

CHAPITRE 2

Processus Stationnaires

Les processus dont les propriétés ne varient pas dans le temps jouent un rôle important dans l'analyse des séries chronologiques. Cette invariance dans le temps est appelée *stationnarité*. Elle est préalable à une prévision des valeurs futures. On ne peut se projeter dans l'avenir (prévision des valeurs futures) que si le processus générateur de nos données est "stable". Cette stabilité est à la base de la stationnarité. Il existe divers types de stationnarité. Nous nous intéressons à :

- la stationnarité faible (ou au second ordre ou en covariance).
- la stationnarité forte ou stricte.

1 Propriétés de base

1.1 Stationnarité au second ordre

Définition 1 Un processus $\{X_t, t \in \mathbb{Z}\}$ est dit **du second ordre** si $E(X_t^2) < \infty, \forall t$.

Remarque 2 En particulier, les moments croisés seront finis, i.e., $E(|X_t X_{t+h}|) < \infty, \forall t, h \in \mathbb{Z}$. Ceci peut être montré à l'aide de l'inégalité de Cauchy-Schwarz.

Définition 3 Un processus $\{X_t, t \in \mathbb{Z}\}$ est dit **stationnaire au second ordre** si :

1. Il est du second ordre.
2. $E(X_t)$ est constante, $\forall t \in \mathbb{Z}$.
3. $Cov(X_t, X_{t-h}) = Cov(X_t, X_{t+h}) = \gamma_h, \forall t, h \in \mathbb{Z}$.

Remarque 4 De tels processus sont dits univariés à temps discret (Voir Chapitre 1 pour plus de détails). Si $t \in [0, 1]$ par exemple, alors le processus est dit à temps continu. Nous ne nous intéresserons qu'aux processus linéaires à temps discret.

Exercice 5 Montrer que $\gamma_h = \gamma_{-h}, \forall h \in \mathbb{Z}$.

La condition 3. ci-dessus signifie que, pour un processus stationnaire au second ordre, γ_h ne dépend que de l'écart entre les instants ($t - (t - h) = h$).

$\{\gamma_h, h \in \mathbb{Z}\}$ est appelée **la fonction d'autocovariance** (FACV ou ACVF en anglais : AutoCoVariance Function). Elle a été définie au chapitre 1. Il en est de même de **la fonction d'autocorrélation** (FAC ou ACF en anglais) $\{\rho_h, h \in \mathbb{Z}\}$ où $\rho_h = \frac{\gamma_h}{\gamma_0}$.

h est appelé le retard ou lag (en anglais).

La FACV (ACVF) ou la FAC (ACF) fournissent une mesure utile du degré de dépendance parmi les valeurs d'une série chronologique en des instants différents et, pour cette raison, elle joue un rôle important dans l'identification du processus générateur, i.e., qui a donné naissance aux données et donc, dans la prévision des valeurs futures de la série en fonction des valeurs présente et passées.

Définition 6 Soit x_1, \dots, x_n une série chronologique. La moyenne empirique est $\bar{x}_n = \frac{\sum_{i=1}^n x_i}{n}$. La fonction d'autocovariance empirique est $\hat{\gamma}_h = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_t - \bar{x}_n)(x_{t+|h|} - \bar{x}_n)$, $-n < h < n$. La fonction d'autocorrélation empirique $\hat{\rho}_h = \frac{\hat{\gamma}_h}{\hat{\gamma}_0}$, $-n < h < n$.

Remarque 7 L'utilisation de la division par n dans $\hat{\gamma}_h$ assure que la matrice de covariance $\hat{\Gamma}_n = (\hat{\gamma}_{i-j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ est semi-définie positive. Il en est de même de $\hat{R}_n = (\hat{\rho}_{i-j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$.

$$\hat{\rho}_0 = 1.$$

On a :

- $\gamma_0 > 0$.
- $|\gamma_h| \leq \gamma_0, \forall h$.

- $\gamma_h = \gamma_{-h}, \forall h.$

Et la fonction d'autocorrélation vérifie :

- $\rho_0 = 1.$
- $|\rho_h| \leq 1, \forall h.$
- $\rho_h = \rho_{-h}, \forall h.$

Proposition 8 *La fonction d'autocovariance de tout processus stationnaire $\{X_t\}$ est semi-définie positive.*

Preuve. Soit $\underline{X}_n = (X_n, \dots, X_1)'$ et soit $\underline{a} = (a_1, \dots, a_n)' \in \mathbb{R}^n$. Et soit $\Gamma_n = \text{Var}(\underline{X}_n)$. Alors $\text{Var}(\underline{a}'\underline{X}_n) = \underline{a}'\Gamma_n\underline{a} = \sum_{i=1}^n \sum_{j=1}^n a_i \gamma_{i-j} a_j \geq 0$, i.e., Γ_n est semi-définie positive. ■

Exercice 9 *Soient A et B deux variables aléatoires centrées-réduites, non corrélées. Soit $X_t = A \cos(\omega t) + B \sin(\omega t), t \in \mathbb{Z}$. Montrer que $\{X_t\}$ est un processus stationnaire.*

Solution 10 $E(X_t^2) = \cos^2(\omega t)E(A^2) + \sin^2(\omega t)E(B^2) = 1 < \infty$. Ce processus est du second ordre.

- $E(X_t) = 0, \forall t$. Il est stationnaire en moyenne.
-

$$\begin{aligned}
 \text{Cov}(X_t, X_{t-h}) &= E(X_t X_{t-h}) \\
 &= E(A \cos(\omega t) + B \sin(\omega t))(A \cos(\omega(t-h)) + B \sin(\omega(t-h))) \\
 &= E(A^2)(\cos(\omega t) \cos(\omega(t-h))) + E(B^2) \sin(\omega t) \sin(\omega(t-h)) \\
 &= \cos(\omega t) \cos(\omega(t-h)) + \sin(\omega t) \sin(\omega(t-h)) \\
 &= \cos(\omega t - \omega(t-h)) = \cos(\omega h) = \gamma_h, \forall h.
 \end{aligned}$$

Ce processus est donc stationnaire au second ordre.

Dans ce qui suit, le processus sera noté $\{X_t\}$.

Exercice 11 Soit $\{X_t\}$ un processus solution de l'équation aux différences stochastiques

$$X_t = \epsilon_t + \theta \epsilon_{t-1}$$

où $\{\epsilon_t\}$ est un processus bruit blanc faible de variance σ_ϵ^2 ($\{\epsilon_t\} \sim BB(0, \sigma_\epsilon^2) \equiv WN(0, \sigma_\epsilon^2)$, voir Chapitre 1).

- 1) Calculer la fonction d'autocovariance de $\{X_t\}$.
- 2) Dédire la fonction d'autocorrélation.

1.2 Stationnarité stricte

Définition 12 Un processus $\{X_t\}$ est dit strictement stationnaire si $(X_1, \dots, X_n)' \stackrel{d}{=} (X_{1+h}, \dots, X_{n+h})'$, $\forall h, n, n \geq 1, h \in \mathbb{Z}$. ($\stackrel{d}{=}$ signifie que les deux vecteurs aléatoires ont la même loi de probabilité conjointe). Cela se traduit par : $\{X_t\}$ est strictement stationnaire si : $F_{X_{t_1}, \dots, X_{t_n}}(x_1, \dots, x_n) = F_{X_{t_1+h}, \dots, X_{t_n+h}}(x_1, \dots, x_n)$, $\forall n, n \geq 1, \forall h, t_1, \dots, t_n \in \mathbb{Z}, \forall x_1, \dots, x_n \in \mathbb{R}$, où $F_{X_{t_1}, \dots, X_{t_n}}$ est la fonction de répartition conjointe.

Propriétés d'un processus strictement stationnaire

1. Les variables aléatoires X_t sont identiquement distribuées.
2. $(X_t, X_{t+h}) \stackrel{d}{=} (X_1, X_{1+h})$, $\forall t, h$.
3. $\{X_t\}$ est aussi faiblement stationnaire si $E(X_t^2) < \infty, \forall t$.
4. La stationnarité faible n'implique pas la stationnarité stricte sauf si le processus est gaussien.

Soit $\{\epsilon_t\}$ un processus bruit blanc fort gaussien de variance 1. Soit $X_t = \begin{cases} \epsilon_t & \text{si } t \text{ est pair} \\ \frac{\epsilon_t^2 - 1}{\sqrt{2}} & \text{si } t \text{ est impair} \end{cases}$

- 1) Montrer que $\{X_t\}$ est stationnaire au second ordre.
- 2) Calculer $P(X_t > 1)$. Dédire que $\{X_t\}$ n'est pas strictement stationnaire.

2 Processus linéaires

La classe des modèles de séries chronologiques linéaires, qui inclut la classe des modèles autorégressifs moyenne mobile (*ARMA* : **A**uto**R**egressive **M**oving **A**verage en anglais) fournit un cadre général pour l'étude des processus stationnaires. Ceci est à la base du théorème de décomposition de Wold qui sera formulé plus loin.

Définition 13 *Le processus aléatoire $\{X_t\}$ est un **processus linéaire** s'il possède la représentation*

$$X_t = \sum_{i=-\infty}^{\infty} \psi_i \epsilon_{t-i}, \forall t \quad (1)$$

où $\{\epsilon_t\}$ est un bruit blanc $WN(0, \sigma_\epsilon^2)$ et $\{\psi_i\}$ est une suite de constantes vérifiant $\sum_{i=-\infty}^{\infty} |\psi_i| < \infty$.

(1) s'écrit de façon plus compacte

$$X_t = \Psi(B)\epsilon_t$$

où $\Psi(B) = \sum_{i=-\infty}^{\infty} \psi_i B^i$.

Un processus linéaire est un processus **moyenne mobile d'ordre ∞** ($MA(\infty)$) si $\psi_i = 0$ pour $i < 0$, i.e.,

$$X_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}, \forall t$$

Remarque 14 *La condition $\sum_{i=-\infty}^{\infty} |\psi_i| < \infty$ assure que la somme infinie dans*

(1) *converge presque sûrement (ps).*

En effet, nous avons :

1) $E(|\epsilon_t|) < \sigma$ car, en vertu de l'inégalité de Cauchy-Schwarz, $E(|\epsilon_t|) \leq (E(\epsilon_t^2))^{1/2} = \sigma$

2) $E(|X_t|) = E\left(\left|\sum_{i=-\infty}^{\infty} \psi_i \epsilon_{t-i}\right|\right) \leq \sum_{i=-\infty}^{\infty} |\psi_i| E(|\epsilon_{t-i}|) = \sum_{i=-\infty}^{\infty} |\psi_i| E(|\epsilon_t|) \leq$

$\sigma \sum_{i=-\infty}^{\infty} |\psi_i| < \infty$ ceci impliquant que $|X_t| < \infty$, ps et donc la somme infinie converge ps.

Remarque 15 La condition donnée dans la remarque 1 assure aussi que

$$\sum_{i=-\infty}^{\infty} \psi_i^2 < \infty. \text{ En effet } \sum_{i=-\infty}^{\infty} |\psi_i| < \infty \Leftrightarrow \left(\sum_{i=-\infty}^{\infty} |\psi_i| \right)^2 < \infty \Leftrightarrow \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} |\psi_i| |\psi_j| < \infty$$

$$\Leftrightarrow \sum_{i=-\infty}^{\infty} \psi_i^2 + \sum_{i=-\infty}^{\infty} \sum_{\substack{j=-\infty \\ i \neq j}}^{\infty} |\psi_i| |\psi_j| < \infty. \text{ Donc la somme infinie dans (1)}$$

converge en moyenne quadratique. En effet $E\left(\sum_{i=-\infty}^{\infty} \psi_i \epsilon_{t-i} - \sum_{i=-n}^n \psi_i \epsilon_{t-i}\right)^2 = \sigma^2 \sum_{|i|>n} \psi_i^2 \rightarrow 0$ quand $n \rightarrow \infty$ car c'est le reste d'une série convergente.

Remarque 16 1) La condition (ψ_i) absolument sommable est plus forte que la condition (ψ_i) de carré sommable. Il existe des processus linéaires pour lesquels la suite (ψ_i) est de carré sommable mais n'est pas absolument sommable.

2) L'opérateur $\Psi(B)$ peut être interprété comme un filtre qui est appliqué au bruit blanc $\{\epsilon_t\}$ qui est l'input pour produire le processus $\{X_t\}$ qui est l'output.

La proposition suivante établit qu'un filtre linéaire appliqué à un processus stationnaire produit un processus stationnaire.

Proposition 17 Soit $\{Z_t\}$ un processus stationnaire de moyenne 0 et de fonction d'autocovariance $\{\gamma_h^Z\}$. Si $\sum_{i=-\infty}^{\infty} |\psi_i| < \infty$, alors le processus $X_t = \sum_{i=-\infty}^{\infty} \psi_i Z_{t-i}$ est stationnaire de moyenne 0 et de fonction d'autocovariance

$$\gamma_h^X = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_{h+k-j}^Z$$

Dans le cas spécial où $\{X_t\}$ est un processus linéaire alors $\gamma_h^X = \sum_{j=-\infty}^{\infty} \psi_j \psi_{j-h} \sigma_\epsilon^2$.

Preuve. 1) $E(X_t) = E\left(\sum_{i=-\infty}^{\infty} \psi_i Z_{t-i}\right) = \sum_{i=-\infty}^{\infty} \psi_i E(Z_{t-i}) = 0$. L'interchangement entre les opérateurs $\sum_{i=-\infty}^{\infty}$ et espérance peut être justifié par l'absolue sommabilité des ψ_i . (La moyenne ne dépend pas de t : elle est constante).

2) $E(X_t X_{t-h}) = E\left(\sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k Z_{t-j} Z_{t-h-k}\right) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k E(Z_{t-j} Z_{t-h-k}) =$
 $\sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_{h+k-j}^Z$ (elle ne dépend que de la différence entre les instants
qui est $t - j - (t - h - k) = h + k - j$).

Le cas spécial d'un processus linéaire se distingue par le fait que $Z_t = \epsilon_t$ et donc

$$\begin{aligned} E(Z_{t-j} Z_{t-h-k}) &= E(\epsilon_{t-j} \epsilon_{t-h-k}) = \begin{cases} \sigma_\epsilon^2 & \text{si } t-j = t-h-k \\ 0 & \text{sinon} \end{cases} \\ &= \begin{cases} \sigma_\epsilon^2 & \text{si } k = j-h \\ 0 & \text{sinon} \end{cases} \end{aligned}$$

■

Ceci montre que le processus $\{X_t\}$ défini par (1) est stationnaire au second ordre. L'absolue convergence de (1) implique que des filtres de la forme $\alpha(B) = \sum_{j=-\infty}^{\infty} \alpha_j B^j$ et $\beta(B) = \sum_{j=-\infty}^{\infty} \beta_j B^j$ avec des coefficients absolument sommables peuvent être appliqués à un processus stationnaire $\{Y_t\}$ pour générer un nouveau processus stationnaire

$$W_t = \sum_{j=-\infty}^{\infty} \psi_j Y_{t-j} = \Psi(B) Y_t$$

où $\Psi(B) = \alpha(B)\beta(B)$ et $\psi_j = \sum_{k=-\infty}^{\infty} \alpha_k \beta_{j-k} = \sum_{k=-\infty}^{\infty} \beta_k \alpha_{j-k}$. (A montrer en exercice).

Exemple 18 Processus AR(1)

Un processus AR(1), $\{X_t\}$ est une solution de l'équation aux différences stochastiques

$$X_t = \phi X_{t-1} + \epsilon_t \quad (2)$$

où $\{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2)$ et ϵ_t est non corrélé avec X_s pour tout $s < t$. Pour montrer qu'une telle solution existe, si $|\phi| < 1$, et est l'unique solution de (2), on considère le processus linéaire

$$X_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i} \quad (3)$$

$$\sum_{i=0}^{\infty} \phi^i \epsilon_{t-i} = \epsilon_t + \sum_{i=1}^{\infty} \phi^i \epsilon_{t-i}. \text{ Posons } j = i - 1. \text{ On obtient } \epsilon_t + \sum_{i=1}^{\infty} \phi^i \epsilon_{t-i} = \epsilon_t + \sum_{j=0}^{\infty} \phi^{j+1} \epsilon_{t-1-j} = \epsilon_t + \phi \sum_{j=0}^{\infty} \phi^j \epsilon_{t-1-j}. \text{ D'où } X_t = \phi X_{t-1} + \epsilon_t.$$

Montrons qu'elle est unique. Pour cela, supposons que $Y_t = \phi Y_{t-1} + \epsilon_t$ où $\{Y_t\}$ est une autre processus. Alors,

$$\begin{aligned} Y_t &= \phi Y_{t-1} + \epsilon_t \\ &= \epsilon_t + \phi \epsilon_{t-1} + \phi^2 Y_{t-1} \\ &= \dots \\ &= \epsilon_t + \phi \epsilon_{t-1} + \phi^2 \epsilon_{t-2} + \dots + \phi^k \epsilon_{t-k} + \phi^{k+1} Y_{t-k-1} \end{aligned}$$

Si $\{Y_t\}$ est stationnaire au second ordre, alors $E(Y_t^2) < \infty$ et est indépendant de t . Ainsi

$$\begin{aligned} E(Y_t - \sum_{i=0}^k \phi^i \epsilon_{t-i})^2 &= E(\phi^{k+1} Y_{t-k-1})^2 \\ &= \phi^{2k+2} E(Y_{t-k-1}^2) \\ &\rightarrow 0 \text{ quand } k \rightarrow \infty \end{aligned}$$

Ceci implique que Y_t est égal à la limite en moyenne quadratique $\sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$ et donc, le processus défini par l'équation (3) est l'unique solution stationnaire au second ordre de (2).

Lorsque $|\phi| > 1$, la série définie par (3) n'est pas convergente. Mais on peut réécrire le modèle (2) de façon que le processus à l'instant t (ou $t-1$ ou ...) s'écrive en fonction de son futur.

$$X_{t-1} = -\frac{1}{\phi} \epsilon_t + \frac{1}{\phi} X_t \quad (4)$$

En itérant (4) on obtient

$$\begin{aligned} X_{t-1} &= -\frac{1}{\phi} \epsilon_t - \frac{1}{\phi^2} \epsilon_{t+1} + \frac{1}{\phi^2} X_{t+1} \\ &= \dots \\ &= -\frac{1}{\phi} \epsilon_t - \frac{1}{\phi^2} \epsilon_{t+1} - \dots - \frac{1}{\phi^{k+1}} \epsilon_{t+k} + \frac{1}{\phi^{k+1}} X_{t+k} \end{aligned}$$

En utilisant les mêmes arguments que ci-dessus, on montre que

$$X_t = -\sum_{i=1}^{\infty} \frac{1}{\phi^i} \epsilon_{t+i-1}$$

est l'unique solution stationnaire au second ordre de (4).

Remarque 19 Cette solution n'est pas très naturelle car X_t est corrélée avec le futur du bruit $\epsilon_{t+1}, \epsilon_{t+2}, \dots$. Ceci contraste avec la solution (3) dans laquelle on voit que X_t est non corrélé avec ϵ_s pour $s > t$. On restreindra notre attention aux processus $AR(1)$ tels que $|\phi| < 1$.

Définition 20 Le processus $\{X_t\}$ est dit **causal** ou **non anticipatif** ou **indépendant du futur** si X_t s'écrit en fonction du présent et du passé du bruit blanc.

Exercice 21 Montrer que le processus $AR(1)$ donné par (2) est non stationnaire si $\phi = \pm 1$.

Exercice 22 Calculer la fonction d'autocovariance et la fonction d'autocorrélation d'un processus $AR(1)$ causal.

Remarque 23 $(2) \iff X_t - \phi X_{t-1} = \epsilon_t \iff (1 - \phi B)X_t = \epsilon_t$. Considérons $(1 - \phi z) = 0, z \in \mathbb{C}$. $|\phi| < 1 \iff |z| > 1$. Les racines de $\Phi(z)$ sont à l'extérieur du cercle unité.

Remarque 24 Pour un processus $AR(1)$, $X_t = \phi X_{t-1} + \epsilon_t$, la solution unique de cette équation aux récurrences stochastiques est $X_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$. On peut aussi la trouver en considérant l'équation $(1 - \phi B)X_t = \epsilon_t$. Et donc $X_t = (1 - \phi B)^{-1} \epsilon_t \iff X_t = \sum_{i=0}^{\infty} \phi^i B^i \epsilon_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$.

Remarque 25 Le processus $AR(1)$ est donc stationnaire au second ordre si et seulement si $\phi \neq \pm 1$. Il est de plus causal si et seulement si $|\phi| < 1$.

Remarque 26 Il est rappelé qu'un modèle est un cadre mathématique sensé refléter au mieux l'évolution des données d'observation. Les modèles linéaires sont les modèles les plus simples. Ce sont des modèles de moyenne conditionnelle.

Exemple 27 Cet exemple introduit le modèle $ARMA(1,1)$. Les modèles $ARMA(p,q)$ seront étudiés dans le chapitre 3.

Définition 28 Un processus $\{X_t\}$ suit un modèle $ARMA(1, 1)$ s'il est solution de l'équation aux différences stochastiques

$$X_t = \phi X_{t-1} + \epsilon_t + \theta \epsilon_{t-1} \quad (5)$$

où $\{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2)$ et ϵ_t est non corrélé avec X_s pour tout $s < t$ et $\phi + \theta \neq 0$.

L'équation (5) peut s'écrire de façon plus compacte :

$$\Phi(B)X_t = \Theta(B)\epsilon_t \quad (6)$$

où $\Phi(B) = 1 - \phi B$ et $\Theta(B) = 1 + \theta B$. Si $|\phi| < 1$, $X_t = \Phi^{-1}(B)\Theta(B)\epsilon_t = \sum_{i=0}^{\infty} \phi^i B^i (1 + \theta B)\epsilon_t = \underbrace{\sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}}_{(1)} + \theta \underbrace{\sum_{i=0}^{\infty} \phi^i \epsilon_{t-1-i}}_{(2)}$.

(1) = $\epsilon_t + \sum_{i=1}^{\infty} \phi^i \epsilon_{t-i} = \epsilon_t + \phi \sum_{i=0}^{\infty} \phi^i \epsilon_{t-1-i}$ (voir calcul ci-dessus, modèle $AR(1)$). Par conséquent

$$X_t = \epsilon_t + (\phi + \theta) \sum_{i=0}^{\infty} \phi^i \epsilon_{t-1-i} = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \quad (7)$$

$$\text{avec } \psi_0 = 1 \text{ et } \psi_i = (\phi + \theta)\phi^{i-1}, i \geq 1 \quad (8)$$

Donc, si $|\phi| < 1$, $\forall \theta$ le processus $\{X_t\}$ s'écrit en fonction du présent et du passé de ϵ_t . $\{X_t\}$ est stationnaire au second ordre et causal. (A montrer en exercice).

Exercice 29 Si $|\phi| > 1$, Montrer que le processus est stationnaire et non causal.

Remarque 30 Si $\phi = \pm 1$, le processus est non stationnaire (à montrer).

En résumé

- Une solution stationnaire des équations (5) existe si et seulement si $\phi \neq \pm 1$.
- Si $|\phi| < 1$ alors la solution stationnaire unique est donnée par (7). Dans ce cas, on dit que $\{X_t\}$ est causal ou est une fonction causale de $\{\epsilon_t\}$ puisque X_t s'exprime en fonction des valeurs présente et passées $\epsilon_s, s \leq t$.

- Si $|\phi| > 1$, la solution (en exercice) est non causale puisque X_t s'exprime en fonction des valeurs présente et futures $\epsilon_s, s \geq t$.

Un processus $\{X_t\}$ est donc **causal** si $X_t = f(\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \dots)$ où f est une fonction mesurable.

Il existe un concept dual de la causalité, c'est l'**inversibilité**. Un processus est **inversible** si ϵ_t s'écrit en fonction de $X_s, s \leq t$, i.e., $\epsilon_t = g(X_t, X_{t-1}, X_{t-2}, \dots)$ où g est une fonction mesurable.

Un processus $ARMA(1, 1)$ est inversible si $|\theta| < 1$. En effet, de (6) on déduit que $\epsilon_t = \Theta^{-1}(B)\Phi(B)X_t$. Pour déterminer $\Theta^{-1}(B)$, posons $\chi(B) = \Theta^{-1}(B) = 1 + \chi_1 B + \chi_2 B^2 + \dots$. Puisque $\Theta(B)\Theta^{-1}(B) = 1$, on a

$$(1 + \theta B)(1 + \chi_1 B + \chi_2 B^2 + \dots) = 1$$

On déduit que $\theta + \chi_1 = 0 \Leftrightarrow \chi_1 = -\theta$, $\theta\chi_1 + \chi_2 = 0$ d'où $\chi_2 = \theta^2, \dots$, $\chi_j = (-\theta)^j, j \in \mathbb{N}$. D'où $\epsilon_t = X_t - (\phi + \theta) \sum_{j=1}^{\infty} (-\theta)^{j-1} X_{t-j}$. Lorsque $|\theta| > 1$, le processus $ARMA(1, 1)$ est **non inversible**. En effet, dans ce cas, ϵ_t s'écrit en fonction du présent et du futur. On ne considèrera pas le cas où $\theta = \pm 1$.

Remarque 31 On conclut que la racine de $\Theta(z)$ est à l'extérieur du cercle (ou du disque) unité si et seulement si le processus $ARMA(1, 1)$ est inversible. En effet $(\Theta(z) = 0 \Rightarrow |z| > 1) \Leftrightarrow |\theta| < 1$. On dira donc que le processus $ARMA(1, 1)$ est inversible si et seulement si la racine de $\Theta(z)$ est à l'extérieur du cercle unité. On restreindra donc notre attention aux processus $ARMA$ stationnaires et causaux (pluriel de causal) et inversibles.

3 Propriétés de la moyenne empirique et de la fonction d'autocorrélation empirique

Pour un processus stationnaire au second ordre, $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ est un estimateur sans biais de $E(X_t) = \mu$.

$$\begin{aligned} \text{L'erreur quadratique moyenne de } \bar{X}_n \text{ est } E(\bar{X}_n - \mu)^2 &= \text{Var}(\bar{X}_n) = \\ \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \underbrace{\text{Cov}(X_i, X_j)}_{=\gamma_{i-j}} &= \frac{1}{n^2} \sum_{i-j=-n+1}^{n-1} (n - |i - j|) \gamma_{i-j} \\ &= \frac{1}{n} \sum_{h=-n+1}^{n-1} \left(1 - \frac{|h|}{n}\right) \gamma_h. \end{aligned}$$

Remarque 32 Si $\gamma_h \rightarrow 0$ quand $h \rightarrow \infty$ alors $\text{Var}(\bar{X}_n) \rightarrow 0$ et \bar{X}_n converge vers μ en moyenne quadratique.

Proposition 33 Si $\{X_t\}$ est un processus stationnaire au second ordre de moyenne μ et de fonction d'autocovariance $\{\gamma_h, h \in \mathbb{Z}\}$ alors :

- $E(\bar{X}_n - \mu)^2 = \text{Var}(\bar{X}_n) \rightarrow 0$ si $\gamma_n \rightarrow 0$ quand $n \rightarrow \infty$
- $nE(\bar{X}_n - \mu)^2 \rightarrow \sum_{|h|<\infty} \gamma_h$ si $\sum_{|h|<\infty} \gamma_h < \infty$

Si le processus $\{X_t\}$ est gaussien de moyenne μ et de fonction d'autocovariance $\{\gamma_h, h \in \mathbb{Z}\}$, alors :

$$\sqrt{n}(\bar{X}_n - \mu) \sim \mathcal{N}(0, \sum_{|h|<n} (1 - \frac{|h|}{n})\gamma_h)$$

Si $\gamma_h, |h| < n$ est connue on peut donner un intervalle de confiance de niveau $1 - \alpha = 0.95$, par exemple, pour μ . Si le processus n'est pas gaussien, une bonne approximation est donnée par

$$]\bar{X}_n - 1.96 \frac{v^{1/2}}{\sqrt{n}}, \bar{X}_n + 1.96 \frac{v^{1/2}}{\sqrt{n}}[$$

où $v = \sum_{|h|<\infty} (1 - \frac{|h|}{n})\gamma_h$. Si $\gamma_h, |h| < \infty$ est inconnue, on l'estime par $\hat{\gamma}_h$ et $\hat{v}_n = \sum_{|h|\leq\sqrt{n}} (1 - \frac{|h|}{n})\hat{\gamma}_h$.

Rappel : $\hat{\gamma}_h = \frac{1}{n} \sum_{t=1}^{n-|h|} (X_{t+|h|} - \bar{X}_n)(X_t - \bar{X}_n)$ et $\hat{\rho}_h = \frac{\hat{\gamma}_h}{\hat{\gamma}_0}$ pour un processus stationnaire au second ordre.

Remarque 34 Box et Jenkins (1970) suggèrent que $n \geq 50$ et $h \leq n/4$.

La fonction d'autocorrélation joue un rôle important dans le choix d'un modèle adéquat pour représenter l'évolution de la série chronologique. On admettra que, pour les modèles linéaires en général et les modèles *ARMA* en particulier, le vecteur $\hat{\underline{\rho}}_k = (\hat{\rho}_1, \dots, \hat{\rho}_k)'$ est **approximativement distribué pour n grand** comme suit :

$$\hat{\underline{\rho}}_k \sim \mathcal{N}(\underline{\rho}_k, \frac{1}{n}W) \quad (9)$$

où $\underline{\rho}_k = (\rho_1, \dots, \rho_k)'$ et W est la matrice de covariance du vecteur $\widehat{\underline{\rho}}_k$ dont le (i, j) - ème élément est donné par **la formule de Bartlett** :

$$\begin{aligned} w_{ij} &= \sum_{k=-\infty}^{\infty} \{ \rho_{k+i} \rho_{k+j} + \rho_{k-i} \rho_{k+j} + 2\rho_i \rho_j \rho_k^2 \\ &\quad - 2\rho_i \rho_k \rho_{k+j} - 2\rho_j \rho_k \rho_{k+i} \} \\ &= \sum_{k=1}^{\infty} \{ \rho_{k+i} + \rho_{k-i} - 2\rho_i \rho_k \} \{ \rho_{k+j} + \rho_{k-j} - 2\rho_j \rho_k \} \end{aligned}$$

Exercice 1) Calculer w_{ij} pour le processus bruit blanc $\{\epsilon_t\}$
 2) Calculer w_{ij} pour le processus $MA(1)$: $X_t = \epsilon_t + \theta \epsilon_{t-1}$
 3) Calculer w_{ij} pour le processus $AR(1)$: $X_t = \phi X_{t-1} + \epsilon_t$ où $|\phi| < 1$.

Solution 35 1) *Le processus bruit blanc est non corrélé par définition. Aussi, $\rho_h = 0$ si $|h| > 0$. Et $w_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$. De (9) on déduit que $\widehat{\underline{\rho}}_k \sim \mathcal{N}(\mathbf{0}_{\mathbb{R}^k}, \frac{1}{n} I_k)$. Donc pour n grand, $\widehat{\rho}_1, \dots, \widehat{\rho}_k$ sont iid de loi normale de moyenne 0 et de variance $\frac{1}{n}$. Ce résultat est à la base du test que des observations forment un bruit blanc fort en utilisant l'ACF. Dans ce cas, on teste, pour tout h*

$H_0 : \rho_h = 0$ vs $H_1 : \rho_h \neq 0$. La statistique de test, sous H_0 est $\frac{\widehat{\rho}_h}{\sqrt{n}}$ et la règle de décision est :

$$\text{si } \left| \frac{\widehat{\rho}_h}{\frac{1}{\sqrt{n}}} \right| > 1.96 \text{ alors on rejette } H_0 \text{ au seuil } \alpha = 0.05$$

Il existe un autre test pour tester si le processus est un bruit blanc. C'est un test global qui se base sur toutes les autocorrélations au lieu de les tester une à une. C'est le test de Box-Pierce (dont la statistique de test sera notée $BP(K)$). Il a été modifié par Box-Ljung (dont la statistique de test sera notée $BL(K)$) lorsque la taille T de la série est modérée.

On teste $H_0 : \rho_1 = \rho_2 = \dots = \rho_K = 0$ vs $H_1 : \exists j \in \{1, \dots, K\}, \rho_j \neq 0$.

$$BP(K) = T \sum_{j=1}^K \widehat{\rho}_j^2 \sim \chi_{K-p}^2$$

$$BL(K) = T(T+2) \sum_{j=1}^K \frac{\widehat{\rho}_j^2}{T-j} \sim \chi_{K-p}^2$$

où p est le nombre de paramètres estimés dans le modèle proposé pour la série chronologique initiale et les $\hat{\rho}_j, j = 1, \dots, K$ sont les autocorrélations des résidus du modèle estimé proposé pour la série.

2) On sait que $\rho_h = 0$ si $|h| \geq 2$. Calculer w_{ij} .

3) a) Calculons d'abord $\rho_h, h \in \mathbb{Z}$

Puisque $|\phi| < 1$, le processus $AR(1)$ est stationnaire et causal. Donc en particulier $\mu = E(X_t) = \phi E(X_{t-1}) + E(\epsilon_t)$

$\Rightarrow \mu = \phi\mu$ (le processus étant stationnaire) $\Leftrightarrow (1 - \phi)\mu = 0 \Rightarrow \mu = 0$ car $|\phi| < 1$.

Donc $\gamma_h = Cov(X_t, X_{t-h}) = E(X_t X_{t-h}) = E((\phi X_{t-1} + \epsilon_t) X_{t-h}) = \phi E(X_{t-1} X_{t-h}) + E(\epsilon_t X_{t-h}) = \phi \gamma_{h-1} = \phi^2 \gamma_{h-2}$, etc.

En particulier, (le processus étant stationnaire $\gamma_h = \gamma_{-h}$) on a $\gamma_h = \phi^{|h|} \gamma_0, |h| \geq 1$.

Calcul de $\gamma_0 = Var(X_t) : \gamma_0 = Cov(X_t, X_t) = Cov(\phi X_{t-1} + \epsilon_t)(\phi X_{t-1} + \epsilon_t) = \phi^2 \gamma_0 + \sigma_\epsilon^2$ (le passé du processus est non corrélé avec le futur du bruit. D'où $\gamma_0 = \frac{\sigma_\epsilon^2}{1-\phi^2}$. Et donc $\rho_h = \frac{\gamma_h}{\gamma_0} = \phi^{|h|}, h \in \mathbb{Z}$.

b) Calculer w_{ij} .

Exercice 36 1) Simuler un processus $AR(1)$ avec $\phi = 0.8$ puis avec $\phi = -0.8$. Tracer la fonction d'autocorrélation.

2) Simuler un processus $MA(1)$ avec $\theta = 0.7$ puis avec $\phi = -0.7$. Tracer la fonction d'autocorrélation.

4 Prédiction d'une série chronologique stationnaire

On veut effectuer des prévisions des valeurs futures $X_{T+h}, h = 1, 2, \dots$ d'une série chronologique stationnaire de moyenne μ et de fonction d'autocovariance γ en fonction des valeurs observées $\{X_T, X_{T-1}, \dots, X_1\}$. On veut trouver la combinaison linéaire de $1, X_T, X_{T-1}, \dots, X_1$ qui fournisse la meilleure prévision $P_n X_{n+1}$ de X_{T+h} au sens du critère de l'erreur quadratique moyenne. Donc $P_n X_{n+1} = a_0 + a_1 X_T + a_2 X_{T-1} + \dots + a_T X_1$. On trouve les valeurs a_0, a_1, \dots, a_T qui minimisent la fonction

$$S(a_0, a_1, \dots, a_T) = E(X_{T+h} - a_0 - a_1 X_T - a_2 X_{T-1} - \dots - a_T X_1)^2 \quad (10)$$

Remarque 37 Dans le cas gaussien $P_T X_{T+h} = \hat{X}_T(h) = \mathbb{E}(X_{T+h} \mid \mathcal{F}_T)$ où \mathcal{F}_T désigne l'information disponible jusqu'à l'instant T .

On résoud le système d'équations :

$$\frac{\partial S(a_0, a_1, \dots, a_T)}{\partial a_j} = 0, j = 0, 1, \dots, T$$

On trouve

$$E(X_{T+h} - a_0 - \sum_{i=1}^T a_i X_{T+1-i}) = 0, j = 0 \quad (11)$$

$$E(X_{T+h} - a_0 - \sum_{i=1}^T a_i X_{T+1-i}) X_{T+1-j} = 0, j = 1, \dots, T \quad (12)$$

De l'équation (11), on déduit que $\mu = \frac{a_0}{1-a_1-\dots-a_T} \Leftrightarrow a_0 = \mu(1-a_1-\dots-a_T)$. En reportant cette valeur dans (12), on obtient le système

$$\mathbb{E} \left[\left\{ (X_{T+h} - \mu) - \sum_{i=1}^T a_i (X_{T+1-i} - \mu) \right\} (X_{T+1-j} - \mu) \right] = 0, j = 1, \dots, T$$

ou encore, après avoir effectué

$$\begin{cases} \gamma_h = a_1 \gamma_0 + a_2 \gamma_1 + \dots + a_T \gamma_{T-1} \\ \gamma_{h+1} = a_1 \gamma_1 + a_2 \gamma_0 + \dots + a_T \gamma_{T-2} \\ \gamma_{h+2} = a_1 \gamma_2 + a_2 \gamma_1 + \dots + a_T \gamma_{T-3} \\ \vdots \\ \gamma_{T+h-1} = a_1 \gamma_{T-1} + a_2 \gamma_{T-2} + \dots + a_T \gamma_0 \end{cases} \quad (13)$$

$$\Leftrightarrow \Gamma_T \underline{a}_T = \gamma_T(h)$$

avec $\Gamma_T = (\gamma_{i-j})_{\substack{1 \leq i \leq T \\ 1 \leq j \leq T}}$, $\underline{a}_T = (a_1, \dots, a_T)'$ et $\gamma_T(h) = (\gamma_h, \gamma_{h+1}, \dots, \gamma_{T+h-1})'$.

Ainsi, $P_T X_{T+h} = \mu + \sum_{i=1}^T a_i (X_{T+1-i} - \mu)$ où $\underline{a}_T = (a_1, \dots, a_T)'$ vérifie (13).

L'erreur de prévision est $E(X_{T+h} - P_T X_{T+h})^2 = E(X_{T+h} - \mu - \sum_{i=1}^T a_i (X_{T+1-i} - \mu))^2 = E(X_{T+h} - \mu)^2 - 2 \sum_{i=1}^T a_i E(X_{T+h} - \mu)(X_{T+1-i} - \mu) + \sum_{i=1}^T \sum_{j=1}^T a_i a_j E(X_{T+1-i} - \mu)(X_{T+1-j} - \mu)$
 $= \gamma_0 - 2 \underline{a}_T' \gamma_T(h) + \underbrace{\underline{a}_T' \Gamma_T \underline{a}_T}_{=\gamma_T(h)} = \gamma_0 - \underline{a}_T' \gamma_T(h).$

Le calcul direct du vecteur \underline{a}_T en résolvant le système d'équations linéaires est coûteux puisque la dimension de la matrice Γ_T est grande lorsque la taille de la série chronologique est grande. **L'algorithme de Durbin-Levinson** est un algorithme récursif qui permet de calculer $P_{h+1}X_{h+2}$ en fonction de P_hX_{h+1} . C'est un algorithme récursif. Posons

$$P_hX_{h+1} = \phi_{h1}X_h + \phi_{h2}X_{h-1} + \cdots + \phi_{hh}X_1$$

où $\underline{\phi}_h = (\phi_{h1}, \phi_{h2}, \dots, \phi_{hh})' = \Gamma_h^{-1}\underline{\gamma}_h$ et $\underline{\gamma}_h = (\gamma_1, \dots, \gamma_h)'$ (13). et l'erreur quadratique moyenne est

$$\mathbb{E}(X_{h+1} - P_hX_{h+1})^2 = v_h = \gamma_0 - \underline{\phi}_h' \underline{\gamma}_h$$

Dans le chapitre 3, la fonction $\{\phi_{hh}, h = 1, 2, \dots\}$ sera **la fonction d'autocorrélation partielle**. ($\phi_{00} = 1 = \rho_0$).

Algorithme de Durbin-Levinson : les coefficients $\phi_{h1}, \phi_{h2}, \dots, \phi_{hh}$ peuvent être calculés de façon récursive à partir des équations

$$\begin{aligned} \phi_{hh} &= [\gamma_h - \sum_{j=1}^{h-1} \phi_{h-1,j} \gamma_{h-j}] v_{h-1}^{-1} \\ \begin{pmatrix} \phi_{h1} \\ \phi_{h2} \\ \vdots \\ \phi_{h,h-1} \end{pmatrix} &= \begin{pmatrix} \phi_{h-1,1} \\ \phi_{h-1,2} \\ \vdots \\ \phi_{h-1,h-1} \end{pmatrix} - \phi_{hh} \begin{pmatrix} \phi_{h-1,h-1} \\ \phi_{h-1,h-2} \\ \vdots \\ \phi_{h-1,1} \end{pmatrix} \\ \text{et } v_h &= v_{h-1}(1 - \phi_{hh}^2) \\ \phi_{11} &= \frac{\gamma_1}{\gamma_0} = \rho_1 \end{aligned}$$

Application

Considérons le processus $AR(1)$ (2). Pour ce processus $\gamma_h = \frac{\phi^{|h|}}{1-\phi^2} \sigma_\epsilon^2, h \in \mathbb{Z}$. L'équation (13) s'écrit dans ce cas : ($h = i - j$)

$$\frac{\sigma_\epsilon^2}{1-\phi^2} \begin{pmatrix} 1 & \phi & \phi^2 & \cdots & \cdots & \phi^{n-1} \\ \phi & 1 & \phi & \phi^2 & \cdots & \phi^{n-2} \\ \phi^2 & & 1 & \phi & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \phi^2 \\ & & & & \ddots & \phi \\ \phi^{n-1} & \cdots & \cdots & \phi & 1 & \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ \vdots \\ a_n \end{pmatrix} = \frac{\sigma_\epsilon^2}{1-\phi^2} \begin{pmatrix} \phi \\ \phi^2 \\ \vdots \\ \phi^n \end{pmatrix}$$

Une solution est $\underline{a}_n = (\phi, 0, \dots, 0)'$. D'où $P_n X_{n+1} = \phi X_n$ est la meilleure prévision au sens de l'erreur quadratique moyenne qui est égale à

$$\mathbb{E}(X_{n+1} - P_n X_{n+1})^2 = \sigma_\epsilon^2$$

En notant $\mathcal{F}_n = \sigma(X_n, X_{n-1}, X_{n-2}, \dots)$ alors $\mathbb{E}(X_{n+1} \mid \mathcal{F}_n) = \mathbb{E}(\phi X_n + \epsilon_{n+1} \mid \mathcal{F}_n) = \phi X_n + \mathbb{E}(\epsilon_{n+1} \mid \mathcal{F}_n) = \phi X_n + \mathbb{E}(\epsilon_{n+1}) = \phi X_n$.

Remarque 38 *L'algorithme de Durbin-Levinson est bien adapté pour fournir des prévisions d'un processus AR(p).*

L'algorithme des innovations

Cet algorithme est applicable à tous les processus du second ordre qu'ils soient stationnaires ou non stationnaires.

Soit $\{X_t\}$ un processus de moyenne zero tel que $\mathbb{E}|X_t|^2 < \infty, \forall t$ et $\mathbb{E}(X_i X_j) = \kappa(i, j)$.

Posons $\hat{X}_n = \begin{cases} 0 & \text{si } n = 1 \\ P_{n-1} X_n & \text{si } n \geq 2 \end{cases}$ et soit $v_n = \mathbb{E}(\mathbb{X}_{n+1} - P_n X_{n+1})^2$.

Introduisons les **innovations** ou erreurs de prévision en une étape : $U_n = X_n - \hat{X}_n$. Alors

$$\underline{U}_n = A_n \underline{X}_n$$

où $\underline{U}_n = (U_1, \dots, U_n)'$ et $\underline{X}_n = (X_1, \dots, X_n)'$

$$\text{et } A_n = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ a_{11} & 1 & 0 & & 0 \\ a_{22} & a_{21} & \ddots & & \vdots \\ \vdots & \vdots & & 1 & 0 \\ a_{n-1,n-1} & a_{n-1,n-2} & \dots & a_{n-1,1} & 1 \end{pmatrix}. \text{ Si } \{X_t\} \text{ est stationnaire,}$$

alors $a_{ij} = -a_i$ ou a_i est donné dans (13) avec $h = 1$. A_n est non singulière.

$$\text{Son inverse est la matrice } C_n = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ \theta_{11} & 1 & 0 & & 0 \\ \theta_{22} & \theta_{21} & \ddots & & \vdots \\ \vdots & \vdots & & 1 & 0 \\ \theta_{n-1,n-1} & \theta_{n-1,n-2} & \dots & \theta_{n-1,1} & 1 \end{pmatrix}$$

Le vecteur des prédicteurs en une étape est $\hat{\underline{X}}_n = (X_1, P_1 X_2, \dots, P_{n-1} X_n)'$. Il peut s'écrire

$$\hat{\underline{X}}_n = \underline{X}_n - \underline{U}_n = C_n \underline{U}_n - \underline{U}_n = (C_n - I_n) \underline{U}_n = \Theta_n (\underline{X}_n - \hat{\underline{X}}_n)$$

$$\text{où } \Theta_n = \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 \\ \theta_{11} & 0 & 0 & & 0 \\ \theta_{22} & \theta_{21} & \ddots & & \vdots \\ \vdots & \vdots & & 0 & 0 \\ \theta_{n-1,n-1} & \theta_{n-1,n-2} & \cdots & \theta_{n-1,1} & 0 \end{pmatrix}. \text{ On déduit que } \hat{X}_{n+1} = \begin{cases} 0 \text{ si } n = 0 \\ \sum_{j=1}^n \theta_{nj}(X_{n+1-j} - \hat{X}_{n+1-j}) \text{ si } n \geq 1 \end{cases}$$

Les prédicteurs $\hat{X}_1, \hat{X}_2, \dots$ peuvent être calculés de façon récursive une fois les $\theta_{nj}, j = 1, \dots, n$ connus. L'algorithme suivant génère ces coefficients ainsi que les erreurs quadratiques moyennes $v_i = \mathbb{E}(X_{i+1} - \hat{X}_{i+1})^2$ partant de $\mathbb{E}(X_i X_j) = \kappa(i, j)$. Les coefficients $\theta_{nj}, j = 1, \dots, n$ se calculent à partir des équations :

L'algorithme des innovations

$$v_0 = \kappa(1, 1)$$

$$\theta_{n,n-k} = v_k^{-1}(\kappa(n+1, k+1) - \sum_{j=0}^{k-1} \theta_{k,k-j} \theta_{n,n-j} v_j), j = 0, 2, \dots, n$$

$$v_n = \kappa(n+1, n+1) - \sum_{j=0}^{n-1} \theta_{n,n-j}^2 v_j, n = 1, 2, \dots$$

Trouver d'abord v_0 , puis $\theta_{11}, v_1; \theta_{22}, \theta_{21}, v_2$; etc.

Exemple 39 En utilisant l'algorithme des innovations, calculer les prévisions du modèle $MA(1)$

$$X_t = \epsilon_t + \theta \epsilon_{t-1}, \{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2)$$

$$\kappa(i, j) = \mathbb{E}(X_i, X_j) = \mathbb{E}(\epsilon_i + \theta \epsilon_{i-1})(\epsilon_j + \theta \epsilon_{j-1}) = \begin{cases} (1 + \theta^2) \sigma_\epsilon^2 & \text{si } i = j \\ \theta \sigma_\epsilon^2 & \text{si } j = i + 1 \text{ ou } i = j + 1 \\ 0 & \text{si } |i - j| > 1 \end{cases}$$

$$v_0 = (1 + \theta^2) \sigma_\epsilon^2$$

$$\text{Pour } n = 1 \text{ et } k = 0 \text{ on a } \theta_{1,1} = v_0^{-1}(\kappa(1+1, 1) - \sum_{j=0}^{0-1} \theta_{k,k-j} \theta_{n,n-j} v_j) =$$

$$((1 + \theta^2) \sigma_\epsilon^2)^{-1} \theta \sigma_\epsilon^2 = \frac{\theta}{1 + \theta^2}$$

$$v_1 = \kappa(1+1, 1+1) - \sum_{j=0}^{1-1} \theta_{1,1-j}^2 v_j = (1 + \theta^2) \sigma_\epsilon^2 - \theta_{1,1}^2 v_0 = (1 + \theta^2) \sigma_\epsilon^2 -$$

$$\frac{\theta^2}{(1 + \theta^2)^2} (1 + \theta^2) \sigma_\epsilon^2 =$$

$$= v_0(1 - \theta_{1,1}^2) = (1 + \theta^2) \sigma_\epsilon^2 (1 - \frac{\theta^2}{(1 + \theta^2)^2}) = \sigma_\epsilon^2 (1 + \theta^2 - \theta^2 v_0^{-1} \sigma_\epsilon^2) \text{ etc.}$$

Appliquer à $\theta = -0.9$. Appliquer l'algorithme de Durbin-Levinson.

5 La décomposition de Wold

Considérons le processus stationnaire

$$\begin{aligned} X_t &= A \cos(\omega t) + B \sin(\omega t) \\ \omega &\in]0, \pi[\text{ } A \text{ et } B \text{ sont non corrélées centrées de variance } \sigma^2 \end{aligned}$$

Un simple calcul montre que $X_n = (2 \cos \omega) X_{n-1} - X_{n-2}$ et donc $X_n = P_{n-1} X_n$. On dit que $\{X_t\}$ est **déterministe**.

Le théorème de décomposition de Wold

Si $\{X_t\}$ est un processus stationnaire non déterministe, alors

- $X_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} + V_t$.
- $\psi_0 = 1, \sum_{i=0}^{\infty} \psi_i^2 < \infty$.
- $\{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2)$ et ϵ_t est non corrélé avec $V_s, \forall s, t$.
- $\{V_t\}$ est déterministe.