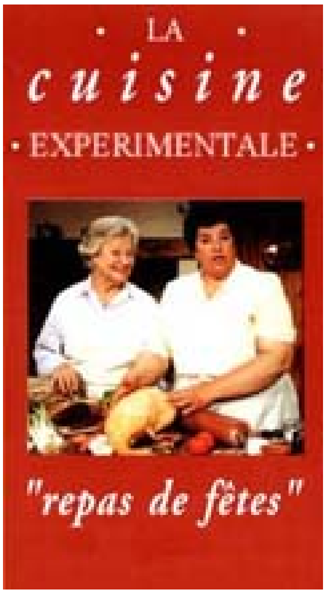
Séries Temporelles

un cadeau de



Librement inspiré du cours de Madame DOZ

2001-2002

Introduction

- Mesures à intervalle régulier : modélisation en temps discret $t \in \mathbb{Z}$, en temps continu $t \in \mathbb{R}$
- Modélisation linéaire
- Séries temporelles : $\rightarrow$ univariées : Une seule série

  $\rightarrow$ multivariées : un vecteur de séries $x_t = \begin{pmatrix} x_{1t} \\ \vdots \\ x_{nt} \end{pmatrix}$

Chapitre 1

Processus réels stationnaires (au second ordre)

Formalisme : On considère des observations x_t, réalisations d'une variable aléatoire X_t. On a donc $x_t = X_t(\omega)$, ω étant un état de la nature.

$(X_t)_{t\in\mathbb{Z}}$ suite de v.a.r. $\longrightarrow (x_t)_{t\in\mathbb{Z}}$ et $x_t = X_t(\omega) \quad \forall t \in \mathbb{Z}$

$(x_t)_{t\in\mathbb{Z}}$: trajectoire du processus $(X_t)_{t\in\mathbb{Z}}$

$(X_t)_{t\in\mathbb{Z}}$: processus stochastique

Remarque : Si $\mathbb{E}(X_t) = m_t$, on a une seule observation (x_t en l'occurence) pour estimer m_t. En revanche si pour tout $t \in \mathbb{Z}$, $\mathbb{E}(X_t) = m$, on peut estimer m par $m = \frac{1}{T}\sum_{t=1}^{T} X_t$. Il parait donc nécessaire de supposer que la suite X_t a certaines propriétés de régularité.

1.1 Processus stationnaires du second ordre

1.1.1 Définitions

Dans toute la suite on considèrera $(X_t)_{t\in\mathbb{Z}}$ avec $X_t \in \mathcal{L}^2(\Omega,\mathcal{A},P)$ pour tout $t \in \mathbb{Z}$

Définition 1.1.1 *$(X_t)_{t\in\mathbb{Z}}$ est un **processus stationnaire au sens strict** si :*
$\forall n \in \mathbb{N},\ \forall (t_1,\ldots,t_n),\ \forall h \in \mathbb{Z}$, la loi de $(X_{t_1},\ldots,X_{t_n})$ est identique à la loi de $(X_{t_1+h},\ldots,X_{t_n+h})$

Théorème 1.1.1 (Théorème de Kolmogorov) *$(X_t)_{t\in\mathbb{Z}}$ est un processus stationnaire au sens strict si et seulement si la loi de $(X_t)_{t\in\mathbb{Z}}$ est identique à la loi de $(Y_t)_{t\in\mathbb{Z}}$ où $Y_t = X_{t+h}$*

Définition 1.1.2 *$(X_t)_{t\in\mathbb{Z}}$ est un **processus stationnaire du second ordre** (ou un processus faiblement stationnaire) s'il vérifie :*

$$\forall t \in \mathbb{Z},\ \mathbb{E}(X_t) = m$$

$$\forall t \in \mathbb{Z},\ \mathbb{V}ar(X_t) = \sigma^2 = \gamma(0)$$

$$\forall t \in \mathbb{Z},\ \forall h \in \mathbb{Z},\ \mathbb{C}ov(X_t,X_{t+h}) = \gamma(h)\ \textit{(ne dépend que de h)}$$

*$\gamma(h)$ est l'**auto-covariance** d'ordre h de X_t.*

Remarque

- Si un processus est stationnaire au sens strict alors il est faiblement stationnaire;
- Dans la suite on considère le cas de la définition 1.1.2 (processus stationnaires du second ordre);
- Si $(X_t)_{t\in\mathbb{Z}}$ est un processus gaussien alors il y'a équivalence entre staionnarités faible et forte;
- $\mathbb{E}\begin{pmatrix} X_{t_1} \\ \vdots \\ X_{t_n} \end{pmatrix} = \begin{pmatrix} m \\ \vdots \\ m \end{pmatrix} \quad \mathbb{V}ar\begin{pmatrix} X_{t_1} \\ \vdots \\ X_{t_n} \end{pmatrix} = \begin{pmatrix} \gamma(0) & & \gamma(t_j - t_i) \\ & \ddots & \\ & & \gamma(0) \end{pmatrix}$

Exemples de processus stationnaires

- $(\varepsilon_t)_{t\in\mathbb{Z}}$: **bruit blanc (faible)** si et seulement si :

$$\mathbb{E}(\varepsilon_t) = 0 \quad \forall t \in \mathbb{Z}$$

$$\mathbb{V}ar(\varepsilon_t) = \sigma^2 \quad \forall t \in \mathbb{Z}$$

$$\mathbb{C}ov(\varepsilon_t,\epsilon_\tau) = 0 \text{ si } t \neq \tau$$

  $\varepsilon_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$

  Remarque : ε_t est un **bruit blanc fort** si et seulement si les ε_t sont i.i.d., $\mathbb{E}(\varepsilon_t) = 0$ et $\mathbb{V}ar(\varepsilon_t) = \sigma^2$.

- Processus MA(1) (moving average of order 1/*moyenne mobile d'ordre 1*)
  Soit $\varepsilon_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$
  Soit $(X_t)_{t\in\mathbb{Z}}$ défini par : $\forall t \in \mathbb{Z}$, $X_t = \varepsilon_t - \theta\varepsilon_{t-1}$
  Alors $(X_t)_{t\in\mathbb{Z}}$ est un processus stationnaire. On dit que $X_t \rightsquigarrow$ MA(1).

Exemples de processus non stationnaires :

- Marche aléatoire
  $\varepsilon_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$, $(X_t)_{t\in\mathbb{Z}}$ est une marche aléatoire (random walk) si et seulement si $X_t = X_{t-1} + \varepsilon_t$, $\forall t \geqslant 0$ et $\mathbb{C}ov(\varepsilon_t,X_{t-k}) = 0$, $\forall\ 0 < k \leqslant t$.

$$\mathbb{E}X_t = \mathbb{E}X_{t-1} \Rightarrow \mathbb{E}X_t = m\ \forall t \in \mathbb{Z}$$

$$\left.\begin{array}{rcl} X_t & = & X_{t-1} + \varepsilon_t \\ X_{t-1} & = & X_{t-2} + \varepsilon_{t-1} \\ & \vdots & \\ X_1 & = & X_0 + \varepsilon_1 \end{array}\right\} \Rightarrow X_t = X_0 + \sum_{s=1}^{t} \varepsilon_s$$

$$\begin{aligned}\Rightarrow \mathbb{V}ar X_t &= \mathbb{V}ar X_0 + 2\sum_{s=1}^{t} \mathbb{C}ov(\varepsilon_s, X_0) + \mathbb{V}ar(\sum_{s=1}^{t} \varepsilon_s) \\ &= \mathbb{V}ar(X_0) + t\sigma^2\end{aligned}$$

Le processus n'est pas stationnaire en variance.

– Processus stationnaire autour d'un trend déterministe
$X_t = a + bt + Y_t$ avec $(Y_t)_{t\in\mathbb{Z}}$ processus stationnaire
$\mathbb{E}X_t = a + bt$, le processus n'est pas stationnaire en espérance.

Définition 1.1.3 (Fonction d'auto-covariance) *Elle est définie comme :*

$$\gamma : \mathbb{Z} \mapsto \mathbb{R}$$

$$\gamma(h) = \mathbb{C}ov(X_t, X_{t-h})$$

γ est une fonction paire de type positif, c'est à dire :

$$\forall n \in \mathbb{N}, \forall (t_1, \ldots, t_n), \forall (a_1, \ldots, t_n) \in \mathbb{R}^n, \sum_{1\leqslant i,j\leqslant n} a_i a_j \gamma(t_i - t_j) \geqslant 0$$

Démonstration

Parité :

$$\begin{aligned}\gamma(h) &= \mathbb{C}ov(X_t, X_{t+h}) = \mathbb{C}ov(X_{t-h}, X_{(t-h)+h}) = \mathbb{C}ov(X_{t-h}, X_t) \\ &= \mathbb{C}ov(X_t, X_{t-h}) = \gamma(-h)\end{aligned}$$

Positivité :

$$\begin{aligned}\mathbb{V}ar(\sum a_i X_{t_i}) &= \mathbb{C}ov(\sum_i a_i X_{t_i}, \sum_j a_j X_{t_j}) \\ &= \sum_{i,j} a_i a_j \mathbb{C}ov(X_{t_i}, X_{t_j}) \\ &= \sum_{i,j} a_i a_j \gamma(t_i - t_j) \geqslant 0\end{aligned}$$

Définition 1.1.4 (Fonction d'auto-corrélation)

$$\forall h \in \mathbb{Z}, \quad \rho(h) = \frac{\gamma(h)}{\gamma(0)} = \mathbb{C}orr(X_t, X_{t+h})$$

ρ est une fonction paire, de type positif, à valeurs dans $]-1;1[$.

Démonstration

$$\mathbb{C}orr(X_t,X_{t+h}) = \frac{\mathbb{C}orr(X_t,X_{t+h})}{\sqrt{\mathbb{V}ar X_t \mathbb{V}ar X_{t+h}}} = \frac{\gamma(h)}{\gamma(0)}$$

Auto-corrélograme (théorique) de $(X_t)_{t\in\mathbb{Z}}$ C'est le graphe de :

$$\begin{array}{ccc} \mathbb{N} & \mapsto &]-1;1[\\ h & \rightarrow & \rho(h) \end{array}$$

1.1.2 Rappels sur $\mathcal{L}^2(\Omega,\mathcal{A},P)$

$\mathcal{L}^2(\Omega,\mathcal{A},P)$ est une espace de Hilbert pour $(X|Y) = \mathbb{E}XY$

$$X_n \underset{\mathcal{L}^2}{\longrightarrow} X \Longleftrightarrow \lim_{n\to+\infty} \|X_n - X\|_2 = 0$$

Si

$$\sum_{j\in\mathbb{Z}} \|a_j X_j\|_2 = \sum_{j\in\mathbb{Z}} |a_j| \|X_j\|_2 < +\infty$$

alors la série $\sum_{j\in\mathbb{Z}} a_j X_j$ est définie p.s. et :

$$\sum_{j=-p}^{q} a_j X_j \xrightarrow[p,q\to+\infty]{} \sum_{j\in\mathbb{Z}} a_j X_j$$

Théorème de projection sur un s.e.v. fermé H **de** $\mathcal{L}^2(\Omega,\mathcal{A},P)$

$$\forall X \in \mathcal{L}^2(\Omega,\mathcal{A},P),\ \exists! X^* \in H /\ \|X - X^*\|_2 = \min_{Y\in H} \|X - Y\|_2$$

$P_H(X) = X^*$ est caractérisé par $X^* \in H$ et $X - X^* \in H^\perp$.

Théorème des trois perpendiculaires Soit H un s.e.v. fermé de $\mathcal{L}^2(\Omega,\mathcal{A},P)$, G un s.e.v. fermé de H, alors :

$$\forall X \in \mathcal{L}^2(\Omega,\mathcal{A},P), \quad P_G(P_H(X)) = P_G(X)$$

1.2 Outils pour l'étude des processus stationnaires

1.2.1 Transformée d'un processus stationnaire par une moyenne mobile infinie

Définition 1.2.1 *Soient* $(X_t)_{t\in\mathbb{Z}}$ *un processus stationnaire,* $(a_j)_{j\in\mathbb{Z}}$ *une suite de réels tels que* $\sum_j |a_j| < +\infty$.

Alors $Y_t = \sum_{j\in\mathbb{Z}} a_j X_{t-j}$ *est défini (p.s.) pour tout* t *et :*

$$Y_t \in \mathcal{L}^2(\Omega,\mathcal{A},P)\ \forall t \in \mathbb{Z}$$

Par ailleurs (Y_t) *est stationnaire et vérifie :*

$$m_Y = m_X \left(\sum_{j\in\mathbb{Z}} a_j\right)$$

$$\forall h \in \mathbb{Z},\quad \gamma_Y(h) = \sum_{j,k} a_j a_k \gamma(h+k-j) = \sum_{j,k} a_j a_k \gamma(h+j-k)$$

On dit que $(Y_t)_{t\in\mathbb{Z}}$ *est la* ***transformée de***$(X_t)_{t\in\mathbb{Z}}$ ***par la moyenne mobile infinie associée aux*** $(a_j)_{j\in\mathbb{Z}}$.

Remarque Si $X_t = \varepsilon_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$ alors $Y_t = \sum_{j\in\mathbb{Z}} a_j \varepsilon_{t-j}$ et on dit que $Y_t \rightsquigarrow MA(\infty)$

Démonstration

$$\sum_j \|a_j X_{t-j}\|_2 = \sum_j |a_j| \|X_{t-j}\|_2 = \left(\sum_j |a_j|\right) (m_X^2 + \gamma_X(0))^{1/2} < +\infty$$

Y_t est donc défini p.s. et $Y_t \in \mathcal{L}^2(\Omega,\mathcal{A},P)$, on a alors :

$$\begin{aligned}
\mathbb{E}Y_t &= \int_\Omega Y_t dP = \int_\Omega \left(\sum_{j\in\mathbb{Z}} a_j X_{t-j}\right) dP = \sum_{j\in\mathbb{Z}} a_j \left(\int_\Omega X_{t-j} dP\right) \text{(Fubini)} \\
&= \mathbb{E}\left(\sum_{j\in\mathbb{Z}} a_j X_{t-j}\right) = \sum_{j\in\mathbb{Z}} a_j \mathbb{E}X_{t-j} = m_X \sum_j a_j
\end{aligned}$$

Enfin :

$$\begin{aligned}
\mathbb{C}ov(Y_t,Y_{t-h}) &= \mathbb{C}ov\left(\sum_{j\in\mathbb{Z}} a_j X_{t-j}, \sum_{k\in\mathbb{Z}} a_k X_{t-h-k}\right) \\
&= \sum_j \sum_k a_j a_k \underbrace{\mathbb{C}ov(X_{t-j},X_{t-h-k})}_{\gamma_X(h+k-j)}
\end{aligned}$$

1.2.2 Régression linéaire ou affine théorique sur un nombre fini de retards

Définition 1.2.2 *Soit* $(X_t)_{t\in\mathbb{Z}}$ *un processus stationnaire :*

- *(i) La* ***régression linéaire théorique*** *de* X_t *sur* $X_{t-1},\ldots,X_{t-p}$ *est la projection orthogonale dans* $\mathcal{L}^2(\Omega,\mathcal{A},P)$ *de* X_t *sur* $H = Vect(X_{t-1},\ldots,X_{t-p})$.
- *(ii) La* ***régression affine théorique*** *de* X_t *sur* $X_{t-1},\ldots,X_{t-p}$ *est la projection orthogonale dans* $\mathcal{L}^2(\Omega,\mathcal{A},P)$ *de* X_t *sur* $H^* = Vect(1,X_{t-1},\ldots,X_{t-p})$.

Propriété : *(i)* et *(ii)* coïncident si et seulement si $m_X = 0$.

Notation : On note géneralement $EL(X_t|X_{t-1},\ldots,X_{t-p})$ la régression affine théorique de X_t sur $X_{t-1},\ldots,X_{t-p}$.

Rappel : Calcul de la régression affine théorique *(ii)*
$H = Vect(1,X_{t-1},\ldots,X_{t-p})$ et $X_t^* = p_H(x_t)$ est caractérisé par $X_t^* \in H$ et $X_t - X_t^* \perp H$.

$$X_t^* \in H \Leftrightarrow \exists\, a_0,a_1,\ldots,a_p/\ X_t^* = a_0 + \sum_{j=1}^{p} a_j X_{t-j}$$

$$\begin{aligned}
X_t - X_t^* \perp H \quad &\Leftrightarrow \quad \begin{cases} (X_t - X_t^*|1) = 0 \\ (X_t - X_t^*|X_{t-j}) = 0 \quad \forall j = 1,\ldots,p \end{cases} \\
&\Leftrightarrow \quad \begin{cases} \mathbb{E}(X_t - X_t^*) = 0 \\ \mathbb{E}[(X_t - X_t^*)X_{t-j}] = 0 \quad \forall j = 1,\ldots,p \end{cases} \\
&\Leftrightarrow \quad \begin{cases} \mathbb{E}X_t = \mathbb{E}(a_0 + \sum_{j=1}^{p} a_j X_{t-j}) = a_0 + m_X \sum_{j=1}^{p} a_j \\ \mathbb{E}(X_t X_{t-j}) = \left[\left(a_0 + \sum_{k=1}^{p} a_k X_{t-k}\right) X_{t-j}\right] \quad \forall j = 1,\ldots,p \end{cases} \\
&\Leftrightarrow \quad \begin{cases} a_0 = m_X \left(1 - \sum_{j=1}^{p} a_j\right) \\ \mathbb{E}(X_t X_{t-j}) = \mathbb{E}\left[m_X\left(1 - \sum_{k=1}^{p} a_k\right) X_{t-j}\right] + \sum_{k=1}^{p} a_k \mathbb{E}(X_t X_{t-k}) \quad \forall j = 1,\ldots,p \end{cases} \\
&\Leftrightarrow \quad \begin{cases} a_0 = m_X \left(1 - \sum_{j=1}^{p} a_j\right) \\ \mathbb{E}(X_t X_{t-j}) = m_X^2 \left(1 - \sum_{k=1}^{p} a_k\right) + \sum_{k=1}^{p} a_k \mathbb{E}(X_t X_{t-k}) \quad \forall j = 1,\ldots,p \end{cases} \\
&\Leftrightarrow \quad \begin{cases} a_0 = m_X \left(1 - \sum_{j=1}^{p} a_j\right) \\ \mathbb{E}(X_t X_{t-j}) = m_X^2 + \sum_{k=1}^{p} a_k \left[\mathbb{E}(X_t X_{t-k}) - m_X^2\right] \quad \forall j = 1,\ldots,p \end{cases}
\end{aligned}$$

On a donc :

$$\forall j = 1,\ldots,p \quad \mathbb{E}(X_t X_{t-j}) - m_X^2 = \sum_{k=1}^{p} a_k \left[\mathbb{E}(X_t X_{t-j}) - m_X^2\right]$$

Soit encore :

$$\forall j = 1,\ldots,p \quad \mathbb{C}ov(X_t,X_{t-j}) = \sum_{k=1}^{p} a_k \mathbb{C}ov(X_t,X_{t-j})$$

$$\gamma(j) = \sum_{k=1}^{p} a_k \gamma(k-j)$$

Et :

$$a_0 = m_X \left(1 - \sum_{j=1}^{p} a_j\right)$$

$$\begin{pmatrix} 1 & \gamma(1) & \dots & \gamma(p-1) \\ \gamma(1) & 1 & \dots & \gamma(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(p-1) & \gamma(p-2) & \dots & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} \gamma(1) \\ \vdots \\ \gamma(p) \end{pmatrix}$$

$$\Leftrightarrow \boxed{\begin{pmatrix} 1 & \rho(1) & \dots & \rho(p-1) \\ \rho(1) & 1 & \dots & \rho(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(p-1) & \rho(p-2) & \dots & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} \rho(1) \\ \vdots \\ \rho(p) \end{pmatrix}}$$

Cette dernière matrice étant inversible si les X_t sont indépendants

Proposition 1.2.1 *(admise) Dans l'équation précédente, le coefficient a_p est égal à :*

$$\begin{aligned} a_p &= \mathbb{C}orr(X_t - EL(X_t|X_{t-1},\dots,X_{t-p+1}), X_{t-p} - EL(X_{t-p}|X_{t-1},\dots,X_{t-p+1})) \\ &= \frac{\mathbb{C}ov(X_t - EL(X_t|X_{t-1},\dots,X_{t-p+1}), X_{t-p} - EL(X_{t-p}|X_{t-1},\dots,X_{t-p+1}))}{[\mathbb{V}ar(X_t - EL(X_t|X_{t-1},\dots,X_{t-p+1}))\mathbb{V}ar(X_{t-p} - EL(X_{t-p}|X_{t-1},\dots,X_{t-p+1}))]^{1/2}} \end{aligned}$$

$$\begin{aligned} a_p = a_p^p \quad \text{où} \quad EL(X_t|X_{t-1},\dots,X_{t-p}) &= \sum_{j=1}^{p} a_j^p X_{t-j} \\ EL(X_t|X_{t-1},\dots,X_{t-p+1}) &= \sum_{j=1}^{p-1} a_j^{p-1} X_{t-j} \\ EL(X_{t-p}|X_{t-1},\dots,X_{t-p+1}) &= \sum_{j=1}^{p-1} \alpha_j^{p-1} X_{t-j} \end{aligned}$$

1.2.3 Régression linéaire théorique sur un nombre infini de retards

Définition 1.2.3 *Soit $(X_t)_{t\in\mathbb{Z}}$ un processus stationnaire :*

- *(i) La* ***régression linéaire théorique*** *de X_t sur $X_{t-1},\dots,X_{t-p},\dots$ est la projection orthogonale dans $\mathcal{L}^2(\Omega,\mathcal{A},P)$ de X_t sur $H = Vect(X_{t-1},\dots,X_{t-p},\dots)$.*
- *(ii) La* ***régression affine théorique*** *de X_t sur $X_{t-1},\dots,X_{t-p},\dots$ est la projection orthogonale dans $\mathcal{L}^2(\Omega,\mathcal{A},P)$ de X_t sur $H^* = Vect(1,X_{t-1},\dots,X_{t-p},\dots)$.*

Les deux notions coïncident si et seulement si $\mathbb{E}X_t = 0$, $\forall t$.

Remarque : $X_t^* = EL(X_t | X_{t-1}, \dots, X_{t-p})$

$$\begin{aligned} \|X_t - X_t^*\|_2 &= \min_{a_0,\dots,a_p} \left\| X_t - \left(a_0 + \sum_{j=1}^{p} a_j X_{t-j} \right) \right\|_2 \\ &= \min_{Y \in H} \|X_t - Y\|_2 \end{aligned}$$

Notations : On note aussi $\overline{\mathcal{L}}(\underline{X_{t-1}})$ l'espace $\overline{\mathcal{L}}(1, X_{t-1}, \dots, X_{t-k}, \dots)$ et :

$$\begin{aligned} EL(X_t | \underline{X_{t-1}}) &= EL(X_t | X_{t-1}, \dots, X_{t-k}, \dots) \\ (&= EL(X_t | 1, X_{t-1}, \dots, X_{t-k}, \dots)) \end{aligned}$$

la régression linéaire (ou affine) sur $\overline{\mathcal{L}}(\underline{X_{t-1}})$.

Propriété : $EL(X_t | \underline{X_{t-1}}) = \lim_{n \to +\infty} EL(X_t | X_{t-1}, \dots, X_{t-n})$ au sens de $\mathcal{L}^2$.

Théorème 1.2.1 *(Admis)*
Soient $(X_t)_{t \in \mathbb{Z}}$, un processus stationnaire $X_t^* = EL(X_t | \underline{X_{t-1}})$ la régression affine de X_t sur $\overline{\mathcal{L}}(1, X_{t-1}, \dots, X_{t-k}, \dots)$ et $\varepsilon_t = X_t - X_t^*$, alors :
- *$(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc*
- *$\mathbb{C}ov(\varepsilon_t, \varepsilon_{t-k}) = 0 \quad \forall k > 0$*

Définition 1.2.4 *Avec les notations du théorème ci-dessus :*
- *$(\varepsilon_t)_{t \in \mathbb{Z}}$ est le processus des innovations de $(X_t)_{t \in \mathbb{Z}}$*
- *ε_t est l'**innovation** de X_t*
- *X_t^* est la **prévision optimale** de X_t à la date $t-1$*

Remarque : $\varepsilon_t = X_t - X_t^* = X_t - EL(X_t | \underline{X_{t-1}})$ donc $\varepsilon_t \perp 1$ et $\varepsilon_t \perp X_{t-k}$, $\forall k > 0$, ce qui peut aussi s'interpréter comme :

$$\mathbb{E}(\varepsilon_t) = 0$$

$$\forall k > 0, \mathbb{E}(\varepsilon_t X_{t-k}) = \mathbb{C}ov(\varepsilon_t, X_{t-k}) = 0$$

Théorème 1.2.2 (De Wold) *Soient $(X_t)_{t \in \mathbb{Z}}$ un processus stationnaire et $(\varepsilon_t)_{t \in \mathbb{Z}}$ le processus des innovations correspondant :*

$$\exists (a_j)_{j \in \mathbb{Z}} / \sum_{j=0}^{+\infty} |a_j| < +\infty \ \text{et} \ X_t = m_X + \sum_{j=0}^{+\infty} a_j \varepsilon_{t-j}$$

1.2.4 Densité spectrale et auto-corrélations inverses

Proposition 1.2.2 *(et définition)*
Soit $(X_t)_{t \in \mathbb{Z}}$ un processus stationnaire de la forme :

$$X_t = m + \sum_{j=0}^{+\infty} a_j \varepsilon_{t-j}$$

où $(\varepsilon_t)_{t\in\mathbb{Z}} \rightsquigarrow \mathcal{BB}$ *et* $\sum_{j=0}^{+\infty} |a_j| < +\infty$, *alors :*
- $\displaystyle\sum_{h\in\mathbb{Z}} |\gamma_X(h)| < +\infty$
- $\forall\omega \in [-\pi;\pi]$, $\boxed{f_X(\omega) = \dfrac{1}{2\pi}\displaystyle\sum_{h\in\mathbb{Z}} \gamma_X(h)e^{i\omega h}}$

f_X *est la* ***densité spectrale de*** $(X_t)_{t\in\mathbb{Z}}$.

Démonstration :

$$\sum_{h\in\mathbb{Z}} |\gamma_X(h)| = \sum_{h\in\mathbb{Z}} |\sum_{j,k} a_j a_k \gamma_\varepsilon(h+j-k)|$$

Et on a :

$$\gamma_\varepsilon(h+j-k) = \begin{cases} 0 \text{ si } h+j-k \neq 0 \\ \sigma_\varepsilon^2 \text{ si } h+j-k = 0 \end{cases}$$

Donc :

$$\begin{aligned} \sum_{h\in\mathbb{Z}} |\gamma_X(h)| &= \sum_{h\in\mathbb{Z}} |\sigma_\varepsilon^2 \sum_j a_j a_{h+j}| \\ &\leqslant \sigma_\varepsilon^2 \sum_{h,j} |a_j||a_{h+j}| = \sigma_\varepsilon^2 \left(\sum_j a_j\right)^2 < +\infty \end{aligned}$$

Proposition 1.2.3

$$f_X(\omega) = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} \gamma_X(h)\cos(\omega h)$$

Démonstration

$$\begin{aligned} f_X(\omega) &= \frac{1}{2\pi}\left[\gamma_X(0) + \sum_{h>0} \gamma_X(h)e^{i\omega h} + \sum_{h<0} \gamma_X(-h)e^{i\omega h}\right] \\ &= \frac{1}{2\pi}\left[\gamma_X(0) + \sum_{h>0} \gamma_X(h)e^{i\omega h} + \sum_{h>0} \underbrace{\gamma_X(-h)}_{=\gamma_X(h)} e^{-i\omega h}\right] \\ &= \frac{1}{2\pi}\left[\gamma_X(0) + \sum_{h>0} \gamma_X(h) \underbrace{(e^{i\omega h} + e^{-i\omega h})}_{=2\cos(\omega h)}\right] \\ &= \frac{1}{2\pi}\left[\gamma_X(0) + \sum_{h\neq 0} \gamma_X(h)\cos(\omega h)\right] \\ &= \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} \gamma_X(h)\cos(\omega h) \end{aligned}$$

Remarque :

$$(\varepsilon_t)_{t\in\mathbb{Z}} \leadsto \mathcal{BB}(0,\sigma^2) \Rightarrow f_\varepsilon(\omega) = \frac{\sigma_\varepsilon}{2\pi}$$

Théorème 1.2.3 *(d'injectivité)*
Avec les notations précédentes :

$$\forall h \in \mathbb{Z}, \quad \gamma_X(h) = \int_{[-\pi;\pi]} f_X(\omega)e^{-i\omega h}d\omega = \int_{[-\pi;\pi]} f_X(\omega)\cos(\omega h)d\omega$$

Démonstration :

$$\begin{aligned}
\int_{[-\pi;\pi]} f_X(\omega)e^{-i\omega h}d\omega &= \frac{1}{2\pi}\int_{[-\pi;\pi]}\left(\sum_{k\in\mathbb{Z}}\gamma_X(k)e^{i\omega k}\right)e^{-i\omega h}d\omega \\
&= \frac{1}{2\pi}\sum_{k\in\mathbb{Z}}\gamma_X(k)\underbrace{\left(\int_{[-\pi;\pi]} e^{i\omega(k-h)}d\omega\right)}_{=\begin{cases} 0 \text{ si } k\neq h \\ 2\pi \text{ si } k=h\end{cases}} \quad \text{(d'après Fubini)} \\
&= \gamma_X(h)
\end{aligned}$$

Proposition 1.2.4 *Soient :*

$$(X_t)_{t\in\mathbb{Z}}/ \quad X_t = \sum_{j\in\mathbb{Z}} a_j\varepsilon_{t-j} \ \textit{et}\ \varepsilon_t \leadsto \mathcal{BB}, \quad \sum_j |a_j| < +\infty$$

$$Y_t = \sum_{k\in\mathbb{Z}} b_k X_{t-k} \ \textit{avec}\ \sum_k |b_k| < +\infty$$

Alors :

- $Y_t = \sum_{k\in\mathbb{Z}} c_k\varepsilon_{t-k}$
- $f_Y(\omega) = f_X(\omega)|\sum_{k\in\mathbb{Z}} b_k e^{i\omega k}|^2$

Démonstration :

$$\begin{aligned}
Y_t = \sum_{k\in\mathbb{Z}} b_k X_{t-k} &= \sum_{k\in\mathbb{Z}} b_k\left(\sum_{j\in\mathbb{Z}} a_j\varepsilon_{t-k-j}\right) \\
&= \sum_{j,k\in\mathbb{Z}} a_j b_k \varepsilon_{t-(k+j)} \\
&= \sum_{j,h\in\mathbb{Z}} a_j b_{h-j}\varepsilon_{t-h} \\
&= \sum_{h\in\mathbb{Z}}\underbrace{\left(\sum_{j\in\mathbb{Z}} a_j b_{h-j}\right)}_{c_h}\varepsilon_{t-h}
\end{aligned}$$

Et :

$$\begin{aligned}
f_Y(\omega) &= \frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\gamma_Y(h)e^{i\omega h}\\
&= \frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\left(\sum_{j,k\in\mathbb{Z}} b_j b_k \gamma_X(h+j-k)\right)e^{i\omega h}\\
&= \frac{1}{2\pi}\sum_{h,j,k\in\mathbb{Z}} b_j b_k \gamma_X(h+j-k)e^{i\omega(h+j-k)}e^{-i\omega j}e^{i\omega k}\\
&= \frac{1}{2\pi}\left(\sum_{l\in\mathbb{Z}}\gamma_X(l)e^{i\omega l}\right)\left(\sum_{j\in\mathbb{Z}} b_j e^{i\omega j}\right)\left(\sum_{k\in\mathbb{Z}} b_k e^{-i\omega k}\right)\\
&= f_X(\omega)|\sum_{k\in\mathbb{Z}} b_k e^{i\omega k}|^2
\end{aligned}$$

Définition 1.2.5 (Autocorrélations inverses) *Soit :*

$$(X_t)_{t\in\mathbb{Z}}/\quad X_t = \sum_{j\in\mathbb{Z}} a_j\varepsilon_{t-j} \text{ et } \varepsilon_t \rightsquigarrow \mathcal{BB},\quad \sum_j |a_j| < +\infty$$

On suppose $\omega \to \frac{1}{f_X(\omega)}e^{-i\omega h}$ *intégrable sur* $[-\pi;\pi]$.

On appelle ***auto-covariance inverse*** *d'ordre h de* $(X_t)_{t\in\mathbb{Z}}$ *:*

$$\gamma_X^i = \int_{[-\pi;\pi]}\frac{1}{f_X(\omega)}e^{-i\omega h}d\omega$$

*L'**autocorrélation inverse** d'ordre h est alors définie comme :* $\rho_X^i(h) = \dfrac{\gamma_X^i(h)}{\gamma_X^i(0)}$

1.2.5 Estimateurs associés et lois limites

On considère un processus stationnaire $(X_t)_{t\in\mathbb{Z}}$ tel que pour tout $t\in\mathbb{Z}$, $\mathbb{E}X_t = m$. On cherche à estimer les grandeurs associées $\gamma_X(h)$, $\rho(h) = \frac{\gamma_X(h)}{\gamma_X(0)}$, $r(h) = a_h^h$ [1], $f_X(\omega)$ et $\rho_X^i(h)$ et ceci sachant qu'on observe $X_1,\ldots,X_T$.

On prend :

- $\hat{m} = \dfrac{1}{T}\sum_{t=1}^{T} X_t = \bar{X}_T$ moyenne empirique,
- $\hat{\gamma}_X(h) = \dfrac{1}{T-h}\sum_{t=h+1}^{T}(X_t-\bar{X}_T)(X_{t-h}-\bar{X}_T)$ auto-covariance empirique d'ordre h (estimation acceptable si h n'est pas trop grand),
- $\hat{\rho}_X(h) = \dfrac{\hat{\gamma}_X(h)}{\hat{\gamma}_X(0)}$,
- $\hat{r}(h) = \hat{a}_h^h$ dans la régression empirique (m.c.o.) de x_t sur $1, x_{t-1},\ldots,x_{t-h}$,

1. cf page 8

– $\hat{f}_X(\omega) = \frac{1}{2\pi}\sum_{h=-H}^{H}\hat{\gamma}_X(h)e^{i\omega h}$, le problème ici étant que l'on voudrait un H suffisamment grand mais prendre un H trop grand est risqué pour l'estimation de $\hat{\gamma}_X(h)$,
– On ne prend pas $\hat{\rho}_i = \frac{\hat{\gamma}^i(h)}{\hat{\gamma}^i(0)}$ où $\hat{\gamma}^i(h) = \int_{[-\pi;\pi]}\frac{1}{\hat{f}_X(\omega)}e^{-i\omega h}d\omega$, il existe d'autres façons de l'obtenir.

Proposition 1.2.5 *Tous les estimateurs présentés ci-dessus sont convergents (loi des grands nombres).*

Proposition 1.2.6 *Si $X_t = m + \sum_{j=0}^{+\infty} a_j\varepsilon_{t-j}$ où $\mathbb{E}(\varepsilon_t^4) = \eta < +\infty$, alors tous ces estimateurs sont asymptotiquement gaussiens (jointement) :*
– $\sqrt{T}(\bar{X}_T - m) \xrightarrow{L} \mathcal{N}(0, \sum_{h\in\mathbb{Z}}\gamma_X(h))$

$$\sqrt{T}\begin{pmatrix}\hat{\gamma}(0)-\gamma(0)\\ \vdots \\ \hat{\gamma}(h)-\gamma(h)\end{pmatrix} \xrightarrow{L} \mathcal{N}(0,\Omega_h)$$

– $\forall h,\ \sqrt{T}\begin{pmatrix}\hat{\rho}(0)-\rho(0)\\ \vdots \\ \hat{\rho}(h)-\rho(h)\end{pmatrix} \xrightarrow{L} \mathcal{N}(0,\mathcal{W}_h)$

$$\sqrt{T}\begin{pmatrix}\hat{r}(0)-r(0)\\ \vdots \\ \hat{r}(h)-r(h)\end{pmatrix} \xrightarrow{L} \mathcal{N}(0,\Sigma_h)$$

Ω_h, $\mathcal{W}_h$ et Σ_h étant calculables.

Remarque : Les auto-corrélogrammes (direct, partiel et inverse) associés aux valeurs estimées sont appelés auto-corrélogrammes empiriques.

1.3 Polynômes retard et avance

1.3.1 Définition et propriétés

Définition 1.3.1 *L'opérateur retard L (lag) ou B (backward) est défini comme étant :*

$$L\ :\ (X_t)_{t\in\mathbb{Z}} \to (Y_t)_{t\in\mathbb{Z}} \text{ où } Y_t = X_{t-1}$$

On note : $LX_t = X_{t-1}$

De la même façon, l'opérateur avance F (forward) correspond à $FX_t = X_{t+1}$

Propriété : $L^k = \underbrace{L\circ\cdots\circ L}_{k fois}$ vérifie $L^kX_t = X_{t-k}$.

Notation : $L^0 = Id$ est noté $L^0 = 1$ ($L^0 X_t = X_t$).

Définition 1.3.2 *Soit P un polynôme, $P(z) = \sum_{k=0}^{p} a_k z^k$, $a_k \in \mathbb{R}$, on lui associe le polynôme retard $P(L)$ défini comme suit :*

$$P(L) = \sum_{k=0}^{p} a_k L^k$$

Et :

$$P(L)X_t = \left(\sum_{k=0}^{p} a_k L^k\right) X_t = \sum_{k=0}^{p} a_k X_{t-k}$$

De façon similaire on obtient le polynôme avance $P(F)$:

$$P(F)X_t = \left(\sum_{k=0}^{p} a_k F^k\right) X_t = \sum_{k=0}^{p} a_k X_{t+k}$$

Séries en L (ou en F) (polynômes de degré infini)

$A(z) = \sum_{k=0}^{+\infty} a_k z^k$ et $Y_t = A(L)X_t = \sum_{k=0}^{+\infty} a_k X_{t-k}$. $(Y_t)_{t\in\mathbb{Z}}$ est bien un processus stationnaire car $\sum_{k=0}^{+\infty} |a_k| < +\infty$

Propriétés : On suppose donc $\sum_{k=0}^{+\infty} |a_k| < +\infty$ et $\sum_{k=0}^{+\infty} |b_k| < +\infty$ et :

$$A(L) = \sum_{k=0}^{+\infty} a_k L^k \qquad B(L) = \sum_{k=0}^{+\infty} b_k L^k$$

Alors :

- *(i)* $\forall \alpha \in \mathbb{R}, \quad \alpha A(L) = (\alpha A)(L) = \alpha \left(\sum_{k=0}^{+\infty} a_k L^k\right) = \sum_{k=0}^{+\infty} (\alpha a_k) L^k$
- *(ii)* $A(L) + B(L) = (A+B)(L) = \left(\sum_{k=0}^{+\infty} a_k L^k\right) + \left(\sum_{k=0}^{+\infty} b_k L^k\right) = \sum_{k=0}^{+\infty} (a_k + b_k) L^k$
- *(iii)* $A(L) \circ B(L) = (AB)(L) = B(L) \circ A(L)$ avec $(AB)(L) = (BA)(L) = C(L) = \sum_{k=0}^{+\infty} c_k L^k$ et $c_k = \sum_{j=0}^{+\infty} a_j b_{k-j}$

1.3.2 Inversibilité des polynômes en L (vrais polynômes)

On suppose $P(L) = \sum_{k=0}^{p} a_k L^k$ et $Y_t = P(L)X_t$ et on désire savoir si X_t peut s'exprimer en fonction de Y_t ($X_t = P(L)^{-1} Y_t$).

On peut décomposer notre polynôme de la façon suivante :

$$\begin{aligned} P(z) &= \prod_{i=1}^{p} (z - z_i) \qquad z_i \in \mathbb{C} \\ &= \prod_{i=1}^{p} (-z_i) \prod_{i=1}^{p} \left(1 - \frac{z}{z_i}\right) \\ &= \alpha \prod_{i=1}^{p} (1 - \lambda_i z) \qquad \lambda_i = \frac{1}{z_i} \end{aligned}$$

Finalement $P(L) = \alpha \prod_{i=1}^{p}(1 - \lambda_i L)$

1.3.2.1 Inversibilité de $1 - \lambda L$

Proposition 1.3.1

(i) Si $|\lambda| < 1$, $1 - \lambda L$ *est inversible et* $(1 - \lambda L)^{-1} = \sum_{k=0}^{+\infty} \lambda^k L^k$

(ii) Si $|\lambda| > 1$, $1 - \lambda L$ *est inversible et* $(1 - \lambda L)^{-1} = -\sum_{k=1}^{+\infty} \frac{1}{\lambda^k} F^k$

(iii) Si $|\lambda| < 1$, $1 - \lambda L$ *n'est pas inversible*

Démonstration :

(i) Si $|\lambda| < 1$ alors $(1 - \lambda L)^{-1} = \sum_{k=0}^{+\infty} |\lambda^k| = \frac{1}{1 - |\lambda|} < +\infty$, donc $A(L) = \sum_{k=0}^{+\infty} \lambda^k L^k$ est bien défini.

On a ainsi $(1 - \lambda L)A(L) = C(L) = \sum_{j=0}^{+\infty} c_j L^j$ et $c_j = \sum_{k=0}^{+\infty} a_k b_{k-j}$ avec $b_0 = 1$, $b_1 = -\lambda$, $b_k = 0$ si $k > 1$ et $a_k = \lambda^k$. On trouve $c_0 = 1$ et $c_j = 0$ si $j \neq 0$, soit encore $C(L) = 1$. On en déduit que $(1 - \lambda L)$ est inversible et $(1 - \lambda L)^{-1} = A(L)$.

On pouvait aussi montrer ce résultat en écrivant :

$$(1 - \lambda L)A(L) = \lim_{k \to +\infty} (1 - \lambda L) \left(\sum_{j=0}^{k} \lambda^j L^j \right) = \lim_{k \to +\infty} 1 - \lambda^{k+1} L^{k+1} = 1$$

(ii) Si $|\lambda| > 1$ alors $1 - \lambda L = -\lambda \left(L - \frac{1}{\lambda} \right) = -\lambda L \left(1 - \frac{F}{\lambda} \right)$. On a alors :

$$(\lambda L)^{-1} = \frac{1}{\lambda} F \text{ et } \left(1 - \frac{F}{\lambda} \right)^{-1} = \sum_{k=0}^{+\infty} \frac{1}{\lambda^k} F^k \text{ car } |\lambda| > 1$$

En combinant ces deux résultats :

$$\begin{aligned}
(1 - \lambda L) &= (-\lambda L)^{-1} \left(1 - \frac{F}{\lambda} \right)^{-1} \\
&= -\frac{1}{\lambda} F \left(\sum_{k=0}^{+\infty} \frac{1}{\lambda^k} F^k \right) \\
&= -\sum_{k=1}^{+\infty} \frac{1}{\lambda^k} F^k \\
&= -\sum_{k=-\infty}^{-1} \lambda^k L^k
\end{aligned}$$

Remarque : Dans ce cas, $X_t = (1-\lambda L)^{-1} Y_t = -\sum_{k=1}^{+\infty} \frac{1}{\lambda^k} Y_{t+k}$

(iii) Si $\lambda = 1$, $1-L$ n'est pas inversible. Supposons $Y_t = (1-L)X_t = X_t - X_{t-1}$. On a vu que (cf. page 3, l'exemple où $Y_t = \varepsilon_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$) si $(Y_t)_{t\in\mathbb{Z}}$ est staitionnaire alors $(X_t)_{t\in\mathbb{Z}}$ ne l'est pas.

- On n'a pas $X_t = \sum_{k\in\mathbb{Z}} a_k Y_{t-k}$ avec $\sum_{k\in\mathbb{Z}} |a_k| < +\infty$;
- Il n'existe pas $A(L) = \sum_{k\in\mathbb{Z}} a_k L^k$, $\sum_{k\in\mathbb{Z}} |a_k| < +\infty$ tel que $(1-L)A(L) = 1$.

Remarque : On peut le voir à la main :

$$(1-L)A(L) = 1 \Rightarrow |a_k| = |a_{k-1}| \text{ et donc ne tend pas vers } 0$$

Dans ces conditions $\sum_{k\in\mathbb{Z}} |a_k| = +\infty$.

1.3.2.2 Inversibilité des polynômes en L

Soit ϕ un polynôme de degré p à coefficients réels :

$$\phi(z) = 1 + \varphi_1 z + \cdots + \varphi_p z^p$$

$$\phi(L) = 1 + \varphi_1 L + \cdots + \varphi_p L^p \quad (\phi(0) = 1)$$

ϕ possède p racines $(z_1, \ldots, z_p)$ dans $\mathbb{C}$, on peut donc le décomposer en :

$$\begin{aligned} \phi(z) = \varphi_p \prod_{j=1}^{p} (z - z_j) &= \varphi_p \prod_{j=1}^{p} (-z_j) \prod_{j=1}^{p} \left(1 - \frac{z}{z_j}\right) \\ &= \prod_{j=1}^{p} (1 - \lambda_j z) \quad \lambda_j = \frac{1}{z_j} \end{aligned}$$

Par conséquent :

$$\phi(L) = \prod_{j=1}^{p} (1 - \lambda_j z)$$

Proposition 1.3.2 *Avec les notations précédentes :*
(i) Si $|\lambda_j| \neq 1$ pour tout $j = 1, \ldots, p$, alors $\phi(L)$ est inversible
(ii) Si $|\lambda_j| < 1$ pour tout $j = 1, \ldots, p$, alors $\phi(L)$ est inversible et :

$$\phi(L)^{-1} = \sum_{k=0}^{+\infty} a_k L^k \quad a_0 = 1,\ a_k \in \mathbb{R} \quad \sum_{k=0}^{+\infty} a_k < +\infty$$

Démonstration :

(i) $\forall j$, $(1-\lambda_j L)^{-1}$ est bien défini, de la forme $\sum_{k\in\mathbb{Z}} a_{j,k}L^k$ et $\phi(L)^{-1} = \prod_{j=1}^{p}(1-\lambda_j L)^{-1}$ est donc aussi défini.

Mais $\phi(L)^{-1}$ peut contenir des termes en L^k, $k>0$ qui sont des termes concernant le futur et donc peu utilisables en pratique.

(ii) Si $|\lambda_j|<1$ pour tout j alors $(1-\lambda_j L)^{-1} = \sum_{k=0}^{+\infty}\lambda_j^k L^k$ et :

$$\phi(L)^{-1} = \prod_{j=1}^{p}(1-\lambda_j L)^{-1} = \sum_{k=0}^{+\infty} a_k L^k \quad \sum_{k=0}^{+\infty} a_k < +\infty$$

Par ailleurs :

$$\phi(z) = \prod_{j=1}^{p}(1-\lambda_j z) \quad \phi(z)\phi(z)^{-1} = 1 \Leftrightarrow \prod_{j=1}^{p}(1-\lambda_j z)\left(\sum_{k=0}^{+\infty} a_k z^k\right) = 1$$

Donc :

$$\phi(0)\phi(0)^{-1} = 1\times a_0 = 1 \Rightarrow a_0 = 1$$

S'il existe j tel que $\lambda_j \in \mathbb{C}\backslash\mathbb{R}$ alors $\phi(L) = (1-\lambda_j)(1-\overline{\lambda}_j)P(L)$ et :

$$\begin{aligned}(1-\lambda_j)^{-1}(1-\overline{\lambda}_j)^{-1} &= \left(\sum_{k=0}^{+\infty}\lambda_j^k L^k\right)\left(\sum_{k=0}^{+\infty}\overline{\lambda}_j^k L^k\right)\\ &= \sum_{k=0}^{+\infty}\alpha_k L^k \quad \alpha_k\in\mathbb{R},\, \alpha_0 = 1, \sum_{k=0}^{+\infty}|\alpha_k| < +\infty\end{aligned}$$

Remarque : $|\lambda_j|<1 \Leftrightarrow |z_j| = \dfrac{1}{\lambda_j} > 1$

Proposition 1.3.3 *ϕ un polynôme :*
- *(i) Si $\phi(z)$ a toutes ses racines de module différent de 1, alors $\phi(L)$ est inversible*
- *(ii) Si $\phi(z)$ a toutes ses racines de module strictement supérieur à 1, alors $\phi(L)$ est inversible et :*

$$\phi(L)^{-1} = \sum_{k=0}^{+\infty} a_k L^k \quad a_k\in\mathbb{R},\, a_0 = 1, \sum_{k=0}^{+\infty}|a_k| < +\infty$$

1.3.2.3 Méthodes pratiques de calcul de $\phi(L)^{-1}$

On se place dans le cadre défini précédemment où :

$$\phi(L) = \prod_{j=1}^{p}(1-\lambda_j L)$$

Alors :

$$\phi(L)^{-1} = \prod_{j=1}^{p} \left(\sum_{k=0}^{+\infty} \lambda_j^k L^k \right)$$

On peut utiliser directement cette méthode de calcul pour p petit ($p = 1{,}2$) mais elle s'avère fastidieuse en général.

Identification : On écrit que :

$$\phi(L) \left(\sum_{k=0}^{+\infty} a_k L^k \right) = (1 + \varphi_1 L + \cdots + \varphi_p L^p) \left(\sum_{k=0}^{+\infty} a_k L^k \right) = 1$$

Les a_k sont obtenus par récurrence puis identification.

Décomposition en éléments simples :

$$\phi(L)^{-1} = \prod_{j=1}^{p} \frac{1}{1 - \lambda_j L}$$

On décompose cette fraction rationnelle en éléments simples.

Division selon les puissances croissantes de 1 par $\phi(z)$

Exemple : $1 = \phi(z) Q_r(z) + z^{r+1} R_r(z)$

$$\lim_{r \to +\infty} Q_r(z) = \phi^{-1}(z)$$

Chapitre 2

Processus ARMA et ARIMA

Les ARMA sont des processus stationnaires et les ARIMA des processus non stationnaires integrés.

2.1 Processus $AR(p)$ (auto-régressif d'ordre p)

2.1.1 Définition et représentation canonique

Définition 2.1.1 *$(X_t)_{t\in\mathbb{Z}}$ est un processus $AR(p)$ si :*
- *Il est stationnaire*
- *Il vérifie une équation $X_t = \mu + \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} + \varepsilon_t$ avec $\varphi_p \neq 0$ et $\varepsilon_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$*

On note $\phi(L)X_t = \mu + \varepsilon_t$ où $\phi(L) = 1 - (\varphi_1 L + \cdots + \varphi_p L^p)$

Exemple : $X_t \rightsquigarrow AR(1)$

$$X_t = \mu + \varphi X_{t-1} + \varepsilon_t \qquad (1 - \varphi L)X_t = \mu + \varepsilon_t$$

En général on pose d'emblée $|\varphi| < 1$.

Remarque : Il existe des processus non stationnaires (en espérance) qui vérifient la même équation.

Exemple : Soient :
- $(X_t)_{t\in\mathbb{Z}}$ stationnaire $\phi(L)X_t = \mu + \varepsilon_t$,
- $(Y_t)_{t\in\mathbb{Z}}$ déterministe i.e. $\phi(L)Y_t = 0$.

On posera pour simplifier $\phi(L) = 1 - L$ et donc $Y_t = \varphi Y_{t-1} = \cdots = \varphi^t Y_0$.
Soit par ailleurs $Z_t = X_t + Y_t$, alors :

$$\forall t,\, \phi(L)Z_t = \varepsilon_t \qquad \mathbb{E}Z_t = m_X + \varphi^t \mathbb{E}Y_0 \neq cte$$

Proposition 2.1.1 *Soit $X_t \rightsquigarrow AR(p)$ tel que $\phi(L)X_t = \mu + \varepsilon_t$, alors :*

$$\mathbb{E}X_t = \frac{\mu}{\phi(1)} = \frac{\mu}{1 - (\varphi_1 + \cdots + \varphi_p)}$$

Démonstration :

$$
\begin{aligned}
X_t &= \mu + \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} + \varepsilon_t \\
\mathbb{E}X_t &= \mu + \varphi_1 \mathbb{E}X_{t-1} + \cdots + \varphi_p \mathbb{E}X_{t-p} + \mathbb{E}\varepsilon_t \\
m &= \mu + \varphi_1 m + \cdots + \varphi_p m \\
m &= \frac{\mu}{1-(\varphi_1 + \cdots + \varphi_p)}
\end{aligned}
$$

On retrouve bien le résultat annoncé.

Proposition 2.1.2 *Si $X_t \rightsquigarrow AR(p)$ est tel que $\phi(L)X_t = \mu + \varepsilon_t$ et si l'on pose $Y_t = X_t - m$ ($m = \mathbb{E}X_t$), on a alors :*

$$\phi(L)Y_t = \varepsilon_t \text{ et } \mathbb{E}Y_t = 0$$

Démonstration :

$\mathbb{E}X_t = m\dfrac{\mu}{\phi(1)}$ et :

$$\phi(L)(X_t - m) = \phi(L)X_t - \phi(L)m$$

Or $Lm = m$ et par conséquent :

$$\phi(L)m = (1 - (\varphi_1 + \cdots + \varphi_p))m = \phi(1)m$$

Finalement :

$$\phi(L)(X_t - m) = \phi(L)X_t - \mu = \varepsilon_t$$

Ecriture de $MA(\infty)$ quand les racines de ϕ sont de module strictement supérieur à 1

On suppose $\phi(L)X_t = \mu + \varepsilon_t$ où $\phi(L) = 1 - (\varphi_1 L + \cdots + \varphi_p L)$ et aussi que $|z| \leqslant 1 \Rightarrow \phi(z) \neq 0$.

Proposition 2.1.3 *Sous les hypothèses précédentes, $(X_t)_{t\in\mathbb{Z}}$ admet une représentation $MA(\infty)$ i.e. :*

$$X_t = m + \sum_{k=0}^{+\infty} a_k \varepsilon_{t-k} \quad a_0 = 1,\, a_k \in \mathbb{R},\, \sum_{k=0}^{+\infty} |a_k| < +\infty$$

On sait que $\phi(L)(X_t - m) = \varepsilon_t$, donc $X_t - m = \phi(L)^{-1}(\varepsilon_t)$.

Proposition 2.1.4 *Sous les hypothèses précédentes :*

- *(i)* $\overline{\mathcal{L}}(\underline{X_t}) = \overline{\mathcal{L}}(\underline{\varepsilon_t})$
- *(ii)* ε_t *est l'innovation de* X_t

Rappel :

$$\overline{\mathcal{L}}(\underline{X_t}) = \overline{\mathcal{L}}(1,X_t,X_{t-1},\dots,X_{t-p},\dots)$$
$$\overline{\mathcal{L}}(\underline{\varepsilon_t}) = \overline{\mathcal{L}}(1,\varepsilon_t,\varepsilon_{t-1},\dots,\varepsilon_{t-p},\dots)$$

Démonstration :

(i) $X_t = \mu + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t$

On a vu que $X_t = \eta + \sum_{k=0}^{+\infty} a_t \varepsilon_{t-k}$

$$\Rightarrow X_t \in \overline{\mathcal{L}}(\underline{\varepsilon_t}) = \overline{\mathcal{L}}(1,\varepsilon_t,\varepsilon_{t-1},\dots,\varepsilon_{t-k},\dots)$$

Donc $\forall k \geqslant 0$, $X_{t-k} \subset \overline{\mathcal{L}}(\underline{\varepsilon_{t-k}}) \subset \overline{\mathcal{L}}(\underline{\varepsilon_t})$

$$\begin{aligned}
&\Rightarrow \quad \mathcal{L}(1,X_t,X_{t-1},\dots,X_{t-k},\dots) \subset \overline{\mathcal{L}}(\underline{\varepsilon_t}) \\
&\Rightarrow \quad \mathcal{L}(\underline{X_t}) \subset \overline{\mathcal{L}}(\underline{\varepsilon_t}) \\
&\Rightarrow \quad \overline{\mathcal{L}}(\underline{X_t}) \subset \overline{\mathcal{L}}(\underline{\varepsilon_t})
\end{aligned}$$

De la même façon et comme :

$$\varepsilon_t = X_t - (\mu + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p})$$

On obtient l'inclusion réciproque et finalement $\overline{\mathcal{L}}(\underline{X_t}) = \overline{\mathcal{L}}(\underline{\varepsilon_t})$.

(ii) L'innovation de X_t vaut, par définition, $X_t - X_t^*$, or :

$$\begin{aligned}
X_t^* &= EL(X_t|\underline{X_{t-1}}) = EL(X_t|1,X_{t-1},\dots,X_{t-k},\dots) \\
&= EL(\underbrace{\mu + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p}}_{\in \overline{\mathcal{L}}(\underline{X_{t-1}})} + \varepsilon_t|\underline{X_{t-1}}) \\
&= \mu + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + EL(\varepsilon_t|\underline{X_{t-1}})
\end{aligned}$$

Comme $\overline{\mathcal{L}}(\underline{X_{t-1}}) = \overline{\mathcal{L}}(\underline{\varepsilon_{t-1}})$, on a :

$$EL(\varepsilon_t|\underline{X_{t-1}}) = EL(\varepsilon_t|\underline{\varepsilon_{t-1}}) = 0 \text{ car } \varepsilon_t \rightsquigarrow \mathcal{BB}$$

Finalement $X_t^* = \mu + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p}$ et $X_t - X_t^* = \varepsilon_t$, ε_t est bien l'innovation de X_t.

Définition 2.1.2 *Soient $X_t \rightsquigarrow AR(p)$ et ϕ un polynôme vérifiant :*
- $\phi(L)X_t = \mu + \varepsilon_t$
- $|z| \leqslant 1 \Rightarrow \phi(z) \neq 0$

On dit que la représentation $\phi(L)X_t = \mu + \varepsilon_t$ est la ***représentation canonique*** *de $(X_t)_{t\in\mathbb{Z}}$.*

Cas où ϕ admet des racines de module inférieur à 1

Remarque : Si $(X_t)_{t\in\mathbb{Z}}$ est supposé statique alors ϕ n'a pas de racines de module égal à 1.

On suppose $\phi(L)X_t = \mu + \varepsilon_t$ avec :

$$\phi(L) = \prod_{j=1}^{p}(1-\lambda_j L) = \left[\prod_{j/\ |\lambda_j|<1}(1-\lambda_j L)\right]\left[\prod_{j/\ |\lambda_j|>1}(1-\lambda_j L)\right]$$

On sait que $X_t = m + \sum_{k\in\mathbb{Z}} a_k\varepsilon_{t-k}$ n'est pas utilisable (on n'aura pas $\overline{\mathcal{L}}(\underline{X_t}) = \overline{\mathcal{L}}(\underline{\varepsilon_t})$, ε_t n'est pas l'innovation).

Pour obtenir la représentation canonique il faut changer le polynôme ϕ et le bruit blanc.

$$\phi(z) = \left[\prod_{j/\ |\lambda_j|<1}(1-\lambda_j z)\right]\left[\prod_{j/\ |\lambda_j|>1}(1-\lambda_j z)\right]$$

On pose :

$$\phi^*(z) = \left[\prod_{j/\ |\lambda_j|<1}(1-\lambda_j z)\right]\left[\prod_{j/\ |\lambda_j|>1}(1-\frac{z}{\lambda_j})\right]$$

ϕ^* a toutes ses racines de module strictement supérieur à 1.

On définit le processus $(\eta_t)_{t\in\mathbb{Z}}$ tel que $\eta_t = \phi^*(L)X_t$. On montre alors que $\eta_t \rightsquigarrow \mathcal{BB}(0,\sigma_\eta^2)$ et on peut calculer $f_\eta(\omega)$:

$$f_\eta(\omega) = f_X(\omega)|\phi^*(e^{i\omega})|^2$$

Comme $\phi(L)X_t = \varepsilon_t$, on a aussi :

$$f_X(\omega)|\phi(e^{i\omega})|^2 = f_\varepsilon(\omega) = \frac{\sigma_\varepsilon^2}{2\pi}$$

Ceci nous mène à :

$$\begin{aligned}
f_\eta(\omega) &= \frac{\sigma_\varepsilon^2}{2\pi}\frac{1}{|\phi(e^{i\omega})|^2}|\phi^*(e^{i\omega})|^2 \\
&= \frac{\sigma_\varepsilon^2}{2\pi}\frac{\left[\prod_{j/\ |\lambda_j|<1}|1-\lambda_j e^{i\omega}|^2\right]\left[\prod_{j/\ |\lambda_j|>1}|1-\frac{e^{i\omega}}{\lambda_j}|^2\right]}{\left[\prod_{j/\ |\lambda_j|<1}|1-\lambda_j e^{i\omega}|^2\right]\left[\prod_{j/\ |\lambda_j|>1}|1-\lambda_j e^{i\omega}|^2\right]} \\
&= \frac{\sigma_\varepsilon^2}{2\pi}\prod_{j,\ |\lambda_j|>1}\frac{1}{|\lambda_j|^2}\frac{|\lambda_j - e^{i\omega}|^2}{|1-\lambda_j e^{i\omega}|^2}
\end{aligned}$$

Or :

$$\prod_{j/\ |\lambda_j|>1}\frac{|\lambda_j - e^{i\omega}|^2}{|1-\lambda_j e^{i\omega}|^2} = 1$$

En effet:

– Si $\lambda_j \in \mathbb{R}$, $|1-\lambda_j e^{i\omega}|^2 = |1-\lambda_j e^{-i\omega}|^2 = |1-\lambda_j e^{i\omega}|^2$

– Si $\lambda_j \in \mathbb{R}\backslash\mathbb{C}$, $\dfrac{|\lambda_j - e^{i\omega}|^2|\overline{\lambda_j} - e^{i\omega}|^2}{|1-\lambda_j e^{i\omega}|^2|1-\overline{\lambda_j}e^{i\omega}|^2} = 1$, $\overline{\lambda_j}$ étant aussi une racine de ϕ puisque celui-ci est à coefficients réels.

On a donc $f_\eta(\omega) = \dfrac{\sigma_\varepsilon^2\alpha}{2\pi} = \dfrac{\sigma_\eta^2}{2\pi}$ avec $\alpha = \displaystyle\prod_{j,\ |\lambda_j|>1} \frac{1}{|\lambda_j|^2} < 1$ et finalement $\eta_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$.

La représentation $\phi^*(L)X_t = \eta_t$ est la représentation canonique de $(X_t)_{t\in\mathbb{Z}}$ (car ϕ^* a toutes ses racines de module strictement supérieur à 1) et η_t est l'innovation de X_t.

2.1.2 Propriétés des processus $AR(p)$

On suppose que $\phi(L)X_t = \mu + \varepsilon_t$ et que les racines de ϕ sont de module strictement supérieur à 1. On peut se ramener ensuite à $\mu = 0$ par centrage : $\phi(L)(X_t - m) = \varepsilon_t$.

On considère donc le cas où $\phi(L)X_t = \varepsilon_t$ (et $\mathbb{E}X_t = 0$).

Auto-covariance et auto-corrélations de $(X_t)_{t\in\mathbb{Z}}$ $\gamma(h) = \mathbb{C}ov(X_t,X_{t-h}) = \mathbb{E}(X_t X_{t-h})$ pour $h \geqslant 0$ (car $m_X = 0$), et :

$$X_t = \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} + \varepsilon_t$$

Donc :

$$X_t^2 = \varphi_1 X_t X_{t-1} + \cdots + \varphi_p X_t X_{t-p} + X_t\varepsilon_t$$
$$\gamma(0) = \varphi_1\gamma(1) + \cdots + \varphi_p\gamma(p) + \mathbb{E}(X_t\varepsilon_t)$$

Or :

$$\mathbb{E}(X_t\varepsilon_t) = \underbrace{\mathbb{E}[(\varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p})\varepsilon_t]}_{=0 \text{ car } \varepsilon_t \perp \underline{\mathcal{L}}(\overline{X_{t-1}})} + \mathbb{E}(\varepsilon_t^2)$$

D'où :

$$\gamma(0) = \varphi_1\gamma(1) + \cdots + \varphi_p\gamma(p) + \sigma_\varepsilon^2$$

Si h¿0, on procède de la même façon:

$$X_t X_{t-h} = \varphi_1 X_{t-1}X_{t-h} + \cdots + \varphi_p X_{t-p}X_{t-h} + \varepsilon_t X_{t-h}$$
$$\Rightarrow \gamma(h) = \varphi_1\gamma(h-1) + \cdots + \varphi_p\gamma(h-p) + \underbrace{\mathbb{E}(\varepsilon_t X_{t-h})}_{=0 \text{ car } \varepsilon_t \perp X_{t-h}}$$
$$\Rightarrow \rho(h) = \varphi_1\rho(h-1) + \cdots + \varphi_p\rho(h-p) \quad \forall h \geqslant 0$$

Ces équations, sont appelées **équations de Yule-Walker**. Pour $h > 0$, les $\gamma(h)$ et les $\rho(h)$ vérifient une relation de récurrence d'ordre p et :

$$1 = \varphi_1\rho(1) + \cdots + \varphi_p\rho(p) + \frac{\sigma_\varepsilon^2}{\gamma(0)}$$

$$\Rightarrow \gamma(0) = \sigma_\varepsilon^2 \frac{1}{1 - (\varphi_1\rho(1) + \cdots + \varphi_p\rho(p))}$$

Les équations de Yule-Walker pour $h = 1,\ldots,p$ peuvent s'écrire :

$$\begin{pmatrix} 1 & \rho(1) & \cdots & \rho(p-1) \\ \rho(1) & 1 & \rho(1) & \rho(p-2) \\ \vdots & & \ddots & \vdots \\ \rho(p-1) & & \cdots & 1 \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_p \end{pmatrix} = \begin{pmatrix} \rho(1) \\ \vdots \\ \rho(p) \end{pmatrix}$$

Si $\rho(1),\ldots,\rho(p)$ connus elles permettent d'obtenir $\varphi_1,\ldots,\varphi_p$. En particulier elles donneront une estimation préliminaire de $\hat{\varphi}_1,\ldots,\hat{\varphi}_p$ en fonction de $\hat{\rho}_T(1),\ldots,\hat{\rho}_T(p)$.

$$\left\{ \begin{array}{rcl} \rho(1) & = & \varphi_1 + \varphi_2\rho(1) + \cdots + \varphi_p\rho(p-1) \\ & \cdots & \\ \rho(p) & = & \varphi_1\rho(p-1) + \cdots + \varphi_{p-1}\rho(1) + \varphi_p \end{array} \right.$$

$$\Leftrightarrow \left\{ \begin{array}{rcl} \varphi_1 & = & (1-\varphi_2)\rho(1) - \cdots - \varphi_p\rho(p-1) \\ & \cdots & \\ \varphi_p & = & \rho(p) - \varphi_1\rho(p-1) - \cdots + \varphi_{p-1}\rho(1) \end{array} \right.$$

On peut donc aussi obtenir $\rho(1),\ldots,\rho(p)$ en fonction de $\varphi_1,\ldots,\varphi_p$.

Proposition 2.1.5 *Les $\rho(h)$, $h > 0$ sont solutions d'une équation de récurrence d'ordre p dont le polynôme caractéristique est $z^p\phi(\frac{1}{z})$ (ce à quoi il faut ajouter les équations de Yule-Walker).*

Les $\rho(h)$ décroissent exponentiellement vers 0.

Démonstration :

$$\forall h > 0,\ \rho(h) - \varphi_1\rho(h-1) - \cdots - \varphi_p\rho(h-p) = 0$$

Le polynôme caractéristique de cette relation de récurrence est :

$$\begin{array}{rl} & z^p - \varphi_1 z^{p-1} - \cdots - \varphi_{p-1}z - \varphi_p \\ = & z^p\left(1 - \frac{\varphi_1}{z} - \cdots - \frac{\varphi_{p-1}}{z^{p-1}} - \frac{\varphi_p}{z^p}\right) \\ = & z^p\phi(\frac{1}{z}) \end{array}$$

Avec $\phi(L)X_t = \varepsilon_t$ et $\phi(L) = 1 - \varphi_1 L - \cdots \varphi_p L^p$. Les racines du polynôme caractéristique sont les $\lambda_i = \frac{1}{z_i}$ (les z_i étant les racines de ϕ) avec $|\lambda_i| < 1$. La forme générale de la solution est, si $z_1,\ldots,z_n$ sont des racines distinctes de ϕ de multiplicités respectives $m_1,\ldots,m_n$:

$$\rho(h) = \sum_{i=1}^{n}\sum_{k=0}^{m_i-1}\alpha_{ik}\lambda_i^k h^k$$

$|\lambda_i| < 1$ donc $\rho(h)$ décroit vers 0 exponentiellement avec h.

2.1.3 Autocorrélations partielles et inverse d'un processus *AR(p)*

Proposition 2.1.6 *Si $(X_t)_{t\in\mathbb{Z}} \rightsquigarrow AR(p)$ et si $\phi(L)X_t = \mu+\varepsilon_t$ est sa représentation canonique, alors :*
- *$r(h) = 0$ si $h > p$.*
- *$r(p) \neq 0$*

Démonstration :

$r(h)$ est le coefficient de X_{t-h} dans $EL(X_t|X_{t-1},\ldots,X_{t-h})$ et :

$$X_t = \mu + \underbrace{\varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p}}_{\in\mathcal{L}(1,X_t,\ldots,X_{t-p})\subset\mathcal{L}(1,X_t,\ldots,X_{t-h})} + \varepsilon_t$$

$$\begin{aligned}\Rightarrow EL(X_t|X_{t-1},\ldots,X_{t-h}) &= \mu + \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} + EL(\varepsilon_t|X_{t-1},\ldots,X_{t-h})\\ &= \mu + \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} + 0\end{aligned}$$

Si $h > p$, le coefficient de X_{t-h} est 0.
Si $h = p$, le coefficient de X_{t-p} est $\varphi_p \neq 0$.

Proposition 2.1.7 *Soit $(X_t)_{t\in\mathbb{Z}} \rightsquigarrow AR(p)$:*
- *Si $h > p$, $\rho_i(h) = 0$*
- *Si $h = p$, $\rho_i(h) \neq 0$*

Démonstration :

$\rho_i(h) = \dfrac{\gamma_i(h)}{\gamma_i(0)}$ où :

$$\gamma_i(h) = \int_{-\pi}^{\pi} \frac{1}{f_X(\omega)} e^{i\omega h} d\omega$$

$\phi(L)X_t = \varepsilon_t$ (en remplaçant éventuellement X_t par $X_t - m$) :

$$f_X(\omega)|\phi(e^{i\omega})|^2 = f_\varepsilon(\omega) = \frac{\sigma_\varepsilon^2}{2\pi}$$

$$\Rightarrow f_X(\omega) = \frac{\sigma_\varepsilon^2}{2\pi}\frac{1}{|\phi(e^{i\omega})|^2}$$

Et par conséquent :

$$\frac{1}{f_X(\omega)} = \frac{2\pi}{\sigma_\varepsilon^2}|\phi(e^{i\omega})|^2$$

$$\phi(z) = 1 - \varphi_1 z - \cdots - \varphi_p z^p = \sum_{k=0}^{p} \psi_k z^k$$

Avec $\psi_0 = 1$ et $\psi_k = -\varphi_k$, $k > 0$,

$$\begin{aligned}
\frac{1}{f_X(\omega)} &= \frac{2\pi}{\sigma_\varepsilon^2}\left(\sum_{k=0}^{p}\psi_k e^{i\omega k}\right)\left(\sum_{k=0}^{p}\psi_k e^{-i\omega k}\right) \\
&= \frac{2\pi}{\sigma_\varepsilon^2}\sum_{0\leqslant k,l\leqslant p}\psi_k\psi_l e^{i\omega(k-l)}
\end{aligned}$$

Ainsi :

$$\gamma_i(h) = \frac{2\pi}{\sigma_\varepsilon^2}\sum_{0\leqslant k,l\leqslant p}\psi_k\psi_l\underbrace{\int_{-\pi}^{\pi}e^{i\omega(k-l+h)}d\omega}_{=0 \text{ sauf si } k-l+h=0}$$

Or $k - l \in [\![-p;p]\!]$ donc si $h > p$, $\gamma_i(h) = 0$. En revanche si $h = p$:

$$\begin{aligned}
\textstyle\int_{-\pi}^{\pi} e^{i\omega(k-l+h)}d\omega \quad &\Leftrightarrow \quad p = l - k \\
&\Leftrightarrow \quad \begin{cases} l = p \\ k = 0 \end{cases}
\end{aligned}$$

Donc :

$$\gamma_i(p) = \frac{4\pi^2}{\sigma_\varepsilon^2}\psi_0\psi_p = -\frac{4\pi^2}{\sigma_\varepsilon^2}\varphi_p \neq 0$$

2.2 Processus $MA(q)$

2.2.1 Définition et représentation canonique

Définition 2.2.1 *$(X_t)_{t\in\mathbb{Z}} \rightsquigarrow MA(q)$ s'il existe $\varepsilon_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$ et $\theta_1,\ldots,\theta_q$ tels que :*

$$X_t = m + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}$$

Remarques :

1. $(X_t)_{t\in\mathbb{Z}}$ est nécessairement stationnaire.
2. $X_t = m + \theta(L)\varepsilon_t$ où $\theta(L) = 1 - \theta_1 L - \cdots - \theta_p$.

Proposition 2.2.1 $\mathbb{E}X_t = m$

Remarque : $X_t - m = \theta(L)\varepsilon_t$, par centrage $\mathbb{E}X_t = 0$

2.2.1.1 Ecriture $AR(\infty)$ quand $|\theta_i| < 1$

On suppose que $\theta(z) = 1 - \theta_1 z - \cdots - \theta_q z^q = \prod_{i=1}^{q}(1-\lambda_i z)$ avec $\lambda_i = \dfrac{1}{z_i}$ et $\theta(z_i) = 0$, on a alors :

$$\forall i,\, |z_i| > 1 \Rightarrow |\lambda_i| < 1 \Rightarrow (1-\lambda_i L)^{-1} = \sum_{k=0}^{+\infty}\lambda_i^k L^k$$

$\theta(L)$ est donc inversible et $\theta(L)^{-1} = \sum_{k=0}^{+\infty} a_k L^k$ avec $a_0 = 1$ et $\sum |a_k| < +\infty$. Il s'en suit que :

$$X_t - m = \theta(L)\varepsilon_t \iff \theta(L)^{-1}(X_t - m) = \varepsilon_t \iff \theta(L)^{-1}X_t - \frac{m}{\theta(1)} = \varepsilon_t$$

Soit encore :

$$\sum_{k=0}^{+\infty} a_k X_{t-k} - \mu = \varepsilon_t \text{ où } \mu = \frac{m}{\theta(1)}$$

Par ailleurs :

$$X_t = m + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q} \Rightarrow X_t \in \mathcal{L}(1,\varepsilon_t,\ldots,\varepsilon_{t-q}) \subset \overline{\mathcal{L}}(\underline{\varepsilon_t})$$

Ce qui nous amène à (cf. cas AR page 21 pour plus de détails):

$$\left.\begin{array}{c} \overline{\mathcal{L}}(\underline{X_t}) \subset \overline{\mathcal{L}}(\underline{\varepsilon_t}) \\ \varepsilon_t = \sum_{k=0}^{+\infty} a_k X_{t-k} - \mu \Rightarrow \overline{\mathcal{L}}(\underline{\varepsilon_t}) \subset \overline{\mathcal{L}}(\underline{X_t}) \end{array}\right\} \Rightarrow \overline{\mathcal{L}}(\underline{X_t}) = \overline{\mathcal{L}}(\underline{\varepsilon_t})$$

Proposition 2.2.2 *Sous les hypothèses précédentes, ε_t est l'innovation de X_t.*

Démonstration

$$\begin{aligned} X_t^* &= EL(X_t|\underline{X_{t-1}}) \\ &= EL(m + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}|\underline{X_{t-1}}) \\ &= EL(m + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}|\underline{\varepsilon_{t-1}}) \\ &= m + 0 - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q} \\ &= X_t - \varepsilon_t \end{aligned}$$

Donc $X_t - X_t^* = \varepsilon_t$, X_t^* est bien l'innovation de X_t.

2.2.1.2 Cas où $|\theta_i| \neq 1$ et $\exists i / |\theta_i| > 1$

On suppose qu'on s'est ramené à $X_t = \theta(L)\varepsilon_t$ par centrage :

$$X_t = \left[\prod_{i/\ |\lambda_i|<1} (1 - \lambda_i L)\right]\left[\prod_{i/\ |\lambda_i|>1} (1 - \lambda_i L)\right]\varepsilon_t$$

Comme précédemment, on définit (cf. page 22) :

$$\theta^*(L) = \left[\prod_{i/\ |\lambda_i|<1} (1 - \lambda_i L)\right]\left[\prod_{i/\ |\lambda_i|>1} (1 - \frac{1}{\lambda_i} L)\right]$$

2.2.2 Propriétés des processus $MA(q)$

Proposition 2.2.3 *Si* $X_t = m + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}$*, alors :*

$$\gamma(h) = \begin{cases} 0 \text{ si } |h| > q \\ -\theta_q\sigma_\varepsilon^2 \neq 0 \text{ si } |h| = q \\ \sigma_\varepsilon^2(-\theta_h + \displaystyle\sum_{i=h+1}^{q} \theta_i\theta_{i-h}) \text{ si } 1 \leqslant |h| < q \\ \sigma_\varepsilon^2(1 + \displaystyle\sum_{i=1}^{q} \theta_i^2) \text{ si } h = 0 \end{cases}$$

En particulier, $\rho(h) = 0$ *si* $|h| > q$ *et* $\rho(q) \neq 0$.

Démonstration :

$X_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}$ après centrage.

- Si $h = 0$:
$$\gamma(0) = \mathbb{V}arX_t = \sigma_\varepsilon^2(1 + \theta_1^2 + \cdots + \theta_q^2) \neq 0$$
Car $\mathbb{C}ov(\varepsilon_{t-j},\varepsilon_{t-k}) = 0$ si $j \neq k$.
- Si $h > q$:
$$\begin{aligned} \gamma(h) &= \mathbb{C}ov(X_t,X_{t-h}) \\ &= \mathbb{C}ov[\underbrace{(\varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q})}_{\varepsilon_{t-j},\ j\in[\![0,q]\!]}, \underbrace{(\varepsilon_{t-h} - \theta_1\varepsilon_{t-h-1} - \cdots - \theta_q\varepsilon_{t-h-q})}_{\varepsilon_{t-k},\ k\in[\![h,q+h]\!]}] \\ &\Rightarrow t - j \neq t - k \\ &\Rightarrow \gamma(h) = 0 \end{aligned}$$
- Si $|h| = q$
$$\begin{aligned} \gamma(q) &= \mathbb{C}ov[(\varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}),(\varepsilon_{t-q} - \theta_1\varepsilon_{t-q-1} - \cdots - \theta_q\varepsilon_{t-2q})] \\ &= -\theta_q\sigma_\varepsilon^2 \end{aligned}$$
- Si $1 \leqslant |h| < q$:
$$\begin{aligned} \gamma(h) &= \mathbb{C}ov[(\varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}),(\varepsilon_{t-h} - \theta_1\varepsilon_{t-h-1} - \cdots - \theta_q\varepsilon_{t-h-q})] \\ &= -\sum_{i=1}^{q}\theta_i Cov[\varepsilon_{t-i},\varepsilon_{t-h} - \sum_{k=1}^{q}\theta_k\varepsilon_{t-h-k}] \quad \text{car } \mathbb{C}ov(\varepsilon_t,\varepsilon_{t-i}) = 0\ \forall i > 0 \\ &= -\theta_h\sigma_\varepsilon^2 + \sum_{i=1}^{q}\sum_{k=1}^{q}\theta_i\theta_k \underbrace{\mathbb{C}ov(\varepsilon_{t-i},\varepsilon_{t-h-k})}_{=0 \text{ si } i\neq h+k} \\ &= -\theta_h\sigma_\varepsilon^2 + \left(\sum_{i=h+1}^{q}\theta_i\theta_{i-h}\right)\sigma_\varepsilon^2 \end{aligned}$$

Remarque : On n'a pas de résultat particulier pour les autocorrélations partielles.

Proposition 2.2.4 $\rho_i(h)$ *décroit exponentiellement avec* h.

Démonstration :

$\rho_i(h) = \dfrac{\gamma_i(h)}{\gamma_i(0)}$ avec, on le rappelle $\gamma_i(h) = \displaystyle\int_{-\pi}^{\pi} \frac{1}{f_X(\omega)} e^{i\omega h} d\omega$ et :

$$\begin{aligned} X_t = \theta(L)\varepsilon_t \quad &\Rightarrow \quad f_X(\omega) = \frac{\sigma_\varepsilon^2}{2\pi}|\theta(e^{i\omega})|^2 \\ &\Rightarrow \quad \frac{1}{f_X(\omega)} = \frac{2\pi}{\sigma_\varepsilon^2|\theta(e^{i\omega})|^2} \end{aligned}$$

Soit $(y_t)_{t\in\mathbb{Z}}$ un processus tel que $\theta(L)y_t = \eta_t$ et $y_t \rightsquigarrow AR(q)$:

$$\frac{\sigma_\eta^2}{2\pi} = f_y(\omega)|\theta(e^{i\omega})|^2$$

Donc :

$$f_y(\omega) = \frac{\sigma_\eta^2}{2\pi}\frac{1}{|\theta(e^{i\omega})|^2}$$

On a ainsi :

$$f_y(\omega) = \frac{1}{f_X(\omega)} \iff \frac{2\pi}{\sigma_\varepsilon^2} = \frac{\sigma_\eta^2}{2\pi} \iff \sigma_\eta^2 = \frac{4\pi^2}{\sigma_\varepsilon^2}$$

Les autocorrélations inverses d'un processus $MA(q)$ ont les mêmes propriétés que les autocorrélations d'un $AR(q)$:

	$AR(p)$	$MA(q)$
$\rho(h)$	décroit exponentiellement vers 0 avec h	0 si $\|h\| > q$ et non nul si $h = q$
$r(h)$	0 si $h > p$ et non nul si $h = p$	-
$\rho_i(h)$	0 si $h > p$ et non nul si $h = p$	décroit exponentiellement vers 0 avec h

2.3 Processus $ARMA$

2.3.1 Définition et représentation canonique minimale

Définition 2.3.1 *un processus stationnaire $(X_t)_{t\in\mathbb{Z}}$ admet une représentation $ARMA(p,q)$ canonique minimale s'il vérifie une équation :*

$$\phi(L)X_t = \mu + \theta(L)\varepsilon_t$$

où :

1. $\varepsilon_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$
2. $\phi(L) = 1 - \varphi_1 L - \cdots - \varphi_p L^p$, *avec* $\varphi_p \neq 0$
3. $\theta(L) = 1 - \theta_1 L - \cdots - \theta_q L^q$, *avec* $\theta_q \neq 0$
4. *ϕ et θ ont toutes leurs racines de module strictement supérieur à 1 (représentation canonique).*
5. *ϕ et θ n'ont pas de racines communes (représentation minimale).*

Remarques : Si ϕ et θ ont des racines de module strictement supérieur à 1 on peut se ramener au cas idéal (cf. la section précédente).

Si ϕ et θ avaient une racine commune :

$$\phi(L) = (1 - \lambda L)\varphi_0(L) \qquad \theta(L) = (1 - \lambda L)\theta_0(L)$$

alors :

$$\varphi_0(L)X_t = \frac{\mu}{1-\lambda} + \theta_0(L)\varepsilon_t \;\Rightarrow\; X_t \rightsquigarrow ARMA(p-1, q-1)$$

Proposition 2.3.1

1. $\mathbb{E}X_t = \dfrac{\mu}{\phi(1)} = m$
2. $\phi(L)(X_t - m) = \theta(L)\varepsilon_t$

Par centrage on peut donc se ramener au cas où $\mu = 0$.

Démonstration :

1- On a :

$$\mathbb{E}(X_t - \varphi_1 X_{t-1} - \cdots \varphi_p X_{t-p}) = \mathbb{E}(\mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q})$$

et comme (X_t) est supposé stationnaire :

$$m(1 - \varphi_1 - \cdots - \varphi_p) = \mu + 0 \;\Rightarrow\; m = \frac{\mu}{\phi(1)}$$

2- Ensuite :

$$\begin{aligned} \phi(L)X_t &= \phi(1)m + \theta(L)\varepsilon_t \\ &= \phi(L)m + \theta(L)\varepsilon_t \end{aligned}$$

Donc $\phi(L)(X_t - m) = \theta(L)\varepsilon_t$.

Remarque : Si Y_t est déterministe tel que $\phi(L)Y_t = 0$:

→ Y_t est non stationnaire en espérance.

→ $Z_t = X_t + Y_t$ vérifie $\phi(L)Z_t = \mu + \theta(L)\varepsilon_t$ mais (Z_t) est non stationnaire en espérance.

Proposition 2.3.2 *Sous les hypothèses précédentes :*

(i) (X_t) *admet une représentation* $AR(\infty)$, $\displaystyle\sum_{k=0}^{+\infty} a_k X_{t-k} = \mu_0 + \varepsilon_t$ *où* $a_0 = 1$ *et* $\sum_k |a_k| < +\infty$

(ii) (X_t) *admet une représentation* $MA(\infty)$, $X_t = m + \displaystyle\sum_{k=0}^{+\infty} b_k\varepsilon_{t-k}$ *où* $b_0 = 1$ *et* $\sum_k |b_k| < +\infty$

(iii) $\overline{\mathcal{L}}(\underline{X_t}) = \overline{\mathcal{L}}(\underline{\varepsilon_t})$

(iv) ε_t *est l'innovation de* X_t.

Démonstration :

(i) On sait que $\phi(L)(X_t - m) = \theta(L)\varepsilon_t$, celà nous permet d'écrire :

$$\underbrace{\theta(L)^{-1}\phi(L)}_{A(L)}(X_t - m) = \varepsilon_t$$

$$\Rightarrow A(L)X_t - A(1)m = \varepsilon_t$$

Et ce avec $A(1)m = \dfrac{\phi(1)}{\theta(1)} = \dfrac{\mu}{\theta(1)}$.

(ii) De la même façon $\phi(L)X_t = \mu + \theta(L)\varepsilon_t$ amène :

$$X_t = \frac{\mu}{\phi(1)} + \underbrace{\phi(L)^{-1}\theta(L)}_{B(L)=\sum b_k L^k}\varepsilon_t$$

(iii) Etant donné que X_t est de la forme $AR(\infty)$:

$$\forall t, \varepsilon_t \in \overline{\mathcal{L}}(\underline{X_t}) \;\Rightarrow\; \mathcal{L}(\underline{\varepsilon_t}) \subset \overline{\mathcal{L}}(\underline{X_t}) \;\Rightarrow\; \overline{\mathcal{L}}(\underline{\varepsilon_t}) \subset \overline{\mathcal{L}}(\underline{X_t})$$

Par un raisonnement identique et tenant compte du fait que X_t est également de la forme $MA(\infty)$ on obtient :

$$\overline{\mathcal{L}}(\underline{X_t}) \subset \overline{\mathcal{L}}(\underline{\varepsilon_t})$$

Les deux résultats nous permettent alors de dire que :

$$\overline{\mathcal{L}}(\underline{X_t}) = \overline{\mathcal{L}}(\underline{\varepsilon_t})$$

(iv) Calculons l'innovation de X_t :

$$\begin{aligned}
X_t - X_t^* &= X_t - EL(X_t|\underline{X_{t-1}}) \\
&= X_t - EL(-\sum_{1=0}^{+\infty} a_k X_{t-k} + \mu_0 + \varepsilon_t|\underline{X_{t-1}}) \\
&= X_t + \sum_{1=0}^{+\infty} a_k X_{t-k} - \mu_0 - \underbrace{EL(\varepsilon_t|\underline{\varepsilon_{t-1}})}_{=0} \\
&= \varepsilon_t
\end{aligned}$$

Remarques :

- $AR(p) = ARMA(p,0)$
- $MA(q) = ARMA(0,q)$
- $ARMA(p,q) \equiv AR(\infty) \# AR(P)$ si P grand
  $\equiv MA(\infty) \# MA(Q)$ si Q grand
  Souvent l'un des paramètres (p ou q) est petit alors que l'autre est grand. Avec l'approximation précédente on a alors moins de paramètres à estimer.
- En vertu du théorème de Wold, $X_t = m + B(L)\varepsilon_t$, où ε_t est le processus des innovations, si de plus $X_t \rightsquigarrow ARMA(p,q)$ alors $B(L) = \dfrac{\theta(L)}{\phi(L)}$

2.3.2 Propriétés

On considére un processus $ARMA(p,q)$ tel que :
- $\phi(L)X_t = \theta(L)\varepsilon_t$, éventuellement après centrage.
- $\phi(L) = 1 - \varphi_1 L - \cdots - \varphi_p L^p$
- $\theta(L) = 1 - \theta_1 L - \cdots - \theta_q L^q$

Proposition 2.3.3 (Autocovariances et auto-corrélations)
Pour $k > q$, les $\gamma(h)$ et les $\rho(h)$ vérifient les équations de récurrence d'ordre p :

$$\gamma(k) - \varphi_1\gamma(k-1) - \cdots - \varphi_p\gamma(k-p) = 0$$

$$\rho(k) - \varphi_1\rho(k-1) - \cdots - \varphi_p\rho(k-p) = 0$$

Conséquences : Les $\gamma(h)$ et les $\rho(h)$ vérifient une équation de récurrence dont le polynôme caractéristique est $z^{p+1}\phi(\frac{1}{z})$.

Elles décroissent doc vers 0 exponentiellement avec k, pour $k > q$. Les conditions initiales sont $\gamma(q),\gamma(q-1),\ldots,\gamma(q-p+1)$ et $\rho(q),\rho(q-1),\ldots,\rho(q-p+1)$.

Démonstration de la proposition :

$X_t = \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}$ et par conséquent :

$$\begin{aligned}
\gamma(h) &= \mathbb{E}(X_t X_{t-h}) \\
&= \varphi_1\mathbb{E}(X_{t-1}X_{t-h}) + \cdots + \varphi_p\mathbb{E}(X_{t-p}X_{t-h}) - \theta_1 \underbrace{\mathbb{E}(\varepsilon_{t-1}X_{t-h})}_{=0} - \cdots - \theta_q \underbrace{\mathbb{E}(\varepsilon_{t-q}X_{t-h})}_{=0} \\
&= \varphi_1\gamma(h-1) + \cdots + \varphi_p\gamma(h-p)
\end{aligned}$$

Il s'en suit que :

$$\rho(h) = \varphi_1\rho(h-1) + \cdots + \varphi_p\rho(h-p)$$

Equations de Yule-Walker : L'équation précédente pour $k = q+1,\ldots,q+p$ donne :

$$\begin{pmatrix} \rho(q) & \ldots & \rho(q+p-1) \\ \vdots & \ddots & \vdots \\ \rho(q+p-1) & \ldots & \rho(q) \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_p \end{pmatrix} = \begin{pmatrix} \rho(p+1) \\ \vdots \\ \rho(p+q) \end{pmatrix}$$

2.4 Processus $ARIMA$

Introduction : On considére un processus $(X_t)_{t\in\mathbb{Z}}$ correspondant à une marche aléatoire c'est à dire $\varepsilon_t \leadsto \mathcal{BB}(0,\sigma^2)$ tel que :

$$\forall t > 0,\ \mathbb{C}ov(\varepsilon_t,X_0) = 0$$

$$\forall t \geqslant 0,\ X_t = X_{t-1} + \varepsilon_t$$

Alors :

$$\begin{aligned} X_t &= X_0 + \sum_{k=1}^{t} \varepsilon_k = X_0 + \sum_{j=0}^{t-1} \varepsilon_{t-j} \\ &= X_{-1} + \sum_{k=0}^{t} \varepsilon_k = X_{-1} + \sum_{j=0}^{t} \varepsilon_{t-j} \end{aligned}$$

On ne peut pas itérer le procédé car $\sum_{j=0}^{+\infty} \varepsilon_{t-j}$ **n'est pas définie**, on peut alors penser à considérer :

$$(1-L)X_t = X_t - X_{t-1} = \Delta X_t = \varepsilon_t$$

Idée : $X_t \rightsquigarrow ARIMA(p,d,q)$ si et seulement si $(1-L)^d X_t$ est stationnaire alors que $(1-L)^{d-1} X_t$ ne l'est pas (dans le cas de la marche aléatoire, $d = 1$).

Définition 2.4.1 *$(X_t)_{t \geqslant -pd}$ est un processus $ARIMA(p,d,q)$ en représentation canonique minimale s'il vérifie une équation du type :*

$$\forall t \geqslant 0, (1-L)^d \phi(L) X_t = \mu + \theta(L)\varepsilon_t$$

Et ceci avec :
- $\phi(L) = 1 - \varphi_1 L - \cdots - \varphi_p L^p$ *où* $\varphi_p \neq 0$
- $\theta(L) = 1 - \theta_1 L - \cdots - \theta_q L^q$ *où* $\theta_q \neq 0$
- $|z| \leqslant 1 \Rightarrow \phi(z) \neq 0$ *et* $\theta(z) \neq 0$
- ϕ *et* θ *n'ont pas de racines communes*
- $\varepsilon_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$ *avec* $\mathbb{C}ov(\varepsilon_t, Z) = 0$

$Z = (X_{-1}, \ldots, X_{-pd}, \varepsilon_{-1}, \ldots, \varepsilon_{-q})$ *représentant les* ***conditions initiales****.*

Proposition 2.4.1 *$(1-L)^d X_t = Y_t$ est alors asymptotiquement équivalent à un processus ARMA (donc stationnaire), ce qui signifie que il existe un processus stationnaire $(Z_t)_{t \in \mathbb{Z}}$ tel que :*

$$\phi(L) Z_t = \mu + \theta(L)\varepsilon_t$$

$$\lim_{t \to +\infty} \|Y_t - Z_t\|_2 = 0$$

Définition 2.4.2 *Si $(1-L)^d X_t$ est asymptotiquement équivalent à un processus stationnaire et si $(1-L)^{d-1} X_t$ ne l'est pas alors on dit que (X_t) est intégré d'ordre d et on note $X_t \rightsquigarrow I(d)$.*

Remarque : Si (X_t) stationnaire alors $X_t \rightsquigarrow I(0)$.

2.4.1 Approximation auto-régressive d'un $ARIMA(p,d,q)$

Proposition 2.4.2 *Avec les notations précédentes, il existe $A_t(L)$, μ_0 et $h(t)$ tels que*[1] *:*

- $A_t(L) = \sum_{j=0}^{t} a_j^t L^j$ *et* $a_t^0 = 1$
- $h(t) \in \mathbb{R}^{p+d+q}$ *et* $\lim_{t\to+\infty} h(t) = 0$
- $A_t(L)X_t = \mu_0 + \varepsilon_t + h(t)'Z \iff X_t = -\sum_{j=1}^{t} a_j^t X_{t-j} + \varepsilon_t + h(t)'Z$

Démonstration :

On pose $\psi(L) = (1-L)^d\phi(L)$, avec cette notation :

$$\psi(L)X_t = \mu + \theta(L)\varepsilon_t \quad d^o\psi = p+d,\, d^o\theta = q$$

On effectue la division selon les puissances croissantes à l'ordre t de 1 par $\theta(z)$:

$$1 = \theta(z)Q_t(z) + z^{t+1}R_t(z) \text{ où } d^oQ_t = t,\, d^oR_t = q-1$$

Ce qui implique :

$$1 = \theta(L)Q_t(L) + L^{t+1}R_t(L)$$

Or :

$$\begin{aligned}\underbrace{\psi(L)Q_t(L)}_{d^o=p+d+t}X_t &= Q_t(1)\mu + Q_t(L)\theta(L)\varepsilon_t \\ &= Q_t(1)\mu + (1-L^{t+1}R_t(L))\varepsilon_t\end{aligned}$$

Ainsi :

$$\sum_{j=0}^{p+d+t} a_j^t X_{t-j} = \mu_0 + \varepsilon_t - R_t(L)\varepsilon_{-1}$$

En décomposant la somme :

$$\sum_{j=0}^{t} a_j^t X_{t-j} = \mu_0 + \varepsilon_t - \sum_{j=t+1}^{p+d+t} a_j^t X_{t-j} - \sum_{k=0}^{q-1} r_k^t \varepsilon_{-1-k}$$

On effectue le changement d'indice $k = t-j$ dans $\sum_{j=t+1}^{p+d+t} a_j^t X_{t-j}$:

$$\sum_{j=0}^{t} a_j^t X_{t-j} = \mu_0 + \varepsilon_t \underbrace{- \sum_{k=-p-d}^{-1} a_{t-k}^t X_k - \sum_{k=0}^{q-1} r_k^t \varepsilon_{-1-k}}_{h(t)'Z}$$

---

[1] Ici la notation $h(t)'$ désigne la transposée de $h(t)$.

2.4.2 Approximation moyenne mobile d'un $ARIMA(p,d,q)$

Proposition 2.4.3 *En utilisant de nouveau les notations utilisées plus haut, il existe $B_t(L)$, μ_1 et $\tilde{h}(t)$ tels que :*

- $B_t(L) = \sum_{j=0}^{t} b_j^t L^j$ *et* $b_t^0 = 1$
- $\tilde{h}(t) \in \mathbb{R}^{p+d+q}$ *et* $\lim_{t \to +\infty} \tilde{h}(t) = 0$
- $X_t = \mu_1 + B_t(L)\varepsilon_t + \tilde{h}(t)'Z$

Proposition 2.4.4 (Calcul de $\mathbb{E}X_t$) *Si l'on note $m_t = \mathbb{E}X_t$ alors m_t vérifie $\psi(L)m_t = \mu$. On obtient ainsi :*

$\rightarrow$ *une équation de récurrence dont le polynôme caractéristique est* $z^{p+d+1}\psi(\frac{1}{z})$,

$\rightarrow$ *une forme générale de la solution (pour $\mu = 0$ et $\mu \neq 0$).*

Exemple 1 : Si $X_t - X_{t-1} = \mu + \varepsilon_t$ alors $m_t - m_{t-1} = \mu$, dans ces conditions deux cas peuvent se produire :

- $\mu = 0$ et alors $m_t = m_0 \; \forall t$;
- Ou $\mu \neq 0$ et alors $m_t = m_0 + \mu t \; \forall t$.

Remarque :

On a vu que : $X_t = X_0 + \sum_{k=1}^{t} \varepsilon_k$ si $\mu = 0 \Rightarrow \mathbb{E}X_t = \mathbb{E}X_0$

$$X_t = X_0 + \mu t + \sum_{k=1}^{t} \varepsilon_k \text{ si } \mu \neq 0 \Rightarrow \mathbb{E}X_t = \mathbb{E}X_0 + \mu t$$

Exemple 2 : $(1 - L)(1 - \varphi L)X_t = \varepsilon_t$ et alors $m_t = \alpha + \beta\varphi t$.

Chapitre 3

Identification et estimation d'un modèle $ARMA$ ou $ARIMA$

Introduction

On observe $x_1,\ldots,x_T$

On suppose que les observations proviennent d'un processus (X_t) et que :

- $(X_t)_{t\in\mathbb{Z}}$ est un processus stationnaire auquel cas il faut estimer un $ARMA(p,q)$,
- ou $(X_t) \rightsquigarrow I(d)$ et est donc non stationnaire ($(1-L)^d X_t$ est stationnaire), dans ce cas il faut estimer un $ARIMA(p,d,q)$.

A faire :

(i) Choix de d

(ii) Choix de (p,q)

(iii) Estimer $\varphi_1,\ldots,\varphi_p,\theta_1,\ldots,\theta_q$ (ce qui peut se faire par le maximum de vraisemblance sous l'hypothèse que les $\varepsilon_t \rightsquigarrow \mathcal{N}(0,\sigma^2)$ sont i.i.d.) et σ^2

(iv) Phase de vérification : $\rightarrow \varphi_p \neq 0$?
$\rightarrow \varphi_p \neq 0$?
$\rightarrow \varepsilon_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$?

Les deux premières étapes constituent la phase d'identification du processus et pour vérifier la non nullité des coefficients lors de la phase de vérification il faudra définir les tests auxquels on aura recours.

Remarque : En ce qui concerne le choix de d on peut procéder de façon empirique (en observant les auto-corrélogrammes) ou en effectuant des tests de racine unité :

$$\begin{cases} H_0 : d = 1 \quad H_1 : d = 0 & \rightarrow \text{DF (ADF), PP, SP} \\ H_0 : d = 0 \quad H_1 : d = 1 & \rightarrow \text{KPSS} \end{cases}$$

3.1 Première phase de l'identification : choix de d

3.1.1 Approche descriptive

On a vu que :

- si $X_t \rightsquigarrow ARMA(p,q)$, les $\rho(h)$ décroissent exponentiellement vers 0 avec h (pour $h > q$)
- si (X_t) est stationnaire $\hat{\rho}_T(h) \xrightarrow{P} \rho(h)$
- Sous des hypothèses suffisantes ($\mathbb{E}(\varepsilon_t^4) = \text{cte}$) :

$$\sqrt{T}\begin{pmatrix} \hat{\rho}_T(1) - \rho(1) \\ \hat{\rho}_T(H) - \rho(H) \end{pmatrix} \xrightarrow{L} \mathcal{N}(0,*), \quad \forall H$$

Remarque : Si (X_t) n'est pas stationnaire, la proposition $\rho(h)$ décroit exponentiellement vers 0 avec h n'est plus vraie (cf. persistance des chocs).

Exemple : On considère un processus (X_t) tel que $X_t - X_{t-1} = \varepsilon_t$ où $\varepsilon_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$ et $\mathbb{C}ov(\varepsilon_t,X_0) = 0$ si $t \geqslant 0$.

Remarque : Supposons $(1 - \varphi L)X_t = \varepsilon_t$ et $|\varphi| < 1$ alors :

$$X_t = \sum_{k=0}^{+\infty} \varphi^k \varepsilon_{t-k}$$

$$X_{t+h} = \sum_{k=0}^{+\infty} \varphi^k \varepsilon_{t+h-k}$$

$$X_{t+1} = \sum_{k=0}^{+\infty} \varphi^k \varepsilon_{t+1-k} = \varepsilon_{t+1} + \varphi\left(\varepsilon_t + \sum_{k=1}^{+\infty} \varphi^k \varepsilon_{t-k}\right)$$

$$\begin{aligned} \rho(h) &= \frac{\mathbb{C}ov(X_t,X_{t+h})}{(\mathbb{V}ar(X_t)\mathbb{V}ar(X_{t+h}))^{1/2}} = \frac{\mathbb{C}ov\left(X_0 + \sum_{k=1}^{t} \varepsilon_k, X_0 + \sum_{j=1}^{t+h} \varepsilon_j\right)}{(\mathbb{V}arX_0 + t\sigma^2)^{1/2}(\mathbb{V}arX_0 + (t+h)\sigma^2)^{1/2}} \\ &= \frac{\mathbb{V}arX_0 + t\sigma^2}{(\mathbb{V}arX_0 + t\sigma^2)^{1/2}(\mathbb{V}arX_0 + (t+h)\sigma^2)^{1/2}} \end{aligned}$$

Pour t grand et $h << t$, $\rho(h) \# \dfrac{t}{\sqrt{t(t+h)}} = \dfrac{1}{\sqrt{1+\frac{h}{t}}} \# \; 1 - \dfrac{h}{2t}$.

La décroissance est lente et linéaire en h. D'où une règle pratique : **Si les $\hat{\rho}_T(h)$ restent proches de 1 ou décroissent linéairement avec h alors le processus est sans doute non stationnaire**.

3.1.2 Approche par les tests de racine unité

La plupart des tests confrontent l'hypothèse $H_0 : d = 1$ (existence d'une racine unité) à l'hypothèse alternative $H_1 : d = 0$ (pas de racine unité). Ils présentent cependant le problème majeur d'être peu puissants. Il existe un test (KPSS) qui confronte $H_0 : d = 0$ à $H_1 : d = 1$.

3.1.2.1 Les tests de Dickey-Fuller

Cas 1 : Test de Dickey-Fuller simple.

On suppose $X_t = \rho X_{t-1} + \varepsilon_t$ avec comme de coutume $\varepsilon_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$ et pour changer un peu $|\rho| \leqslant 1$. On teste alors :

$$\begin{array}{ccc} H_0 : \rho < 1 & \text{contre} & H_1 : \rho = 1 \\ (1-L)X_t = \varepsilon_t & & AR(1) \\ \text{marche aléatoire} & & \end{array}$$

On a :

$$\hat{\rho}_T = \frac{\sum_{t=2}^{T} x_t x_{t-1}}{\sum_{t=2}^{T} x_{t-1}^2}$$

On rejette H_0 si $\hat{\rho}_T < c_\alpha$ où $P_{H_0}(\hat{\rho}_T < c_\alpha) = \alpha$. On est alors amené à se demander quelle est la loi de $\hat{\rho}_T$ sous H_0. En effet sous cette hypothèse (X_t) n'est pas stationnaire donc on n'a pas :

$$\sqrt{T}(\hat{\rho}_T - 1) \xrightarrow{L} \mathcal{N}(0,*)$$

Cependant on sait calculer la loi de $T\hat{\rho}_T$ sous H_0 :

$\rightarrow$ On sait tabuler cette loi

$\rightarrow$ On sait déterminer c_α tel que $P_{H_0}(T\hat{\rho}_T < c_\alpha) = \alpha$

$\rightarrow$ On rejettera H_0 si $T\hat{\rho}_T < c_\alpha$

3.2 Deuxième phase de l'identification : choix de p et q

On suppose que l'on a déjà déterminé d et on travaille éventuellement sur $Y_t = (1-L)^d X_t$. On assimile alors Y_t à un processus $ARMA(p,q)$. On se propose donc de déterminer les valeurs de p et q.

3.2.1 Résultats préliminaires

On a vu que si $Y_t \rightsquigarrow AR(p)$:

$r(h) = 0$ si $h > p$ et $r(p) \neq 0$

$\rho_i(h) = 0$ si $h > p$ et $\rho_i(p) \neq 0$

Et si $Y_t \rightsquigarrow MA(q)$:

$\rho(h) = 0$ si $h > q$ et $\rho(q) \neq 0$

Enfin on sait que si $Y_t \rightsquigarrow I(0)$:

$\hat{r}_T(h) \xrightarrow{P} r(h)$

$\hat{\rho}_{iT}(h) \xrightarrow{P} \rho_i(h)$

$\hat{\rho}_T(h) \xrightarrow{P} \rho(h)$

De plus si (ε_t) est stationnaire à l'ordre 4 ($\mathbb{E}(\varepsilon_t^4) = \mu > +\infty$) alors tous ces estimateurs sont asymptotiquement normaux.

Remarque : Calcul de $\hat{\rho}_{iT}$ pour $Y_t \rightsquigarrow ARMA(p,q)$

Si $\phi(L)Y_t = \theta(L)\varepsilon_t$, on a vu que pour $\eta_t \rightsquigarrow \mathcal{BB}$ bien choisi, le processus (Z_t) respectant l'équation $\theta(L)Z_t = \phi(L)\eta_t$ vérifie $\rho_Z(h) = \rho_{iY}(h)$.

On a :

$$Z_t = \theta(L)^{-1}\phi(L)\eta_t = A(L)\eta_t = \sum_{j=0}^{+\infty} a_j \eta_{t-j}$$

et :

$$\rho_Z(h) = \frac{\sum_{j=0}^{+\infty} a_j a_{j+h}}{\sum_{j=0}^{+\infty} a_j^2} = \rho_{iY}(h)$$

On prendra :

$$\hat{\rho}_{iY}(h) = \hat{\rho}_Z(h) = \frac{\sum_{j=0}^{K} \hat{a}_j \hat{a}_{j+h}}{\sum_{j=0}^{K} \hat{a}_j^2}$$

où K est suffisamment grand.

Par ailleurs : $\theta(L)^{-1}\phi(L)Y_t = \varepsilon_t$

$\Rightarrow A(L)Y_t = \varepsilon_t$

$\Rightarrow Y_t = -\sum_{j=1}^{+\infty} a_j Y_{t-j} + \varepsilon_t$

Il est possible d'obtenir $\hat{a}_1, \ldots, \hat{a}_K$ en régressant Y_t sur $Y_{t-1}, \ldots, Y_{t-K}$.

3.2.2 Choix de p pour $Y_t \rightsquigarrow AR(p)$

3.2.3 Choix de (p,q) pour un $ARMA(p,q)$

Rappel : $AR(p) = ARMA(p,0)$ et $MA(q) = ARMA(0,q)$.

Si $Y_t \rightsquigarrow ARMA(p,q)$ tel que $\phi(L)Y_t = \theta(L)\varepsilon_t$ alors Y_t admet une représentation $AR(\infty)$ donnée par :

$$\theta(L)^{-1}\phi(L)Y_t = \sum_{k=0}^{+\infty} a_k Y_{t-k} = \varepsilon_t$$

où $a_0 = 1$ et $\sum_{k=0}^{+\infty} |a_k| < +\infty$. Cette représentation peut être approximée par un $AR(P)$ pour P assez grand :

$$\sum_{k=0}^{P} a_k Y_{t-k} \# \varepsilon_t$$

De la même façon Y_t admet une représentation $MA(\infty)$ donnée par :

$$Y_t = \phi(L)^{-1}\theta(L)\varepsilon_t = \sum_{k=0}^{+\infty} b_k \varepsilon_{t-k}$$

qui peut aussi être approximée par un $MA(Q)$ pour Q assez grand :

$$\sum_{k=0}^{Q} b_k \varepsilon_{t-k} \# Y_t$$

Exemple : Si $\hat{r}_T(h)$ et $\hat{\rho}^i_T(h)$ sont significativement non nuls pour $h \leqslant 3$ seulement et si $\hat{\rho}_T(h)$ est significativement non nul pour $h \leqslant 4$ seulement, on est amené à estimer un $AR(3)$ et un $MA(4)$ ce qui pousse à choisir un $ARMA(p,q)$ avec $p \leqslant 3$ et $q \leqslant 4$.

3.3 Estimation

On recherche l'EMV (MLE) sous l'hypothèse que les $\varepsilon_t \rightsquigarrow \mathcal{N}(0,\sigma^2)$ sont i.i.d., on suppose (p,d,q) fixés tels que :

$$(1-L)^d\phi(L)X_t = \mu + \theta(L)\varepsilon_t$$

Les paramètres à estimer sont $\omega = (\varphi_1,\dots,\varphi_p,\theta_1,\dots,\theta_q)$ et σ^2.

On a recours à des prosèdures numériques de maximisation de la vraisemblance :

- La valeur initiale $\dfrac{\mu}{\phi(1)}$ est estimée par la moyenne empirique de $(1-L)^d X_t$.
- Les équations de Yule-Walker donnent un estimateur initial pour $(\varphi_1,\dots,\varphi_p)$:

$$\begin{pmatrix} \rho(q+1) \\ \vdots \\ \rho(q+p) \end{pmatrix} = \begin{pmatrix} \hat{\rho}(q) & \dots & \hat{\rho}(p+q-1) \\ \vdots & \ddots & \vdots \\ \hat{\rho}(p+q-1) & \dots & \hat{\rho}(q) \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_p \end{pmatrix}$$

- On obtient de la même façon $(\theta_1,\dots,\theta_q)$ à partir des $\hat{\rho}^i_T(h)$

3.3.1 Cas d'un $AR(p)$

On se place dans le cadre suivant :

$$y_t = (1-L)^d(x_t - m)$$

$$\phi(L)y_t = \varepsilon_t$$

$$y_t = \varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p} + \varepsilon_t \quad \varepsilon_t \rightsquigarrow \mathcal{N}(0,\sigma^2) \text{ i.i.d.}$$

$\begin{pmatrix} y_{p+1} \\ \vdots \\ y_T \end{pmatrix}$ suit, conditionnellement à $\begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$, une loi normale dont la densité s'écrit :

$$\ell(y_{p+1},\ldots,y_T|y_1,\ldots,y_p) = \prod_{t=p+1}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{y_t - (\varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p})}{2\sigma^2}\right)$$

$$\Rightarrow \ln \ell(y_{p+1},\ldots,y_T|y_1,\ldots,y_p) = -\frac{T}{2}\ln 2\pi - \frac{T}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=p+1}^{T}[y_t - (\varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p})]$$

La maximixsation de la vraisemblance conditionnelle donne $\hat{\varphi}_1,\ldots,\hat{\varphi}_p$ qui peuvent être calculés en régressant (MCO) y_t sur $y_{t-1},\ldots,y_{t-p}$.

CLS (méthode approximative) On a :

$$\ell(y_1,\ldots,y_T) = \ell(y_{p+1},\ldots,y_T|y_1,\ldots,y_p) \times \ell(y_1,\ldots,y_p)$$

où :

$$\begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} \rightsquigarrow \mathcal{N}\left(0, \underbrace{\begin{pmatrix} \gamma(0) & \ldots & \gamma(p-1) \\ \vdots & \ddots & \vdots \\ \gamma(p-1) & \ldots & \gamma(0) \end{pmatrix}}_{V_p}\right)$$

V_p étant calculable en fonction de $\varphi_1,\ldots,\varphi_p$, il en découle :

$$\ell(\underbrace{y_1,\ldots,y_p}_{z}) = \frac{1}{\sqrt[p]{2\pi}}\frac{1}{\sqrt{\det V_p}} e^{-\frac{1}{2}z'V_p z}$$

$$\max_{\sigma^2,\varphi_1,\ldots,\varphi_p} \ln \ell(y_1,\ldots,y_p) = \max_{\sigma^2,\varphi_1,\ldots,\varphi_p} -\frac{p}{2}\ln 2\pi - \frac{1}{2}\ln\det V_p - \frac{1}{2}z'V_p z$$

Pour plus de détails sur le cas $AR(p)$ on pourra consulter [2].

3.3.2 Cas d'un $MA(q)$

On considère :

$$(1-L)^d Y_t = \theta(L)\varepsilon_t$$

$$Y_t = X_t - m_t \text{ où } \mathbb{E}X_t = m_t$$

$$\varepsilon_t \rightsquigarrow \mathcal{N}(0,\sigma^2) \text{ i.i.d.}$$
$$\theta(L)\varepsilon_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}$$

La vraisemblance du modèle est $\ell(y_1,\ldots,y_T|\varepsilon_0,\varepsilon_{-1},\ldots,\varepsilon_{-q+1})$, fonction dont le maximum doit être calculé de façon numérique.

- Soit on suppose $\varepsilon_0 = \varepsilon_{-1} = \cdots = \varepsilon_{-q+1} = 0 \rightarrow$ méthode numérique $\rightarrow 2^{\text{è}}$ approximation $\rightarrow$ CLS
- Soit on ne fait pas cette hypothèse $\rightarrow$ méthode numérique $\rightarrow$ approximation $\rightarrow$ ULLS

Pour plus de précisions cf. [1].

3.3.3 Cas d'un $ARMA(p,q)$

Diverses approximations : MLE, CLS, ULLS (cf. [1]).

3.3.4 Propriétés asymptotiques de l'EMV

On note que $\sqrt{T}\begin{pmatrix} \hat{\omega}_T - \omega \\ \hat{\sigma}^2_T - \sigma^2 \end{pmatrix} \xrightarrow{\text{loi}} \mathcal{N}\left(0, \begin{pmatrix} \Omega & 0 \\ 0 & \alpha \end{pmatrix}\right)$. Il existe des tests asymptotiques sur $\omega = (\varphi_1,\ldots,\varphi_p,\theta_1,\ldots,\theta_q,\mu)'$.

3.4 Vérifications a posteriori

3.4.1 Tests sur les paramètres

On désire tester $H_0 : \varphi_p = 0$ contre l'hypothèse alternative $H_1 : \varphi_p \neq 0$, on a :

$$\sqrt{T}(\hat{\varphi}_{T,p} - \varphi_p) \xrightarrow{L} \mathcal{N}(0, V_{as}(\sqrt{T}\hat{\varphi}_{T,p}))$$

φ_p est significatif au seuil de 5% (asymptotiquement) si :

$$\underbrace{\left| \frac{\hat{\varphi}_{T,p}}{(V_{as}(\hat{\varphi}_{T,p})^{1/2})} \right|}_{|t_{\hat{\varphi}_{T,p}}|} > 1{,}96$$

Le même test s'applique avec θ_q à la place de φ_p, on refuse alors H_0 (i.e. θ_q est significatif) si $|t_{\hat{\varphi}_{T,p}}| > 1{,}96$.

Remarque : Économétrie asymptotique :

$$\hat{\beta} = \beta + \frac{(X'X)^{-1}}{T}\frac{X'}{T}\varepsilon \text{ et } \frac{X'X}{T} \rightarrow Q$$

$$\sqrt{T}\frac{X'\varepsilon}{T} \xrightarrow{L} \mathcal{N}\left(0, \underbrace{\lim\left(\frac{X'X}{T}\right)}_{Q}\right)$$

$$y_t = a + bx_t + \varepsilon_t$$

$$\frac{\hat{b}-b}{\hat{\sigma}_{\hat{b}}} \overset{H_0}{\rightsquigarrow} \tau(N-2)$$

Si $\hat{\varphi}_p$ n'est pas significatif (ou θ_q) on relance l'estimation en remplaçant p par $p-1$ (ou q par $q-1$).

Si $\hat{\mu}$ n'est pas significatif, on relance l'estimation sans terme constant.

3.4.2 Tests sur les résidus

On se place toujours dans les mêmes conditions :

$$\phi(L)Y_t = \mu + \theta(L)\varepsilon_t$$

$$\Rightarrow Y_t - \varphi_1 Y_{t-1} - \cdots - \varphi_p Y_{t-p} = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}$$

$$\Rightarrow \hat{\varepsilon}_t = Y_t - \hat{\varphi}_1 Y_{t-1} - \cdots - \hat{\varphi}_p Y_{t-p} - \hat{\mu} + \hat{\theta}_1\varepsilon_{t-1} + \cdots + \hat{\theta}_q\varepsilon_{t-q}$$

Les résidus estimés (à savoir $\hat{\varepsilon_t}$) sont-ils compatibles avec l'hypothèse de bruit blanc de ε_t ? Pour celà on calcule :

$$\hat{\rho}_k(\hat{\varepsilon}) = \frac{1}{T-1}\sum_{t=k+1}^{T} \hat{\varepsilon}_t\hat{\varepsilon}_{t-k}$$

l'auto-corrélation empirique d'ordre k de $\hat{\varepsilon}_t$, soit alors:

$$Q_K = T\sum_{k=1}^{K} \hat{\rho}_k^2(\hat{\varepsilon})$$

calculée pour K assez grand ($K \geqslant 12$), on montre que si $(\varepsilon_t) \rightsquigarrow \mathcal{BB}(0,\sigma^2)$ alors $Q_K \xrightarrow{L} \chi^2(K-p-q)$.

On refuse $H_0 : \varepsilon_t \rightsquigarrow \mathcal{BB}$ au seuil α si $Q_K > \chi^2_{1-\alpha}(K-p-q)$

Remarque : Q_K peut être éventuellement remplacé par :

$$Q'_K = T(T+2)\sum_{k=1}^{K} \frac{1}{T-k}\hat{\rho}_k^2(\hat{\varepsilon})$$

3.5 Choix du modèle

A l'issue des phases d'estimation et de vérification il reste en général plusieurs modèles possibles pour représenter les données. Pour choisir un modèle on peut se fier à plusieurs critères :

- $\hat{\sigma}^2$ petit
- Critère de parcimonie : $p+q$ minimal
- Critère de qualité de qualité de la prévision (cf. plus loin)
- Critère d'information

On suppose $\varepsilon_t \rightsquigarrow \mathcal{BB}(0,\sigma^2)$ i.i.d. On considère $f_0(x) = f(x,\omega_0,\sigma_0^2)$ la vraie loi (inconnue) du processus et $f(x) = f(x,\omega,\sigma^2)$ la famille de loi correspondant au modèle $ARMA(p,q)$ estimé. L'écart entre ces lois est mesuré par :

$$\Delta(f,f_0) = \mathbb{E}_0\left[-2\ln\frac{f(x)}{f_0(x)}\right]$$

Cette quantité est positive (d'après Jensen) et est nulle si et seulement si $f(x) = f_0(x)$ p.s.

En pratique on cherche à minimiser l'écart entre f_0 et f. Il existe différentes façons d'approximer ce critère d'information :

- $AIC(p,q) = \ln\hat{\sigma^2} + \frac{2(p+q)}{T}$ critère d'Akaïke
- $AIC(p,q) = \ln\hat{\sigma^2} + (p+q)\frac{\ln T}{T}$
- $AIC(p,q) = \ln\hat{\sigma^2} + (p+q)c\frac{\ln(\ln T)}{T}$ avec $c > 2$, critère d'Hannan-Quinn

Chapitre 4

Prévision dans les $ARMA$ et les $ARIMA$

Introduction

On suppose (p,d,q) connus. on remplacera $(\varphi_1,\ldots,\varphi_p,\theta_1,\ldots,\theta_q,\mu,\sigma^2)$ par leurs estimateurs.

A la date t, on suppose $x_1,\ldots,x_t$ observés, faire un prévision à l'horizon H c'est prévoir $x_{t+1},\ldots,x_{t+H}$.

- On peut réaliser une prévision optimale, c'est à dire faire comme si les x_{t-k} étaient observés pour tout $k \geqslant 0$, on obtient alors
- Ou une prévision approchèe

Bibliographie

[1] C. GOURIEROUX et A. MONFORT. *Séries temporelles et modèles dynamiques.* Economica.

[2] J. HAMILTON. *Time Series Analysis.* Princeton University Press.

Table des matières

Introduction 1

1 Processus réels stationnaires (au second ordre) 2

1.1 Processus stationnaires du second ordre 2
1.1.1 Définitions 2
1.1.2 Rappels sur $\mathcal{L}^2(\Omega,\mathcal{A},P)$ 5
1.2 Outils pour l'étude des processus stationnaires 5
1.2.1 Transformée d'un processus stationnaire par une moyenne mobile infinie 5
1.2.2 Régression linéaire ou affine théorique sur un nombre fini de retards 6
1.2.3 Régression linéaire théorique sur un nombre infini de retards 8
1.2.4 Densité spectrale et auto-corrélations inverses 9
1.2.5 Estimateurs associés et lois limites 12
1.3 Polynômes retard et avance 13
1.3.1 Définition et propriétés 13
1.3.2 Inversibilité des polynômes en L (vrais polynômes) 14

2 Processus ARMA et ARIMA 19

2.1 Processus $AR(p)$ (auto-régressif d'ordre p) 19
2.1.1 Définition et représentation canonique 19
2.1.2 Propriétés des processus $AR(p)$ 23
2.1.3 Autocorrélations partielles et inverse d'un processus $AR(p)$ 25
2.2 Processus $MA(q)$ 26
2.2.1 Définition et représentation canonique 26
2.2.2 Propriétés des processus $MA(q)$ 28
2.3 Processus $ARMA$ 29
2.3.1 Définition et représentation canonique minimale 29
2.3.2 Propriétés 32
2.4 Processus $ARIMA$ 32
2.4.1 Approximation auto-régressive d'un $ARIMA(p,d,q)$ . . . 34
2.4.2 Approximation moyenne mobile d'un $ARIMA(p,d,q)$ . . . 35

3 Identification et estimation d'un modèle $ARMA$ ou $ARIMA$ **36**

3.1 Première phase de l'identification : choix de d 37
3.1.1 Approche descriptive 37
3.1.2 Approche par les tests de racine unité 37
3.2 Deuxième phase de l'identification : choix de p et q 39

3.2.1 Résultats préliminaires 39
3.2.2 Choix de p pour $Y_t \rightsquigarrow AR(p)$ 39
3.2.3 Choix de (p,q) pour un $ARMA(p,q)$ 39
3.3 Estimation 40
3.3.1 Cas d'un $AR(p)$ 41
3.3.2 Cas d'un $MA(q)$ 41
3.3.3 Cas d'un $ARMA(p,q)$ 42
3.3.4 Propriétés asymptotiques de l'EMV 42
3.4 Vérifications a posteriori 42
3.4.1 Tests sur les paramètres 42
3.4.2 Tests sur les résidus 43
3.5 Choix du modèle 43

4 Prévision dans les $ARMA$ et les $ARIMA$ **45**

Bibliographie **46**

Index

Akaïké (critère d'), 43
auto-corrélation, 4
auto-covariance, 4
autocorrélation inverse, 12

bruit blanc, 3
  faible, 3
  fort, 3

critère
  d'information, 43
  de parcimonie, 43
  de qualité de la prévision, 43

densité spectrale, 9
Dickey-Fuller (test de), 36

Hannan-Quinn (critère d'), 43

innovation, 9

marche aléatoire, 3
moyenne mobile
  infinie, 6

opérateur, 13
  avance, 13
  retard, 13

polynôme, 14
  avance, 14
  retard, 14
processus AR, 19
processus ARIMA, 32
processus ARMA, 29
processus des innovations, 9
processus déterministe, 19
processus MA, 25
processus stationnaire, 2
  du second ordre, 2
  strict, 2
prévision optimale, 9

représentation canonique d'un processus, 21
représentation canonique minimale, 29
régression, 6
  affine théorique (retards finis), 6
  affine théorique (retards infinis), 8
  linéaire théorique (retards finis), 6
  linéaire théorique (retards infinis), 8

Wold (théorème de), 9

Yule-Walker (équations de), 23, 32