Cours des Séries chronologiques

Auto-corrélation et Autocorrélation partielle

SACI Oualid

Ecole supérieure en sciences et technologies de l'informatique et du numérique

Introduction

La connaissance d'un processus stationnaire se ramène entièrement à l'étude de la fonction d'auto covariance $\gamma(h) = cov(X_t, X_{t+h})$, ou bien à la connaissance de la variance $Var(X_t)$ et à la fonction d'auto corrélation $\rho(h) = \frac{\gamma(h)}{\gamma(0)}$.
Notons que ces notions n'ont de sens que pour les processus stationnaires.
Les pcopriétés sont semblables pour γ et ρ (voir le cours 1).

Notion de corrélation partielle

Un marchand de glaces, situé près de la tour Eiffel, cherche à calculer le coefficient de corrélation entre ses vente (x_1) et le nombre de touristes visitant ce monument (x_2).
Ces deux variables sont influencées par le climat : la consomation de glaces est plus importante lorsqu'il fait chaud et les touristes sont peu enclins à visiter le monument en cas de froid ou de pluie, on appelle x_3 cette variable climatique.

Notion de corrélation partielle

Nous pouvons penser que la corrélation entre x_1 et x_2 est positive, cependant un calcul de coefficient de corrélation simple n'est pas vraiment révélateur du degré de liaison réelle entre ces deux variables ; en effet la variable climatique influence la vente des glaces et la fréquentation des touristes.
En d'autres termes, le coefficient de corrélation simple calculé ainsi intègre l'apport de la variabilité des conditions climatiques sans pouvoir isoler l'influence relative du nombre de touristes.

Exemple introductif

Considèrons le processus

$$X_t = 0.8X_{t-1} + \epsilon_t \ (*)$$

où $(\epsilon_t)_t$ est un bruit blanc indépendant de X_{t-1}.

De par la définition de (X_t), il y a une forte corrélation entre X_t et X_{t-1} ($\rho(1) = 0.8$), qui se répercute entre X_t et X_{t-2} ($\rho(2) = 0.64$), entre X_t et X_{t-3} ($\rho(3) = 0.512$),...
Pourtant le processus (*) semble indiquer qu'il n'y a pas de corrélation "directe" entre X_t et X_{t-h} pour $h > 1$.
L'autororrélation partielle permet en fixant le niveau des variables intermédiaires, de mesurer cette dépendance.

Autocorrélation partielle

Définition

Si (X_t) est un processus stationnaire, l'autocorrélation partielle d'ordre h, notée $\pi(h)$, est définie par:

$$\pi(h) = Corr(X_t - E_c(X_t|X_{t+1},...,X_{t+h-1}); X_{t+h} - E_c(X_{t+h}|X_{t+1},...,X_{t+h-1})),$$

pour $h \geq 2$, où E_c désigne l'espérance conditionnelle linéaire.
On convient que $\pi(0) = 1$ et $\pi(1) = \rho(1)$.

Autocorrélation partielle

Le coefficient de corrélation partielle (PACF) mesure la liaison entre deux variables lorsque l'influence d'une ou des autres variables explicatives est retirée.

Autocorrélation partielle

remarque

Une propriété caractéristique des $AR(p)$ est que la PACF d'un $AR(p)$ est nulle au delà de p : $\pi(h) = 0$ si $h > p$.
Cette propriété est opposée à celle d'un $MA(q)$ pour lequel l'ACF vérifie : $\rho(h) = 0$ si $h > q$.

Autocorrélation partielle

Exemple

Soit X un procéssus centré tel que $X_t = 0.9X_{t-1} + \epsilon_t$ avec ϵ un bruit blanc gaussien centré et réduit et ϵ_t est indépendant de X_s pour tout $s < t$.

Autocorrélation partielle

On calcule d'abord la fonction d'autocorrélation.

$$Corr(X_t, X_{t-1}) = Corr(0.9X_{t-1} + \epsilon_t; X_{t-1}) = 0.9$$

Intéressons nous maintenant aux projections sur le passé et/ou le futur.

$$E_c(X_k|X_2, ..., X_{k-1}) = 0.9X_{k-1} \text{ et } E_c(X_1|X_2, ..., X_{k-1}) = 0.9X_1.$$

On vérifie aussi que $E_c(X_k|X_2, ..., X_{k-1}) = E_c(0.9X_{k-1} + \epsilon_k|X_2, ..., X_{k-1}) = 0.9X_1$. Or ϵ_t est, par définition, indépendant du passé. On a donc bien

$$E_c(X_k|X_2, ..., X_{k-1}) = 0.9X_{k-1}$$

Caractéristiques des processus ARMA : ACF et PACF

On va étudier les caractéristiques propres aux processus AR, MA et ARMA en termes d'autocorrélations simple et partielle afin de donner un schéma d'identification

Caractéristiques des processus ARMA : ACF et PACF

Corrélogramme simple

Cas de l'AR(p) : Soit le processus AR(p)

$$X_t - \sum_{i=1}^{p} \phi_i X_{t-i} = \epsilon_t, \ \forall t \in \mathbb{Z} \text{ ...(2)}$$

Pour calculer sa variance, on multiplie cette équation par X_t et on en prend son espérance :

$$E(X_t^2) - \sum_{i=1}^{p} \phi_i E(X_t X_{t-i}) = E(X_t \epsilon_t)$$

on obtient alors :

$$\gamma(0) = \sum_{i=1}^{p} \phi_i \gamma(i) + \sigma^2$$

Caractéristiques des processus ARMA : ACF et PACF

Corrélogramme simple

Ainsi, puisque $\gamma(i) = \gamma(0)\rho(i)$

$$\gamma(0) = \frac{\sigma^2}{1-\sum_{i=1}^{p}\phi_i\rho(i)}$$

Pour obtenir les autocovariances, on multiplie l'équation (2) par X_{t-h} et on en prend son espérence. On obtient alors les équations suivantes :

$$\gamma(h) - \sum_{i=1}^{p}\phi_i\gamma(h-i) = 0 \ \forall h > 0.$$

Les $p+1$ premières équations sont appelées équations de Yule-Walker.

Caractéristiques des processus ARMA : ACF et PACF

Corrélogramme simple

En divisant par $\gamma(0)$, on obtient alors :

$$\rho(h) = \sum_{i=1}^{p} \phi_i \rho(h-i) = 0 \ \forall h > 0$$

soit une équation de récurrence linéaire homogène d'ordre p dont le polynôme caractéristique

$$z^h - \sum_{i=1}^{p} \phi_i z^{h-i} = z^h(1 - \phi_1 z^{-1} - \ldots - \phi_p z^p) = 0 \ \ldots \ (3)$$

Caractéristiques des processus ARMA : ACF et PACF

Corrélogramme simple

Si le polynôme $\Phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$ admet p racines distinctes $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, ..., \frac{1}{\lambda_p}$, alors les racines de l'équation (3) sont $\lambda_1, \lambda_2, ..., \lambda_p$ et le coefficient d'autocorrélation d'ordre h est alors donné par :

$$\rho(h) = a_1(\lambda_1)^h + a_2(\lambda_2)^h + ... + a_p(\lambda_p)^h$$

où les $a_i, i = 1, ..., p$ sont des constantes déterminées par les conditions initiales.
Puisque la représentation est canonique, les racine $\frac{1}{\lambda_i}, i = 1, ..., p$ sont alors de module supérieur à 1 donc $|\lambda_i| < 1$.

Caractéristiques des processus ARMA : ACF et PACF

Corrélogramme simple

Ainsi, la fonction d'autocorrélation (d'un processus stationnaire) décroit, soit de manière exponentielle (si les racines sont réelles), soit selon des cycles amortis (si les racines sont complexes).

Caractéristiques des processus ARMA : ACF et PACF

Corrélogramme simple

Cas du MA(q) : Calculons les autocovariances pour le processus MA(q) :

$$X_t = \epsilon_t - \sum_{i=1}^{q} \theta_i \epsilon_{t-i}, \ \forall t \in \mathbb{Z}.$$

On obtient

$$\gamma(h) = E(X_t X_{t-h}) = E(\epsilon_t - \sum_{i=1}^{q} \theta_i \epsilon_{t-i})(\epsilon_{t-h} - \sum_{i=1}^{q} \theta_i \epsilon_{t-h-i}) = \begin{cases} (1 + \sum_{i=1}^{q} \theta_i^2)\sigma_2 & \text{si } h = 0 \\ (-\theta_h + \sum_{i=h+1}^{q} \theta_i \theta_{i-h})\sigma^2 & \text{si } 1 \leq h \leq q-1 \\ -\theta_q \sigma^2 & \text{si } h = q \\ 0 & \text{si } h > q \end{cases}$$

Caractéristiques des processus ARMA : ACF et PACF

Corrélogramme simple

Ainsi, les autocorrélations simples s'annulent après l'ordre du MA :

$$\rho(h) = 0 \text{ si } h > q$$

et pour $h \leq q$, sont tels que :

$$\rho(h) = \frac{-\theta_h + \sum_{i=h+1}^{q} \theta_i \theta_{i-h}}{1 + \sum_{i=1}^{q} \theta_i^2}$$

Caractéristiques des processus ARMA : ACF et PACF

Corrélogramme simple

Cas de l'ARMA(q) : le corrélogramme d'un processus ARMA(p,q) est le même que celui d'un AR à partir de l'ordre $q+1$.

Caractéristiques des processus ARMA : ACF et PACF

Corrélogramme partiel

Cas de l'AR(p) : les autocorrélations partielles d'un AR(p) sont telles que :

$$\pi(1) = \rho(1), \pi(2) = \frac{\rho(2) - (\rho(1))^2}{1 - (\rho(1))^2}, \ldots$$

Caractéristiques des processus ARMA : ACF et PACF

Corrélogramme partiel

Cas du MA(q) : Les autocorrélations partielles décroissent soit de manière exponentielle (si les racine sont réelles) soit selon des cycles amortis (si les racines sont complexes).

Caractéristiques des processus ARMA : ACF et PACF

Corrélogramme partiel

Cas de l'ARMA(p,q) : le corrélogramme partiel d'un ARMA(p,q) est le même que celui d'un MA à partir de l'ordre $p+1$.

Caractéristiques des processus ARMA : ACF et PACF

Identification du modèle ARMA

Modèle	AR(p)	MA(q)	ARMA(p,q)
Auto-corrélation	$\rho(h) \xrightarrow[h\to+\infty]{} 0$	$\rho(h) = 0 \; si \; h > q$	$\rho(h) \xrightarrow[h\to+\infty]{} 0$
Auto-corrélation partielle	$\pi(h) = 0 \; si \; h > p$	$\pi(h) \xrightarrow[h\to+\infty]{} 0$	$\pi(h) \xrightarrow[h\to+\infty]{} 0$

Figure: Tableau récapitulatif des propriétés des modèles ARMA

Caractéristiques des processus ARMA

Identifier le modèle correspondant aux corrélogramme simple et corréligramme partiel des séries chronologiques suivantes

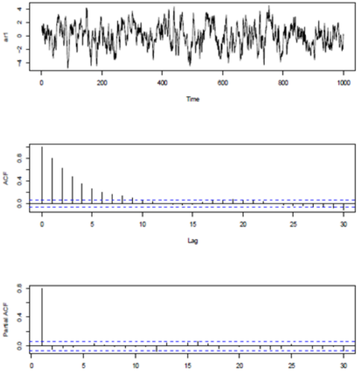

Figure: Série 1

Caractéristiques des processus ARMA

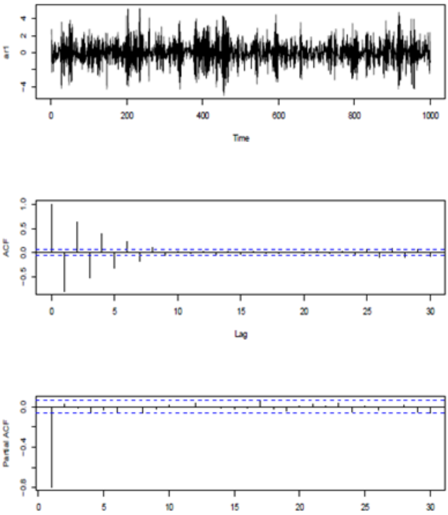

Figure: Série 2

Caractéristiques des processus ARMA

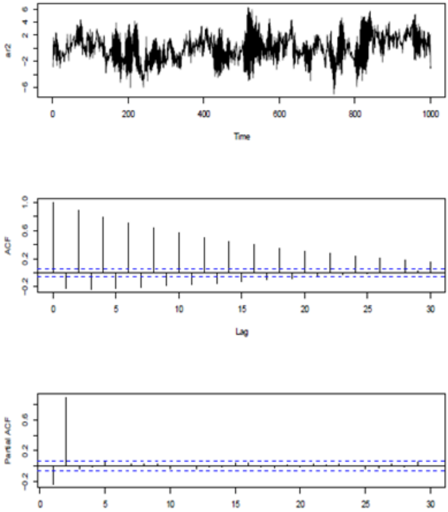

Figure: Série 3

Caractéristiques des processus ARMA

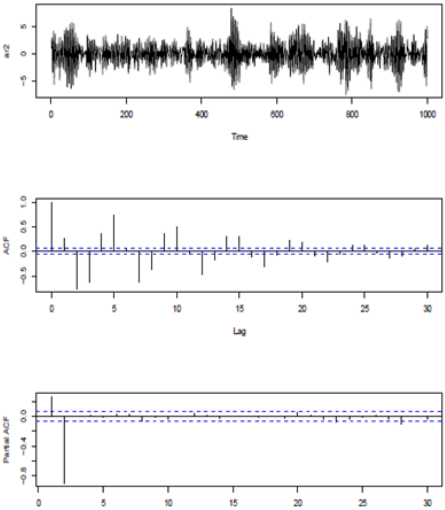

Figure: Série 4

Caractéristiques des processus ARMA

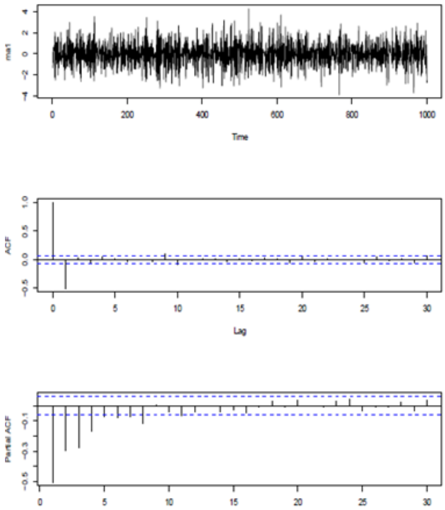

Figure: Série 5

Caractéristiques des processus ARMA

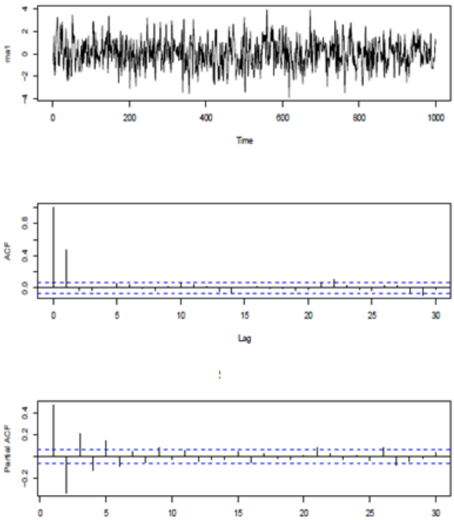

Figure: Série 6

Caractéristiques des processus ARMA

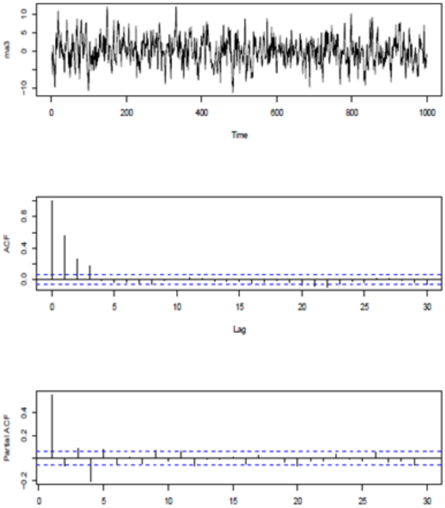

Figure: Série 7

Caractéristiques des processus ARMA

Solution

Série 1 : AR(1)
Série 2 : AR(1)
Série 3 : AR(2)
Série 4 : AR(2)
Série 5 : MA(1)
Série 6 : MA(1)
Série 7 : MA(3)