} + e_t$$

qui peut s'écrire

$$(1 + 0,5B - 0,3B^2)y_t = e_t.$$

Le trinôme du second degré $1 + 0,5x - 0,3x^2$ a un réalisant $(0,5)^2 - 4(-0,3) = 1,45 = (1,204)^2$, dont les racines $(-0,5 \pm 1,204)/(-0,6)$ valent 2,84 et $-1,17$. Notons que le produit des racines vaut $-1/0,3$, l'inverse du coefficient de x^2 . Par conséquent, on a la factorisation :

$$\begin{aligned} 1 + 0,5x - 0,3x^2 &= -0,3[x - 2,84][x - (-1,17)] \\ &= (-0,3) \times (2,84) \times (-1,17) \times (0,352x - 1)(-0,852x - 1) \\ &= (1 - 0,352x)(1 + 0,852x). \end{aligned}$$

En décomposant de même en facteurs l'opérateur autorégressif, le processus peut être considéré comme défini par l'équation

$$(1 + 0,352B)(1 + 0,852B)y_t = e_t.$$

On peut vérifier que $(1 - 0,352B)(1 + 0,852B) = 1 + 0,5B - 0,3B^2$. Le fait que les coefficients de B dans la factorisation soient compris entre -1 et 1 garantit la stationnarité, ce qui correspond à des racines supérieures à 1 en valeur absolue.

Plus généralement, si les x_i racines du polynôme $1 - \phi_1x - \dots - \phi_px^p$ sont telles que $|x_i| > 1$, le processus est stationnaire.

Un polynôme de degré $p - p$ supérieur à $1 - n$ a pas nécessairement p racines réelles. Par exemple, le polynôme $1 + x + 0,6x^2$ a un réalisant négatif,

égal à $1 - 4(0,6) = -1,4$. Le polynôme n'est donc pas factorisable. Dans ce cas, il faut également prendre en considération les *racines complexes* qui vont toujours par paire de nombres *imaginaires conjugués*. Par exemple, en notant i l'unité imaginaire égale à $\sqrt{-1}$, les racines de $1 + x + 0,6x^2$ sont de la forme

$$\frac{-1 \pm \sqrt{-1,4}}{2 \times 0,6} = \frac{-1 \pm 1,183i}{1,2} = -0,833 \pm 0,986i,$$

soit des nombres de *module* $\sqrt{(-0,833)^2 + (\pm 0,986)^2} = 1,291$ (en fait, c'est simplement la racine carrée de l'inverse du terme indépendant du trinôme 0,6). Ce module est supérieur à 1. Le processus AR(2) de polynôme autorégressif $1 + B + 0,6B^2$ est donc stationnaire.

Remarque



Notons que la factorisation complexe fonctionne effectivement parce que

$$\begin{aligned} & 0,6[x - (-0,833 + 0,986i)][x - (-0,833 - 0,986i)] \\ &= 0,6[x^2 - \{(-0,833 + 0,986i) + (-0,833 - 0,986i)\}x + (-0,833 + 0,986i)(-0,833 - 0,986i)] \\ &= 0,6[x^2 - 2(-0,833)x + (-0,833)(-0,833) + (0,986)(-0,986)i^2] \\ &= 0,6[x^2 + 1,667x + (0,833^2 + 0,986^2)] \\ &= 0,6[x^2 + 1,667x + (1,291)^2] \\ &= 0,6x^2 + x + 1 \end{aligned}$$

En revanche, le processus AR(2) qui aurait comme polynôme autorégressif $1 + B + 1,2B^2$ n'est pas stationnaire parce que les racines du polynôme correspondant sont de carré de module $1/(1,2)$, inférieur à 1.

Remarque



Notons que les processus autorégressifs sont toujours inversibles. Il n'y a donc pas de condition d'inversibilité.

A.b UTILISATION DES AUTOCORRÉLATIONS

Les autocorrélations de ces séries ne permettent pas d'identifier les processus AR(2) générateurs. Tout au plus peut-on relever une décroissance exponentielle ou des fluctuations pseudo-périodiques amorties. Ces dernières sont particulièrement visibles dans le cas de la série AR202_9 et sont dues au fait que les racines du polynôme autorégressif du processus autorégressif générateur sont complexes. Remarquons que le graphique des séries ne caractérise pas très bien le processus générateur. Il y a des exceptions, comme la série générée par le processus autorégressif avec $\phi = 0,9$ qui

présente des excursions amples autour de la moyenne 0. Dorénavant, nous ne demanderons plus systématiquement de visualiser le graphique des séries.

Nous avons mentionné que la *condition de stationnarité* d'un processus $AR(p)$ utilise les modules des racines du polynôme autorégressif qui doivent être supérieurs à 1. Cette condition n'est pas toujours vérifiable aisément. Toutefois, le processus $AR(4)$ défini par l'équation

$$y_t - 0,9y_{t-4} = e_t$$

est stationnaire parce que les racines de l'équation $1 - 0,9x^4 = 0$ sont de module $(1/0,9)^{1/4} > 1$. De même, le processus $AR(5)$ d'équation

$$(1 + 0,6B)(1 + 0,8B^4)y_t = e_t$$

est également stationnaire parce que les modules des racines valent $1/0,6$ et $(1/0,8)^{1/4}$.

Il est intéressant de remarquer qu'une racine réelle dans le polynôme autorégressif induit une contribution exponentielle décroissante dans la fonction d'autocorrélation, tandis qu'un couple de racines complexes, imaginaires conjuguées, produit une *cosinusoïde amortie*. Plus précisément, on peut montrer que la fonction d'autocorrélation d'un processus $AR(p)$ est une combinaison linéaire de fonctions exponentielles décroissantes avec ou sans oscillations ainsi que de cosinusoïdes amorties. Une série temporelle générée par un processus $AR(p)$ présentera des fluctuations de comportement plus ou moins fortement périodiques s'il y a des racines complexes dans le polynôme autorégressif du processus générateur.

A.c UTILISATION DES AUTOCORRÉLATIONS PARTIELLES

Il s'ensuit qu'il n'est pas facile d'identifier un processus autorégressif par sa fonction d'autocorrélation, sauf dans le cas du processus $AR(1)$. C'est la raison pour laquelle on utilise les *autocorrélations partielles*. Nous ne pouvons pas établir ici la caractérisation évoquée dans la partie 1 parce qu'elle est trop complexe dans le cadre de ce cours.

A.d CALCUL DES PRÉVISIONS POUR UN MODÈLE $AR(p)$

Plaçons-nous au temps T où la dernière donnée est disponible. Sachant que le modèle est $AR(p)$, d'équation

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + e_t, \quad (*)$$

il faut calculer les prévisions $\hat{y}_T(1), \hat{y}_T(2), \dots, \hat{y}_T(h)$. On écrit l'équation (*) aux temps $T+1, T+2, \dots, T+h$. Pour fixer les idées, supposons $p=2$ et $h=3$. On obtient donc

$$\begin{aligned} y_{T+1} &= \phi_1 y_T + \phi_2 y_{T-1} + e_{T+1} \\ y_{T+2} &= \phi_1 y_{T+1} + \phi_2 y_T + e_{T+2} \\ y_{T+3} &= \phi_1 y_{T+2} + \phi_2 y_{T+1} + e_{T+3}. \end{aligned}$$

Pour obtenir une prévision ponctuelle, on remplace les innovations par leur moyenne qui est zéro, ce qui donne

$$\begin{aligned} \hat{y}_T(1) &= \phi_1 y_T + \phi_2 y_{T-1} \\ \hat{y}_T(2) &= \phi_1 \hat{y}_T(1) + \phi_2 y_T \\ \hat{y}_T(3) &= \phi_1 \hat{y}_T(2) + \phi_2 \hat{y}_T(1), \end{aligned}$$

où il faut encore substituer les valeurs futures inconnues figurant dans les seconds membres par leurs prévisions ponctuelles les plus récentes, donc en $t=T$:

$$\begin{aligned} \hat{y}_T(1) &= \phi_1 y_T + \phi_2 y_{T-1} \\ \hat{y}_T(2) &= \phi_1 \hat{y}_T(1) + \phi_2 y_T \\ \hat{y}_T(3) &= \phi_1 \hat{y}_T(2) + \phi_2 \hat{y}_T(1). \end{aligned}$$

On constate donc que, pour un horizon h supérieur à l'ordre p du modèle, les prévisions d'un modèle AR(p) ne sont pas nulles (contrairement à celles du modèle moyenne mobile) mais ne font plus intervenir directement les valeurs observées. Les prévisions vont se situer sur une courbe, d'une allure bien déterminée, dont la forme ne dépend que des coefficients $\phi_j, j=1, \dots, p$. La détermination de la distribution des valeurs futures et des intervalles de prévision est reportée au chapitre suivant.

A.e CONCLUSIONS

Si la fonction d'autocorrélation partielle d'une série est calculée et si elle paraît tronquée, compte tenu des fluctuations statistiques admises, on peut envisager de modéliser la série par un processus autorégressif. Sauf dans le cas du processus AR(1), la fonction d'autocorrélation est de peu d'utilité pour trouver la spécification du modèle. Tout au plus peut-on dire

(a) qu'une composante exponentielle décroissante des autocorrélations, même mélangée à d'autres fluctuations, fait songer à une racine (réelle) positive du polynôme autorégressif,

- (b) qu'une alternance en signes indique une racine (réelle) négative, et
- (c) qu'un comportement pseudo-périodique plus ou moins amorti des autocorrélations (qui correspond à un comportement pseudo-périodique au sein de la série), est le signe de la présence de racines complexes dans le polynôme autorégressif.

Si les résidus e_t d'un modèle, quel qu'il soit, présentent une fonction d'autocorrélation partielle tronquée, ils sont également susceptibles d'être représentés par un modèle autorégressif. Par exemple, on peut avoir $e_t = \phi e_{t-1} + u_t$, où $\{u_t\}$ est un bruit blanc. Il faut noter que ce modèle pour les résidus est celui qui est le plus souvent considéré comme alternative au bruit blanc (méthodes de Cochrane-Orcutt et Hildreth-Lu en économétrie, par exemple). Nous verrons ultérieurement comment combiner le modèle des erreurs avec le modèle déjà utilisé pour produire les résidus, ce qui devrait permettre de conduire à un meilleur modèle.

SYNTHESE

Contrairement aux résultats établis dans les parties avancées des autres exercices, nous n'avons pas montré la caractérisation des séries produites par un processus $AR(p)$: les autocorrélations partielles sont tronquées au delà du retard p , ce qui veut dire que celles de retard 1 à p sont possiblement non nulles, mais les suivantes ne s'écartent pas de 0 de façon significative. C'est en effet trop complexe dans le cadre de ce cours. En revanche, nous avons discuté du problème de la stationnarité qui repose sur les racines de ce qu'on appelle le polynôme autorégressif. Ces notions algébriques sont en effet indispensable pour une compréhension de la modélisation des séries temporelles.

[Retour au chapitre 9](#)