

Chapitre 9, exercice 3

Instructions pour étudier les séries du répertoire CH09EX03

Le répertoire CH09EX03 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 94.



Objectif Le but est d'étudier les autocorrélations de séries temporelles générées au moyen de processus moyenne mobile d'ordre 1 ou MA(1).



Données Les données sont artificielles. Elles ont été générées au moyen de plusieurs processus moyenne mobile d'ordre 1 ou MA(1). La série BLANC, présentée au chapitre 8, exercice 4, 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 de séries générées au moyen de plusieurs processus moyenne mobile d'ordre 1 ou MA(1). Non seulement plusieurs valeurs du paramètre sont employées mais plusieurs longueurs de séries sont considérées.

Partie 1 Un processus moyenne mobile d'ordre 1 ou MA(1) 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 MA(1). Il est défini par la relation $y_t = e_t - \theta e_{t-1}$. Pour générer des séries artificielles selon un processus MA(1), on génère une série par le processus $\{e_t\}$ puis on calcule les y_t par la relation ci-dessus. On ne peut pas appliquer cette formule pour $t = 1$. Aussi pose-t-on $y_1 = e_1$.

1.1 AUTOCORRÉLATIONS DE LA SERIE BLANC

Nous allons effectuer cette étude en employant Time Series Expert for Windows.

Son utilisation a été déjà présentée lors de la réalisation des exercices 1 et 2 du chapitre 9. 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 du chapitre 9.

- ⇒ Pour lancer le logiciel, suivez les instructions données en annexe de l'introduction du cours.
- ⇒ Choisissez le répertoire de données approprié sur votre disque (pas sur le CD-ROM): menu File ⇒ Open. Choisissez DATA puis CHAP09 puis CH09EX03.
- ⇒ 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 visualiser graphiquement la série BLANC: menu Graphics ⇒ Series. Sélectionnez BLANC.
- ⇒ Pour visualiser les autocorrélations de la série: menu Graphics ⇒ Autocorrelations and partials. Sélectionnez BLANC. Cliquez OK.



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



1.1.1 Votre réponse

1.2 EXAMEN DES AUTOCORRELATIONS D'UNE SÉRIE ARTIFICIELLE

Nous commençons par une série générée par le processus MA(1), définie par $y_t = e_t - \theta e_{t-1}$, avec $\theta = 0,5$, c'est-à-dire $y_t = e_t - 0,5e_{t-1}$. La série s'appelle MA105.

- ⇒ Pour visualiser la série MA105: menu Graphics ⇒ Series. Choisissez MA105. Cliquez OK.
- ⇒ Pour visualiser les autocorrélations de la série: menu Graphics ⇒ Autocorrelations and partials. Choisissez MA105. Cliquez OK.



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



1.2.1 Votre réponse

Nous allons maintenant réduire la longueur de la série. La manière la plus simple est de changer, dans la boîte de dialogue Autocorrelation parameters, soit la date de début sur la ligne First date, soit la date de fin sur la ligne Last date, de manière à avoir le nombre d'observations souhaité. Par exemple :

- ⇒ Pour visualiser les autocorrélations des 100 dernières données de la série MA105 : menu Graphics ⇒ Autocorrelations and partials. Choisissez MA105. Ensuite sur la ligne First date, tapez 301, et cliquez OK.



Remarques

1. Pour changer la date de fin, déplacez le curseur sur la ligne Last date, tapez la date de fin, par exemple 300.

2. La date de début doit être antérieure à la date de fin.

3. Les séries artificielles traitées ici sont non datées. Quand on traite des données annuelles, mensuelles ou trimestrielles, les dates introduites doivent être valides, c'est-à-dire notamment

- des nombres sans décimales, pour des données annuelles,
- des nombres avec une décimale, la partie décimale variant de 1 à 4, pour des données trimestrielles,
- des nombres avec deux décimales, la partie décimale variant de 1 à 12, pour des données annuelles.

⇒ Expérimentez avec différentes longueurs de séries extraites de la série MA105.



Comment pourrait-on résumer les résultats obtenus?



1.2.2 Votre réponse

1.3 EXAMEN DES AUTOCORRELATIONS D'AUTRES SÉRIES ARTIFICIELLES

Nous traitons maintenant des séries générées à partir de trois autres processus MA(1), définis par $y_t = e_t - \theta e_{t-1}$, mais avec respectivement $\theta = -0,5$, $\theta = 0,2$ et $\theta = 0,9$. Les séries s'appellent respectivement MA1_5, MA102 et MA109. Pour chacune de ces 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.

⇒ Visualisez les autocorrélations d'extraits de la série, en procédant comme indiqué au paragraphe 1.2.

Dans les conclusions relatives aux autocorrélations d'un processus MA(1), on dit dans le cours :

- Les séries générées par un processus MA(1) présentent de l'autocorrélation de retard 1, principalement: $r_1 \neq 0$, tandis que les autres r_2, r_3, r_4, \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 du coefficient θ ont un effet sur les résultats.
- On dit qu'il y a *truncation* des autocorrélations *au-delà* du retard 1.
- Il est intéressant de voir dans les graphiques que ce sont souvent pour les mêmes retards que les autocorrélations sont légèrement significatives.
- C'est parce que toutes les séries ont été générées à partir de la même réalisation d'un bruit blanc.



Essayez de justifier ces conclusions.



1.3.1 Votre réponse

On dit aussi :

- On peut en effet montrer (ce sera fait plus loin dans la cadre du **cours avancé**) que les autocorrélations de retard $k > 1$ d'un processus MA(1) sont égales à 0. (Voir partie A du présent exercice, à titre facultatif).

SYNTHESE

Dans le cours, nous avons déjà examiné des séries générées par des processus moyenne mobile d'ordre 1. Le but du présent exercice était de confirmer, par des expériences basées sur des séries simulées, le comportement des autocorrélations. On constate effectivement que l'autocorrélation de retard 1 est en effet statistiquement significative, en général (sauf peut-être dans le cas où le coefficient θ vaut 0,2) mais pas pour la plupart des

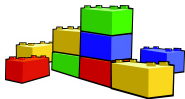
autocorrélations de retard 2, 3, Il en résulte une caractérisation des séries produites par un processus MA(1): les autocorrélations sont tronquées au delà du retard 1, ce qui veut dire que celle de retard 1 est non nulle, les suivantes ne s'écartent pas de 0 de façon significative.



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 123.

Objectif



Le but est de développer, *à titre facultatif*, la théorie relative aux autocorrélations d'un processus MA(1).

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 moyenne mobile avant d'examiner cette partie de l'exercice.

Partie A Dans la partie 1, nous avons considéré les autocorrélations de processus MA(1) de manière purement empirique, en nous basant sur des graphiques obtenus pour quelques séries artificielles. Les expériences ont permis de faire varier la valeur du paramètre et la longueur de la série. Ici, nous procédons à une démonstration mathématique qui explique le résultat des expériences. La démonstration repose sur des concepts relativement simples : évaluation d'une espérance mathématique et d'une variance d'une variable aléatoire qui est la somme, ou plus généralement une combinaison linéaire, de deux variables aléatoires, ainsi que l'évaluation d'une covariance entre deux sommes, ou plus généralement deux combinaisons linéaires, de variables aléatoires.

A.a PRÉALABLE

Il n'est pas question de traiter ici des processus aléatoires de façon approfondie mais seulement d'introduire les éléments nécessaires à la prévision par les modèles ARIMA. Il existe des ouvrages qui abordent le sujet de façon plus complète, notamment Gouriéroux et Monfort [1990].

Commençons par des propriétés de la moyenne aussi appelée espérance mathématique, de la variance et de la covariance de variables aléatoires. Nous traiterons la variance comme cas particulier d'une covariance. En effet la variance de X est la covariance de X avec elle-même : $\text{cov}(X, X) = \text{var}(X)$.

1) La moyenne

Notons $E(X)$ la moyenne ou espérance mathématique de la variable aléatoire X . Pour toutes variables aléatoires X_1 et X_2 et pour toutes constantes a_1 et a_2 , on a

$$E(a_1X_1 + a_2X_2) = a_1 E(X_1) + a_2 E(X_2).$$

Nous emploierons également le théorème suivant.

2) La covariance

Pour toutes variables aléatoires X_1, X_2, Y_1, Y_2 de variances finies et pour toutes constantes a_1, a_2, b_1, b_2 :

$$\begin{aligned} \text{cov}(a_1X_1 + a_2X_2, b_1Y_1 + b_2Y_2) &= a_1b_1 \text{cov}(X_1, Y_1) + a_1b_2 \text{cov}(X_1, Y_2) \\ &\quad + a_2b_1 \text{cov}(X_2, Y_1) + a_2b_2 \text{cov}(X_2, Y_2). \end{aligned}$$

3) La variance

Sous les mêmes suppositions

$$\text{var}(a_1 X_1 + a_2 X_2) = a_1^2 \text{var}(X_1) + a_2^2 \text{var}(X_2) + 2a_1 a_2 \text{cov}(X_1, X_2).$$

Par exemple, on a $\text{var}(X_1 - X_2) = \text{var}(X_1) + \text{var}(X_2) - 2\text{cov}(X_1, X_2)$

et $\text{var}(X_1 + X_2) = \text{var}(X_1) + \text{var}(X_2) + 2\text{cov}(X_1, X_2)$.

En particulier, si les variables X_1 et X_2 sont non corrélées, $\text{cov}(X_1, X_2) = 0$ et donc $\text{var}(X_1 + X_2) = \text{var}(X_1) + \text{var}(X_2) = \text{var}(X_1 - X_2)$. Ceci arrive en particulier si les deux variables X_1 et X_2 sont indépendantes.



Remarque

Pour la démonstration, utiliser le fait que les propriétés relatives aux variances et aux covariances ne sont pas influencées par un changement d'origine pour des variables, par exemple $\text{var}(X) = \text{var}(X + c)$, où c est une constante; en conséquence, on peut supposer sans restreindre la généralité que les variables aléatoires sont toutes de moyenne nulle, de sorte que $\text{cov}(X_1, X_2) = E(X_1 X_2)$; la propriété résulte alors de la distributivité de la multiplication par rapport à l'addition et de ce que $E(cX_1) = cE(X_1)$.

A.b LE PROCESSUS "MOYENNE MOBILE" D'ORDRE 1 OU MA(1)

La variable au temps t est une combinaison linéaire de deux innovations successives

$$y_t = e_t - \theta e_{t-1}.$$

Remarque



Le coefficient de e_t est 1, parce que ce terme représentera l'innovation en t , qui sera aussi l'erreur de prévision d'horizon 1. On verra que le coefficient θ ne peut pas prendre n'importe quelle valeur.

Des processus utilisant plus d'une erreur passée peuvent également être introduits. Ils seront appelés des *processus moyenne mobile* ("moving average") d'ordre q où q est le plus grand retard utilisé. Il ne faut pas les confondre avec les moyennes mobiles utilisées pour le lissage de séries temporelles.

Nous traitons de manière détaillée seulement du processus moyenne mobile d'ordre 1 ou MA(1).

Calculons l'espérance mathématique et la variance de y_t , ainsi que les covariances entre y_t et y_{t-1} , y_{t-2} ,...

1) Moyenne de la variable

Par application du paragraphe précédent, on a :

$$E(y_t) = E[e_t - \theta e_{t-1}] = 0 - \theta \cdot 0 = 0.$$

2) Variance de la variable

$$\begin{aligned} \text{var}(y_t) &= \text{var}(e_t - \theta e_{t-1}) \\ &= \text{var}(e_t) + (-\theta)^2 \text{var}(e_{t-1}) + 2(-\theta) \text{cov}(e_t, e_{t-1}) \\ &= \sigma^2 + \theta^2 \sigma^2 - 2\theta \cdot 0 = (1 + \theta^2) \sigma^2 \end{aligned}$$

3) Autocovariances du processus

$$\begin{aligned} \text{cov}(y_t, y_{t-1}) &= \text{cov}(e_t - \theta e_{t-1}, e_{t-1} - \theta e_{t-2}) \\ &= \text{cov}(e_t, e_{t-1}) - \theta \text{cov}(e_t, e_{t-2}) - \theta \text{cov}(e_{t-1}, e_{t-1}) + \theta^2 \text{cov}(e_{t-1}, e_{t-2}) \\ &= -\theta \text{var}(e_{t-1}) = -\theta \sigma^2 \\ \text{cov}(y_t, y_{t-2}) &= \text{cov}(e_t - \theta e_{t-1}, e_{t-2} - \theta e_{t-3}) = \dots = 0. \end{aligned}$$

4) Autocorrélations du processus

On en déduit :

$$\gamma_0 = \text{var}(y_t) = (1 + \theta^2) \sigma^2$$

et

$$\gamma_1 = \text{cov}(y_t, y_{t-1}) = -\theta \sigma^2 \quad \text{donc} \quad \rho_1 = \gamma_1 / \gamma_0 = -\frac{\theta}{(1 + \theta^2)}$$

$$\gamma_2 = 0 \quad \text{donc} \quad \rho_2 = 0$$

$$\gamma_3 = 0 \quad \text{donc} \quad \rho_3 = 0, \text{ etc.}$$

On remarque que seule l'autocorrélation de retard 1 est différente de 0.

On dit que la fonction d'autocorrélation est *tronquée* au-delà du retard 1.

Pour rendre la notion de processus MA(1) plus compréhensible, nous avons généré dans la partie 1 des séries temporelles artificielles à l'aide de ce processus. Nous avons repris la série $\{e_t\}$ générée selon un processus bruit blanc de l'exercice 4 du chapitre 8. Nous avons utilisé cette série pour déterminer les y_t , à l'aide de l'équation $y_t = e_t - \theta e_{t-1}$, en laissant varier le coefficient θ . Nous avons ainsi généré des séries artificielles de longueur 400 selon plusieurs processus MA(1) pour les valeurs suivantes

du coefficient $\theta = -0,5$, $\theta = 0,2$, $\theta = 0,5$ et $\theta = 0,9$.

A.c NOTION D'INVERSIBILITÉ

Il est également intéressant de considérer le problème inverse : connaissant les autocorrélations du processus, retrouver les coefficients. Dans le cas du processus MA(1), cela revient à résoudre l'équation en θ

$$\rho_1 = -\frac{\theta}{(1+\theta^2)}.$$

Par exemple, supposons pour fixer les idées que $\rho_1 = -0,4$. L'équation du second degré

$$0,4\theta^2 - \theta + 0,4$$

a deux solutions 0,5 et 2. En réalité une seule des deux solutions est acceptable. En effet, si l'on écrit le modèle sous la forme traditionnelle $y_t = f(X_{1t}, X_{2t}, \dots) + e_t$, avec ici les valeurs retardées de y comme variables explicatives, on obtient :

$$\begin{aligned} y_t &= -\theta e_{t-1} + e_t \\ &= -\theta(y_{t-1} + \theta e_{t-2}) + e_t \\ &= -\theta y_{t-1} - \theta^2(y_{t-2} + \theta e_{t-3}) + e_t = \dots \\ &= -\theta y_{t-1} - \theta^2 y_{t-2} - \theta^3 y_{t-3} - \dots + e_t. \end{aligned}$$

Si $\theta > 1$ ou $\theta < -1$, le poids du passé va en grandissant, ce qui est absurde.

Les seules valeurs acceptables de θ sont donc comprises entre -1 et 1 (notons que les valeurs extrêmes -1 et 1 sont ici admissibles). C'est ce qu'on appelle la *condition d'inversibilité*. En imposant cette condition, ce que nous ferons désormais, il n'y a qu'une seule valeur de θ possible pour un processus MA(1) dont l'autocorrélation de retard 1 est donnée. Dans l'exemple traité, $\rho_1 = -0,4$, on trouve $\theta = 0,5$.

A.d SPÉCIFICATION ET ESTIMATION D'UN MODÈLE MA(1)

Supposons que le calcul des autocorrélations r_k d'une série temporelle amène la constatation qu'elles ne sont pas significatives au-delà du retard 1, c'est-à-dire que r_1 est significative mais pas r_2 , r_3 , etc. On peut alors raisonnablement supposer que la série provient d'un processus MA(1). On retient donc pour la série un modèle MA(1) et non un bruit blanc. On dit qu'on a *identifié* ou *spécifié* un modèle MA(1). La question qui se pose alors est de déterminer le coefficient θ_1 , ou plus simplement θ , du modèle ainsi que l'écart-type σ des innovations $\{e_t\}$. C'est un problème d'estima-

tion statistique qui n'est pas simple. La méthode la plus immédiate consiste à évaluer r_1 et ρ_1 . On obtient alors une équation, similaire à une équation écrite plus haut :

$$-\frac{\theta}{(1+\theta^2)} = r_1.$$

Par exemple, si $r_1 = -0,4$, on trouve comme estimation $\hat{\theta} = 0,5$.

SYNTHESE

Nous avons montré, de manière plus générale que dans les exemples de la partie 1, une caractérisation des séries produites par un processus MA(1): les autocorrélations sont tronquées au delà du retard 1, ce qui veut dire que celle de retard 1 est non nulle, les suivantes ne s'écartent pas de 0 de façon significative.

[Retour au chapitre 9](#)