

Séries temporelles S2  
Université de Strasbourg  
M1 DUAS et M1 Statistiques

Davide Giraud

30 avril 2025

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Définition d'une série temporelle . . . . .	2
1.2	Modélisation d'une série temporelle . . . . .	2
<b>2</b>	<b>Etude de la partie déterministe de la série temporelle</b>	<b>4</b>
2.1	Estimation paramétrique . . . . .	4
2.2	Estimation par lissage par moyenne mobile . . . . .	5
2.3	Estimation par lissage exponentiel . . . . .	10
<b>3</b>	<b>Processus stochastiques à temps discret</b>	<b>12</b>
3.1	Introduction aux processus stochastiques . . . . .	12
3.2	Exemples de processus stochastiques . . . . .	15
3.3	Espérance linéaire et innovation . . . . .	18
3.4	Processus défini par une équation de récurrence . . . . .	19
3.5	Processus stationnaires au second ordre : innovation, auto-covariance, prédiction linéaire . . . . .	22
3.6	Estimation de la moyenne et de la fonction d'auto-corrélation . . . . .	25
3.7	Fonction génératrice de l'auto-covariance . . . . .	26
<b>4</b>	<b>Les processus moyennes mobiles</b>	<b>27</b>
4.1	Définition et principales caractéristiques . . . . .	27
4.2	Estimation d'un modèle moyenne mobile . . . . .	28

<b>5</b>	<b>Les processus autorégressifs</b>	<b>29</b>
5.1	Définition et principales caractéristiques . . . . .	29
5.2	Estimation d'un modèle autorégressif . . . . .	32
5.3	Prédiction d'un processus AR ( $p$ ) . . . . .	32
<b>6</b>	<b>Processus ARMA</b>	<b>33</b>
6.1	Définition . . . . .	33
6.2	Représentation d'un processus ARMA ( $p, q$ ) en MA ( $\infty$ ) . . . . .	34
6.3	Représentation d'un ARMA ( $p, q$ ) en AR ( $\infty$ ) . . . . .	34
6.4	Innovation d'un processus ARMA . . . . .	35
<b>7</b>	<b>Processus ARIMA (Auto Regressive Integrated Moving Average)</b>	<b>36</b>

# 1 Introduction

## 1.1 Définition d'une série temporelle

On observe une série temporelle si on observe une suite  $(x_t)_{t=1}^T$  de réels où  $t$  représente la date d'une mesure d'une quantité d'intérêt  $Q$  et  $x_t$  est sa valeur à l'instant  $t$ . Par exemple,  $Q$  peut être la température, le prix d'un bien, le taux d'humidité, etc...

Une série temporelle  $(x_t)_{t=1}^T$  est représentée par l'ensemble des points  $\{(t, x_t), t \in \llbracket 1, T \rrbracket\}$  reliés entre eux, où  $\llbracket a, b \rrbracket = \{k \in \mathbb{N}, a \leq k \leq b\}$ . La plupart du temps, ce type de graphique permet de mettre en évidence les composantes fondamentales d'une série temporelle :

- la tendance, qui représente l'évolution sur le long terme de la série,
- la composante saisonnière, qui induit des creux et sommets sur le graphique et qui correspond à un phénomène périodique (par exemple, la température relevée quotidiennement),
- la composante d'erreur (sans laquelle la série temporelle serait déterministe) qui induit des fluctuations aléatoires,
- la composante accidentelle, qui représente des phénomènes ponctuels mais ayant des conséquences importantes (crash boursier, catastrophe naturelle, etc..). Le traitement de celle-ci dépasse le cadre de cette unité d'enseignement.

L'objectif de ce cours est de présenter des modélisations possibles des séries temporelles afin de faire de la prévision. Plus exactement, si on dispose d'observations  $(x_t)_{t=1}^T$  d'une série temporelle, on cherche à prédire la valeur de la série à l'horizon  $h \in \mathbb{N}^*$ , c'est-à-dire la valeur de  $x_{T+h}$ .

## 1.2 Modélisation d'une série temporelle

Une série temporelle  $(x_t)_{t=1}^T$  est la réalisation aux instants  $(i_t)_{t=1}^T$  d'un processus stochastique  $(X_i)_{i \in E}$ , où, pour tout  $i \in E$ ,  $X_i: \Omega \rightarrow \mathbb{R}$  est une variable aléatoire réelle définie sur l'espace

probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$ . Si  $E = \mathbb{R}$ , on parlera de processus à temps continu et si  $E = \mathbb{Z}$ , de processus à temps discret. Ici, on s'intéressera uniquement au cas discret.

La difficulté réside dans le fait que les variables aléatoires  $(X_t)_{t \in \mathbb{Z}}$  sont dépendantes. On impose un modèle de la forme

$$X_t = f(t, \varepsilon_t), t \in \mathbb{Z}, \quad (1.2.1)$$

où  $\varepsilon_t$  est une variable aléatoire centrée qui représente la composante d'erreur d'une série temporelle et  $f: \mathbb{Z} \times \mathbb{R} \rightarrow \mathbb{R}$  est une fonction inconnue. Cependant, ce modèle est trop général d'un point de vue pratique.

### 1.2.1 Modèle additif

**Définition 1.1.** Soit  $p \in \mathbb{N}^*$ . Une série temporelle  $(X_t)_{t \in \mathbb{Z}}$  est issue d'un modèle additif de période  $p$  s'il existe deux fonctions déterministes inconnues  $T: \mathbb{Z} \rightarrow \mathbb{R}$  et  $S(\cdot; p): \mathbb{Z} \rightarrow \mathbb{R}$  telles que

$$X_t = T(t) + S(t; p) + \varepsilon_t, \quad (1.2.2)$$

où

- $S(\cdot, p)$  est périodique de période  $p \in \mathbb{N}^*$ , i.e., pour tout  $t \in \mathbb{Z}$ ,  $S(t + p; p) = S(t; p)$ ,
- $\varepsilon_t$  est une variable aléatoire.

La fonction  $T(\cdot)$  représente la tendance et la fonction  $S(\cdot; p)$  représente la composante saisonnière.

La tendance est souvent supposée polynomiale, c'est-à-dire

$$T(t) = \sum_{i=0}^q a_i t^i, t \in \mathbb{Z} \quad (1.2.3)$$

où  $(a_i)_{i=0}^q$  sont des réels.

On remarque que la fonction  $S(\cdot; p)$  ne prend qu'au plus  $p$  valeurs.

**Définition 1.2.** Les valeurs  $S_i^{(p)} = S(i; p)$ ,  $i \in \llbracket 1, p \rrbracket$ , sont appelés coefficients saisonniers.

On a donc, pour  $t \in \mathbb{Z}$ ,  $S(t; p) = S_{j(t,p)}^{(p)}$  avec  $j(t, p) \in \llbracket 1, p \rrbracket$  et  $j(t, p) \equiv t \pmod{p}$ .

On supposera souvent que les effets saisonniers se compensent sur une période, c'est-à-dire  $\sum_{i=1}^p S_i^{(p)} = 0$ .

*Exemple 1.3* (Modèle de Buys-Ballot).

$$X_t = \beta_0 + \beta_1 t + S_{j(t,p)}^{(p)} + \varepsilon_t, \quad (1.2.4)$$

où  $j(t, p)$  est défini comme précédemment et  $\sum_{i=1}^p S_i^{(p)} = 0$ .

### 1.2.2 Modèle multiplicatif

**Définition 1.4.** Soit  $p \in \mathbb{N}^*$ . Une série temporelle  $(X_t)_{t \in \mathbb{Z}}$  est issue d'un modèle multiplicatif de période  $p$  s'il existe deux fonctions déterministes inconnues  $T: \mathbb{Z} \rightarrow \mathbb{R}$  et  $S(\cdot; p): \mathbb{Z} \rightarrow \mathbb{R}$  telles que

$$X_t = T(t) \times S(t; p) \times (1 + \varepsilon_t), \quad (1.2.5)$$

où

- $S(\cdot, p)$  est périodique de période  $p \in \mathbb{N}^*$ , i.e., pour tout  $t \in \mathbb{Z}$ ,  $S(t + p; p) = S(t; p)$ ,
- $\varepsilon_t$  est une variable aléatoire telle que pour tout  $t \in \mathbb{Z}$ ,  $\mathbb{P}(\varepsilon_t > -1) = 1$ .

La compensation des effets saisonniers se traduit par  $\prod_{i=1}^p S_i^{(p)} = 1$ .

### 1.2.3 Distinction des modèles

Le choix entre le modèle additif et multiplicatif se fait graphiquement.

modèle	additif	multiplicatif
variations saisonnières	constantes autour de la tendance	proportionnelles à la tendance

On passe du modèle multiplicatif au modèle additif en prenant le logarithme : si  $X_t = T(t) \times S(t; p) \times (1 + \varepsilon_t)$ , alors

$$X_t^* = T^*(t) + S^*(t; p) + \varepsilon_t^*, \quad (1.2.6)$$

avec  $T^*(t) = \log T(t)$ ,  $S^*(t; p) = \log S(t; p)$  et  $\varepsilon_t^* = \log(1 + \varepsilon_t)$ .

## 2 Etude de la partie déterministe de la série temporelle

### 2.1 Estimation paramétrique

Pour estimer la tendance et les coefficients saisonniers, on peut faire appel à un modèle paramétrique. On considère le modèle de Buys-Ballot :

$$X_t = \beta_0 + \beta_1 t + S_{j(t,p)}^{(p)} + \varepsilon_t, \quad (2.1.1)$$

où  $j(t, p)$  est défini comme précédemment et  $\sum_{i=1}^p S_i^{(p)} = 0$ . On dispose seulement d'observations du vecteur aléatoire  $\mathbf{X}_T = (X_1, \dots, X_T)^\top$  et on suppose  $T \gg p + 1$ . En supposant  $\sum_{i=1}^p S_i^{(p)} = 0$ , les paramètres du modèle de Buys-Ballot sont regroupés dans le vecteur  $\theta = (\beta_0, \beta_1, S_1^{(p)}, \dots, S_{p-1}^{(p)})^\top \in \mathbb{R}^{p+1}$ .

**Exercice 1.** Soit le processus stochastique  $(X_t)_{t \in \mathbb{Z}}$  défini par :

$$X_t = \alpha t + \beta + S_{t_p} + \varepsilon_t, \quad t \in \mathbb{Z},$$

où  $t_p \equiv t \pmod{p}$ ,  $t_p \in \{1, \dots, p\}$ . Les variables aléatoires  $(\varepsilon_t)_{t \in \mathbb{Z}}$  sont indépendantes de loi  $\mathcal{N}(0, \sigma^2)$ .

1. On suppose que  $S_1 + \dots + S_p = 0$ . Soit  $\mathbb{X} := (X_1, \dots, X_T)^t$  un vecteur de taille  $T$  extrait du processus stochastique  $(X_t)_{t \in \mathbb{Z}}$ . Montrer que  $\mathbb{X} = A\theta + \varepsilon$ , où  $A$  est une matrice que l'on précisera et  $\theta$  le vecteur des paramètres dont on précisera la dimension.
2. Proposer un estimateur du paramètre  $\theta$ . Quelles sont ses propriétés ?

La méthode des moindres carrés s'applique à des modèles plus généraux de la forme

$$X_t = g(t; \beta) + S_{j(t,p)}^{(p)} + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (2.1.2)$$

où  $g(\cdot; \beta)$  est connue et le vecteur  $\beta$  inconnu.

## 2.2 Estimation par lissage par moyenne mobile

Une moyenne mobile est un opérateur qui s'applique sur les suites (indexées par  $\mathbb{Z}$ ), en particulier, sur les séries temporelles. On note  $\mathcal{S}$  l'ensemble des suites de réels indexées par  $\mathbb{Z}$ .

**Définition 2.1.** Soit  $(x_t)_{t \in \mathbb{Z}}$  une suite de réels. L'opérateur de retard  $B: \mathcal{S} \rightarrow \mathcal{S}$  (Backward shift) est défini par

$$B((x_t)_{t \in \mathbb{Z}}) = (x_{t-1})_{t \in \mathbb{Z}}. \quad (2.2.1)$$

Sa réciproque  $B^{-1}: \mathcal{S} \rightarrow \mathcal{S}$  est l'opérateur avance noté  $F$  (Forward shift) et est défini par

$$F((x_t)_{t \in \mathbb{Z}}) = (x_{t+1})_{t \in \mathbb{Z}}. \quad (2.2.2)$$

Autrement dit, le terme d'indice  $t$  de la suite  $B((x_t)_{t \in \mathbb{Z}})$  est  $x_{t-1}$  et le terme d'indice  $t$  de la suite  $F((x_t)_{t \in \mathbb{Z}})$  est  $x_{t+1}$ .

Les opérateurs  $B$  et  $F$  sont linéaires : si  $\alpha, \beta \in \mathbb{R}$  et  $(x_t)_{t \in \mathbb{Z}}, (y_t)_{t \in \mathbb{Z}}$  sont des suites, alors

$$B(\alpha(x_t)_{t \in \mathbb{Z}} + \beta(y_t)_{t \in \mathbb{Z}}) = \alpha B((x_t)_{t \in \mathbb{Z}}) + \beta B((y_t)_{t \in \mathbb{Z}}). \quad (2.2.3)$$

$$F(\alpha(x_t)_{t \in \mathbb{Z}} + \beta(y_t)_{t \in \mathbb{Z}}) = \alpha F((x_t)_{t \in \mathbb{Z}}) + \beta F((y_t)_{t \in \mathbb{Z}}). \quad (2.2.4)$$

Pour  $k \in \mathbb{N}$ , on note  $B^k = \underbrace{B \circ \dots \circ B}_{k \text{ fois}}$ ,  $F^k = \underbrace{F \circ \dots \circ F}_{k \text{ fois}}$  et  $B^{-k} := F^k$ ,  $F^{-k} = B^k$ . De plus, on note  $I$  l'opérateur identité sur les suites indexées par  $\mathbb{Z}$ .

**Définition 2.2.** La moyenne mobile d'ordres  $(m_-, m_+) \in \mathbb{N}^2$  et de paramètre  $\theta = (\theta_i)_{i=-m_-}^{m_+} \in \mathbb{R}^{m_- + m_+ + 1}$  est l'opérateur linéaire défini par

$$M_{(m_-, m_+, \theta)} = \sum_{i=-m_-}^{m_+} \theta_i B^{-i} = \sum_{i=-m_-}^{m_+} \theta_i F^i. \quad (2.2.5)$$

Étant donnée une série chronologique  $(X_t)_{t \in \mathbb{Z}}$ , on observe seulement une réalisation du vecteur aléatoire  $(X_1, \dots, X_T)$ . Ainsi, on ne peut observer  $M_{(m_-, m_+, \theta)}((X_t)_{t \in \mathbb{Z}})$  que pour les instants  $t \in \llbracket m_- + 1, T - m_+ \rrbracket$ .

**Définition 2.3.** Une moyenne mobile  $M_{(m_-, m_+, \theta)}$  est dite centrée si  $m_- = m_+ = m$ .

**Définition 2.4.** Une moyenne mobile  $M_{(m_-, m_+, \theta)}$  est dite symétrique si elle est centrée et si le paramètre  $\theta$  est tel que  $\theta_{-i} = \theta_i$  pour tout  $i \in \llbracket 1, m \rrbracket$ .

**Exercice 2.** Montrer que la composée de deux moyennes mobiles est encore une moyenne mobile.

**Exercice 3.** Montrer que la composée de deux moyennes mobiles symétriques est encore une moyenne mobile symétrique.

**Définition 2.5.** Le polynôme caractéristique d'une moyenne mobile  $M_{(m_-, m_+, \theta)}$  est le polynôme  $P(\cdot, \theta)$  d'ordre  $m_- + m_+$  tel que  $M_{(m_-, m_+, \theta)} = B^{m_-} P(F, \theta)$ , c'est-à-dire

$$P(z, \theta) = \theta_{-m_-} + \theta_{-m_-+1}z + \cdots + \theta_{m_+}z^{m_-+m_+} = \sum_{i=0}^{m_-+m_+} \theta_{-i-m_-} z^i. \quad (2.2.6)$$

**Exercice 4.** Montrer qu'une moyenne mobile centrée d'ordre  $m \in \mathbb{N}^*$  est symétrique si et seulement si son polynôme caractéristique  $P$  est tel que  $P(z^{-1}) = z^{-2m}P(z)$  pour tout  $z \in \mathbb{C} \setminus \{0\}$ .

On fera le léger abus de notation suivant : on note, pour  $t \in \mathbb{Z}$ ,  $M_{(m_-, m_+, \theta)} x_t$  l'élément d'indice  $t$  de la suite  $M_{(m_-, m_+, \theta)}((x_t)_{t \in \mathbb{Z}})$ .

On cherche une moyenne mobile  $M_{(m_-, m_+, \theta)}$  qui préserve la tendance et annule la composante saisonnière, c'est-à-dire vérifiant

$$M_{(m_-, m_+, \theta)}((T(t)_{t \in \mathbb{Z}})) = (T(t))_{t \in \mathbb{Z}} \text{ et } M_{(m_-, m_+, \theta)}((S(t; p))_{t \in \mathbb{Z}}) = 0. \quad (2.2.7)$$

**Définition 2.6.** L'espace des invariants d'une moyenne mobile  $M_{(m_-, m_+, \theta)}$  est l'ensemble des suites  $(x_t)_{t \in \mathbb{Z}}$  vérifiant

$$\forall t \in \mathbb{Z}, \quad M_{(m_-, m_+, \theta)} x_t = x_t. \quad (2.2.8)$$

Le noyau d'une moyenne mobile  $M_{(m_-, m_+, \theta)}$  est l'ensemble des suites  $(x_t)_{t \in \mathbb{Z}}$  vérifiant

$$\forall t \in \mathbb{Z}, \quad M_{(m_-, m_+, \theta)} x_t = 0. \quad (2.2.9)$$

**Exercice 5.** Cet exercice a pour but l'étude du noyau d'une moyenne mobile.

1. Rappeler la définition d'une moyenne mobile d'ordre  $(m_-, m_+) \in \mathbb{N}^2$  et de paramètres  $\theta_{-m_-}, \dots, \theta_{m_+}$ .
2. Quel est le noyau de la moyenne mobile d'ordre  $(1, 0)$  et de paramètres  $(-1/2, 1)$ .
3. Montrer que l'ensemble des séries  $\{c2^t, c \in \mathbb{R}\}$  est inclus dans le noyau de la moyenne mobile d'ordre  $(2, 0)$  et de paramètres  $(6, -5, 1)$ . Donner un autre ensemble également inclus dans ce noyau.

Dans de nombreux cas, on modélise la tendance d'une série temporelle à l'aide d'un polynôme d'ordre  $q \in \mathbb{N}$ , ce qui conduit à trouver des moyennes mobiles dont l'espace des invariants contient les suites de la forme  $(\beta_0 + \beta_1 t + \cdots + \beta_q t^q)$ .

**Exercice 6.** Après avoir rappelé la définition d'une moyenne mobile centrée et d'une moyenne mobile symétrique, montrer les propriétés suivantes :

1. Une moyenne mobile conserve les constantes si et seulement si la somme de ses paramètres est égale à 1.
2. Une moyenne mobile centrée et symétrique dont la somme de ses paramètres est égale à 1 conserve les polynômes de degré 1.

**Proposition 2.7.** Soit  $q \in \mathbb{N}$ . Si la valeur 1 est une racine d'ordre  $q + 1$  du polynôme  $Q(x, \theta) = P(x, \theta) - x^{m_-}$  alors les polynômes d'ordre au plus  $q \in \mathbb{N}$  appartiennent à l'espace des invariants d'une moyenne mobile  $M_{(m_-, m_+, \theta)}$ .

**Corollaire 2.8.** Soit  $M_{(m_-, m_+, \theta)}$  une moyenne mobile.

1. Si  $\sum_{i=-m_-}^{m_+} \theta_i = 1$ , alors toute suite constante appartient à l'espace des invariants de  $M_{(m_-, m_+, \theta)}$ .
2. Si la moyenne mobile  $M_{(m_-, m_+, \theta)}$  est centrée, symétrique avec  $\sum_{i=-m_-}^{m_+} \theta_i = 1$ , alors pour tous  $a, b \in \mathbb{R}$ , la suite  $(a + bt)_{t \in \mathbb{Z}}$  appartient à l'espace des invariants de  $M_{(m_-, m_+, \theta)}$ .

On s'intéresse à présent au noyau d'une moyenne mobile. Soit  $j(t, p) \in \llbracket 1, p \rrbracket$  tel que  $j(t, p) = t \pmod{p}$ .

**Proposition 2.9.** Soit  $(S(t))_{t \in \mathbb{Z}}$  une suite périodique de période  $p$ , i.e.,  $S(t) = S_{j(t, p)}^{(p)}$ . Soit  $M_{(m_-, m_+, \theta)}$  une moyenne mobile admettant  $P(\cdot, \theta)$  comme polynôme caractéristique.

1. S'il existe un polynôme  $Q(\cdot, \theta)$  tel que

$$P(x, \theta) = Q(x, \theta)(1 - x^p), \quad (2.2.10)$$

alors  $(S(t))_{t \in \mathbb{Z}}$  appartient au noyau de la moyenne mobile  $M_{(m_-, m_+, \theta)}$ .

2. Si  $\sum_{i=1}^p S_i^{(p)} = 0$  et qu'il existe un polynôme  $Q(\cdot, \theta)$  tel que

$$P(x, \theta) = Q(x, \theta) \sum_{i=0}^{p-1} x^i, \quad (2.2.11)$$

alors  $(S(t))_{t \in \mathbb{Z}}$  appartient au noyau de la moyenne mobile  $M_{(m_-, m_+, \theta)}$ .

On donne à présent un exemple de moyenne mobile qui conserve la tendance et annule les variations saisonnières dans le cas où la tendance est un polynôme de degré 1 et la somme des coefficients saisonniers s'annule.

Si  $p$ , la période de variation saisonnière est impaire, on prend  $m_- = m_+ = (p - 1) / 2 = m$  et

$$MX_t = \frac{1}{2m + 1} \sum_{i=-m}^m X_{t+i}, \quad (2.2.12)$$

la moyenne des  $p = 2m + 1$  observations centrées autour de  $X_t$ .

Si  $p$  est pair, on ne peut pas avoir  $p$  observations centrées autour de  $X_t$ . On peut tout de même considérer les moyennes

$$\mu_{1,t} = \frac{1}{p} \sum_{i=t-p/2}^{t+p/2-1} X_i, \quad (2.2.13)$$

$$\mu_{2,t} = \frac{1}{p} \sum_{i=t-p/2+1}^{t+p/2} X_i, \quad (2.2.14)$$

puis prendre la moyenne mobile  $M$  telle que  $MX_t = (\mu_{1,t} + \mu_{2,t})/2$ . Elle admet pour polynôme caractéristique

$$\frac{1}{p} \left( \frac{1}{2} + x + \dots + x^{p-1} + \frac{1}{p} x^p \right) = \underbrace{\frac{1}{2p} (1+x) (1+x+\dots+x^{p-1})}_{Q(x)}. \quad (2.2.15)$$

Par la Proposition 2.9,  $M$  annule les variations saisonnières de période  $P$  et conserve les polynômes de degré 1.

**Définition 2.10.** Soit  $p \in \mathbb{N}^*$ . La moyenne mobile arithmétique  $M_p^{(A)}$  est la moyenne mobile centrée et symétrique d'ordres  $m_- + m_+ =: m_p^{(A)}$  et de paramètre  $\theta_p^{(A)}$ , où

$$m_p^{(A)} = \begin{cases} \frac{p}{2} \text{ et } \theta_p^{(A)} = \left( \frac{1}{2p}, \frac{1}{p}, \dots, \frac{1}{p}, \frac{1}{2p} \right) \in \mathbb{R}^{p+1}, & \text{si } p \text{ est pair} \\ \frac{p-1}{2} \text{ et } \theta_p^{(A)} = \left( \frac{1}{p}, \dots, \frac{1}{p} \right) \in \mathbb{R}^p, & \text{si } p \text{ est impair.} \end{cases} \quad (2.2.16)$$

**Exercice 7.** Soit  $(x_t)_{t \in \mathbb{Z}}$  la suite définie pour tout  $t \in \mathbb{Z}$  par  $x_t = \exp(t)$ . Donner l'expression de la suite  $(Mx_t)_{t \in \mathbb{Z}}$  où  $M$  est une moyenne mobile arithmétique d'ordre  $p$  impair.

**Exercice 8.** Soit le processus stochastique  $(X_t)_{t \in \mathbb{Z}}$  défini par

$$X_t = \beta_1 + \beta_2 t + \beta_3 \cos\left(\frac{2\pi t}{3}\right) + \varepsilon_t,$$

où  $(\varepsilon_t)_{t \in \mathbb{Z}}$  sont des variables aléatoires *i.i.d.* centrées et de variance  $\sigma^2$ .

1. Proposer une moyenne mobile  $M$  pour laquelle  $MX_t = \beta_1 + \beta_2 t + \eta_t$  avec  $\eta_t = M\varepsilon_t$ .
2. Calculer l'espérance et la variance de  $\eta_t$ . Calculer la covariance entre  $\eta_t$  et  $\eta_s$ .

Regardons à présent ce que donne la transformation de l'erreur  $(\varepsilon_t)_{t \in \mathbb{Z}}$  par une moyenne mobile.

**Exercice 9.** Soit  $(\varepsilon_t)_{t \in \mathbb{Z}}$  une suite de variables aléatoires centrées telles que pour tout  $(t, t') \in \mathbb{Z}^2$ ,  $\mathbb{E}[\varepsilon_t \varepsilon_{t'}] = \sigma^2 \delta_{t,t'} < \infty$ . Soit  $M_{(m_-, m_+, \theta)}$  une moyenne mobile avec  $\theta \in \mathbb{R}^{m_- + m_+ + 1}$ . Soit  $Y_t = MX_t$ .

1. Déterminer l'espérance et la variance de  $Y_t$ .
2. On suppose de plus que  $\sum_{i=-m_-}^{m_+} \theta_i = 1$ . Donner une minoration de  $\mathbb{E}[Y_t^2]$ .

### 2.2.1 Estimation de la tendance et de la moyenne saisonnière

On considère le modèle additif

$$X_t = T(t) + S(t; p) + \varepsilon_t. \quad (2.2.17)$$

On cherche à estimer, à partir d'observations  $(x_t)_{t=1}^T$ , la tendance  $T(\cdot)$  et les  $p$  coefficients saisonniers dont la somme fait 0.

**Étape 1** : choix de la moyenne mobile.

On applique à  $(x_t)_{t=1}^T$  une moyenne mobile  $M$  d'ordres  $(m_-, m_+)$  telle que

$$MT(t) = T(t) \quad \text{tendance conservée} \quad (2.2.18)$$

$$MS(t; p) = 0 \quad \text{variation saisonnière supprimée} \quad (2.2.19)$$

Il faut avoir formulé des hypothèses sur  $S$  et  $T$ , par exemple,  $T(t) = a + bt$  et  $S$  périodique de période 12, et dans ce cas on prendrait la moyenne arithmétique.

Pour  $t \in \llbracket m_- + 1, T - m_+ \rrbracket$ , on note  $x_t^* = Mx_t$ .

**Étape 2** : estimation des composantes saisonnières

La série corrigée de la tendance est donnée par

$$\tilde{S}(t; p) = x_t - x_t^*, \quad t \in \llbracket m_- + 1, T - m_+ \rrbracket. \quad (2.2.20)$$

On peut prendre (si  $T - m_+ - m_- \geq p$ )

$$\tilde{S}_k^{(p)} = \sum_{t=m_-+1}^{m_+} \tilde{S}(t; p) \mathbf{1}_{\{j(t,p)=k\}} / \sum_{t=m_-+1}^{m_+} \mathbf{1}_{\{j(t,p)=k\}}. \quad (2.2.21)$$

Afin d'avoir la compensation sur une période, on prend

$$\hat{S}_k^{(p)} = \tilde{S}_k^{(p)} - \frac{1}{p} \sum_{j=1}^p \tilde{S}_j^{(p)}, \quad k \in \llbracket 1, p \rrbracket. \quad (2.2.22)$$

**Étape 3** : estimation de la tendance

La série corrigée des valeurs saisonnières (CVS) est définie pour  $t \in \llbracket 1, T \rrbracket$  par

$$x_t^{\text{CVS}} = x_t - \hat{S}_{j(t,p)}^{(p)}. \quad (2.2.23)$$

Pour estimer la tendance, on impose une forme paramétrique de la forme  $T(t) = g(t, \beta)$ , où  $g(\cdot, \cdot)$  est connue et  $\beta \in \mathbb{R}^q$  inconnu. Il suffit donc d'estimer  $\beta$ , par exemple avec les moindres carrés :

$$\hat{\beta} = \underset{\beta \in \mathbb{R}^q}{\operatorname{argmin}} \sum_{t=1}^T (x_t^{\text{CVS}} - g(t, \beta))^2 \quad (2.2.24)$$

et on estime la tendance par  $\hat{T}(t) = g(t, \hat{\beta})$ .

On peut prédire la valeur de la série à l'horizon  $h \in \mathbb{N}^*$  par

$$\widehat{x}_{T+h} = \widehat{T}(T+h) + \widehat{S}_{j(T+h,p)}^{(p)} \quad (2.2.25)$$

mais on ne peut en général rien dire sur le comportement asymptotique de ces estimateurs. On peut tout de même estimer les valeurs  $\{\varepsilon_1, \dots, \varepsilon_T\}$  de l'erreur par

$$\widehat{\varepsilon}_t = x_t - \left( \widehat{T}(t) + \widehat{S}(t,p) \right), \quad t \in \llbracket 1, T \rrbracket. \quad (2.2.26)$$

En étudiant les résidus (cf. Section 3), on peut savoir si ceux-ci contiennent encore une partie saisonnière ou tendancielle, auquel cas on peut remettre le modèle additif en question.

### 2.3 Estimation par lissage exponentiel

Il s'agit d'une méthode non paramétrique permettant, à partir d'observations  $(x_t)_{t=1}^T$  d'une série temporelle, d'effectuer la prédiction à l'instant  $T+1$ . D'un point de vue théorique, les séries ne doivent pas présenter de tendance ni de variation saisonnière.

**Définition 2.11.** Soit  $\{x_t, t \in \mathbb{Z}, t \leq T\}$  une série bornée, i.e.,

$$\sup_{t \in \mathbb{Z}: t \leq T} |x_t| < \infty. \quad (2.3.1)$$

La prévision à l'instant  $T+1$  par la méthode du lissage exponentiel simple (LES) est donnée par

$$x_{T+1}^{\text{LES}}(\gamma) = (1 - \gamma) \sum_{k=0}^{\infty} \gamma^k x_{T-k}, \quad (2.3.2)$$

où  $\gamma \in ]0, 1[$  est un paramètre de lissage.

*Remarque 2.12.* Le LES a pour idée de prédire la valeur à l'instant  $T+1$  par une moyenne pondérée (car  $\sum_{k=0}^{\infty} (1 - \gamma) \gamma^k = 1$ ).

*Remarque 2.13.* Plus  $\gamma$  est proche de 0, plus le poids des valeurs passées est faible.

On utilise la formule de mise à jour

$$x_{t+1}^{\text{LES}}(\gamma) = \gamma x_t^{\text{LES}}(\gamma) + (1 - \gamma) x_t, \quad t \leq T. \quad (2.3.3)$$

Si  $\gamma = 0$ ,  $x_{t+1}^{\text{LES}}(\gamma) = 0$  : seul le dernier instant compte. Si  $\gamma = 1$ ,  $x_{t+1}^{\text{LES}}(\gamma) = x_t^{\text{LES}}(\gamma)$  (la prévision ne dépend pas du temps).

Obstacle :  $x_{T+1}^{\text{LES}}(\gamma)$  fait appel à une infinité d'observations, ce dont on ne dispose pas en pratique. Pour remédier à ce problème, on procède de la manière suivante.

- On dispose d'observations  $\{x_1, \dots, x_T\}$ .
- On initialise la formule de mise à jour par  $x_1^{\text{LES}}(\gamma) = x_0$ .
- On applique ensuite pour  $t \in \llbracket 1, T \rrbracket$  la formule de mise à jour

$$x_{t+1}^{\text{LES}}(\gamma) = \gamma x_t^{\text{LES}}(\gamma) + (1 - \gamma) x_t, \quad t \leq T. \quad (2.3.4)$$

La série  $\{x_t^{\text{LES}}(\gamma), t \in \llbracket 1, T+1 \rrbracket\}$ , est appelée série lissée. On peut l'utiliser pour

- travailler avec une série plus propre que celle de départ
- estimer une valeur manquante dans la série initiale.

Comment choisir  $\gamma$ ? On peut prendre  $\gamma$  minimisant la distance entre les valeurs observées de la série et les valeurs lissées par la méthode du lissage exponentiel :

$$\hat{\gamma} = \operatorname{argmin}_{\gamma \in ]0,1[} \sum_{t=1}^{\infty} (x_t^{\text{LES}}(\gamma) - x_t)^2. \quad (2.3.5)$$

La justification théorique est la suivante.

**Proposition 2.14.** *Soit  $\{x_t, t \leq T\}$  une série bornée et soit  $\gamma \in ]0,1[$ . Alors*

$$x_{T+1}^{\text{LES}}(\gamma) = \operatorname{argmin}_{b \in \mathbb{R}} \sum_{k=0}^{\infty} \gamma^k (x_{T-k} - b)^2. \quad (2.3.6)$$

$x_{T+1}^{\text{LES}}(\gamma)$  est la meilleure prévision au sens des moindres carrés de la série constante.

Le LES suppose l'absence de composante tendancielle dans la série.

**Exercice 10.** Soit la série  $\{x_t, t \leq T\}$ .

1. Rappeler la définition de la prévision à l'instant  $T+1$  par la méthode de lissage exponentiel simple.
2. Démontrer la formule de mise à jour de cette méthode.

**Définition 2.15.** *Soit  $\{x_t, t \leq T\}$  une série bornée et soit  $\gamma \in ]0,1[$ . La prévision de cette série à l'horizon  $h \in \mathbb{N}^*$  par la méthode du lissage exponentiel double (LED) est donnée par :*

$$x_{T+h}^{(\text{LED})}(\gamma) = \hat{a}_T(\gamma)h + \hat{b}_T, \quad (2.3.7)$$

avec

$$(\hat{a}_T, \hat{b}_T) = \operatorname{argmin}_{(a,b) \in \mathbb{R}^2} \sum_{k=0}^{\infty} \gamma^k (x_{T-k} + ak - b)^2. \quad (2.3.8)$$

Idée : on ajuste à la série de départ  $\{x_t, t \leq T\}$  la série  $\{a(T-t) + b, t \leq T\}$  à l'aide de la méthode des moindres carrés pondérés.

Une éventuelle tendance est prise en compte, mais il faut tout de même que la série ne présente pas de composante saisonnière.

**Proposition 2.16.** *On pose*

$$L_{T+1}^{(1)}(\gamma) = (1-\gamma) \sum_{k=0}^{\infty} \gamma^k x_{T-k}, \quad (2.3.9)$$

$$L_{T+1}^{(2)}(\gamma) = (1-\gamma) \sum_{k=0}^{\infty} \gamma^k L_{T+1-k}^{(1)}. \quad (2.3.10)$$

Les paramètres du lissage exponentiel double sont donnés par

$$\widehat{a}_T(\gamma) = \frac{1-\gamma}{\gamma} \left( L_{T+1}^{(1)}(\gamma) - L_{T+1}^{(2)}(\gamma) \right) \quad (2.3.11)$$

$$\widehat{b}_T(\gamma) = 2 \left( L_{T+1}^{(1)}(\gamma) - L_{T+1}^{(2)}(\gamma) \right) \quad (2.3.12)$$

On effectue deux lissages exponentiels successifs :

- le premier sur la série  $\{x_t, t \leq T\}$ ,
- le second sur la série lissée  $\left\{ L_t^{(1)}(\gamma), t \leq T+1 \right\}$ .

**Proposition 2.17.** *On a également les formules de mise à jour suivantes.*

$$\widehat{b}_t(\gamma) = (1-\gamma^2)x_t + \gamma^2 \left( \widehat{a}_{t-1}(\gamma) - \widehat{b}_{t-1}(\gamma) \right) \quad (2.3.13)$$

$$\widehat{a}_t(\gamma) = \frac{2\gamma}{1+\gamma} \widehat{a}_{t-1}(\gamma) + \frac{1-\gamma}{1+\gamma} \left( \widehat{b}_t(\gamma) - \widehat{b}_{t-1}(\gamma) \right). \quad (2.3.14)$$

### 3 Processus stochastiques à temps discret

Dans la section précédente, nous avons traité la partie déterministe d'une série temporelle de la forme  $X_t = f(t) + \varepsilon_t$ , où la partie déterministe (tendance et composante saisonnière) est représentée par  $f(\cdot)$  et  $\varepsilon_t$  la partie aléatoire. Nous n'avons pas fait d'hypothèse sur cette dernière pour estimer  $f$ . En notant  $\widehat{f}_T(\cdot)$  l'estimation de  $f(\cdot)$  obtenue, on dispose ainsi d'estimations

$$\widehat{\varepsilon}_t := X_t - \widehat{f}_T(t), \quad t \in \llbracket 1, T \rrbracket, \quad (3.0.1)$$

des variables aléatoires  $\varepsilon_1, \dots, \varepsilon_T$ . La collection des  $\widehat{\varepsilon}_t$  est appelée résidus.

Deux cas de figure se présentent.

1. Des tests statistiques laissent à penser que les résidus sont issus de variables aléatoires non corrélées. Dans ce cas,  $\widehat{f}_T(\cdot)$  est une bonne estimation de la tendance.
2. Si on rejette à l'aide de ces test l'hypothèse que les résidus proviennent de réalisations de variables aléatoires non-corrélées. De l'information est encore contenue dans les résidus, pour lesquels une étude plus approfondie est nécessaire.

#### 3.1 Introduction aux processus stochastiques

Nous allons étudier les termes aléatoires d'une série temporelle. Ils seront notés, pour faciliter les notations,  $(X_t)_{t \in \mathbb{Z}}$  au lieu de  $(\varepsilon_t)_{t \in \mathbb{Z}}$ .

Nous supposons que les processus étudiés sont de carré intégrable et nous ferons des rappels sur les espaces  $\mathbb{L}^2$ .

**Définition 3.1.** *Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé.*

1. On note  $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$  l'espace des variables aléatoires  $X: (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  telles que

$$\int_{\Omega} X(\omega)^2 d\mathbb{P}(\omega) = \mathbb{E}[X^2] < \infty. \quad (3.1.1)$$

2. On note  $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$  l'ensemble des classes d'équivalence pour l'égalité presque sûre des variables aléatoires de  $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ .

Un élément de  $\mathbb{L}^2$  est donc une collection de variables aléatoires qui sont toutes égales presque sûrement. On rappelle que  $X = Y$  p.s. s'il existe un ensemble  $N \in \mathcal{F}$  tel que  $\{\omega \in \Omega, X(\omega) \neq Y(\omega)\} \subset N$  et  $\mathbb{P}(N) = 0$ .

Par abus de langage, on dira que  $X \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$  alors que  $X$  est en réalité un représentant de la classe.

**Proposition 3.2.** *L'ensemble  $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$  est un espace vectoriel sur  $\mathbb{R}$ . De plus, l'application  $\langle \cdot, \cdot \rangle$  définie pour  $X, Y \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$  par*

$$\langle X, Y \rangle := \int_{\Omega} X(\omega) Y(\omega) d\mathbb{P}(\omega) = \mathbb{E}[XY] \quad (3.1.2)$$

est une forme hermitienne définie positive, donc un produit scalaire.

La norme induite par le produit scalaire est notée  $\|\cdot\|_2$  :

$$\|X\|_2 = \sqrt{\langle X, X \rangle} = \sqrt{\mathbb{E}[X^2]}. \quad (3.1.3)$$

**Proposition 3.3** (Inégalité de Cauchy-Schwarz). *Soient  $X, Y \in \mathbb{L}^2$ . Alors*

$$|\langle X, Y \rangle| \leq \|X\|_2 \|Y\|_2, \quad (3.1.4)$$

ou, de manière équivalente,

$$(\mathbb{E}[XY])^2 \leq \mathbb{E}[X^2] \mathbb{E}[Y^2]. \quad (3.1.5)$$

**Définition 3.4.** *Soit  $(X_n)_{n \geq 1}$  une suite d'éléments de  $\mathbb{L}^2$  et soit  $X \in \mathbb{L}^2$ . On dit que  $(X_n)_{n \geq 1}$  converge vers  $X$  dans  $\mathbb{L}^2$  si*

$$\lim_{n \rightarrow \infty} \|X_n - X\|_2 = 0 \quad (3.1.6)$$

ou de manière équivalente, si

$$\lim_{n \rightarrow \infty} \mathbb{E}[(X_n - X)^2] = 0. \quad (3.1.7)$$

Le produit scalaire est continu. Plus précisément :

**Proposition 3.5.** *Soient  $(X_n)_{n \geq 1}$  et  $(Y_n)_{n \geq 1}$  des suites d'éléments de  $\mathbb{L}^2$ , convergent respectivement vers  $X \in \mathbb{L}^2$  et  $Y \in \mathbb{L}^2$ . Alors  $\lim_{n \rightarrow \infty} \langle X_n, Y_n \rangle = \langle X, Y \rangle$ .*

**Définition 3.6.** *On dit que  $X, Y \in \mathbb{L}^2$  sont orthogonales si  $\langle X, Y \rangle = 0$ , ou de manière équivalente,  $\mathbb{E}[XY] = 0$ .*

**Théorème 3.7.** *L'espace  $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$  est un espace de Hilbert.*

**Théorème 3.8** (Théorème de la projection orthogonale). *Soit  $\mathbb{H}$  un espace de Hilbert et  $F$  un sous-ensemble convexe fermé. Pour tout  $x \in \mathbb{H}$ , il existe un unique  $p_F(x) \in F$  tel que*

$$\|x - P_F(x)\|_2 = \inf_{y \in F} \|x - y\|_2. \quad (3.1.8)$$

*Si  $F$  est de plus un sous-espace vectoriel fermé, le point  $p_F(x)$  est l'unique point de  $F$  tel que*

$$x - p_F(x) \in F^\perp := \{u \in \mathbb{H}, \forall v \in F, \langle u, v \rangle = 0\}. \quad (3.1.9)$$

La projection orthogonale possède les propriétés suivantes.

**Proposition 3.9.** *Soit  $\mathbb{H}$  un espace de Hilbert et soit  $F$  un sous-espace vectoriel fermé.*

1. *L'application  $p_F: \mathbb{H} \rightarrow F$  est linéaire.*
2. *Si  $x \in F^\perp$ , alors  $p_F(x) = 0$ .*
3. *Si  $F$  est la somme de deux sous-espaces fermés orthogonaux  $F_1$  et  $F_2$ , alors  $P_F(x) = P_{F_1}(x) + P_{F_2}(x)$ .*
4. *Si  $F_1$  et  $F_2$  sont des sous-espaces vectoriels fermés tels que  $F_1 \subset F_2$ , alors pour tout  $x \in \mathbb{H}$ ,*

$$P_{F_1}(P_{F_2}(x)) = P_{F_2}(P_{F_1}(x)) = P_{F_1}(x). \quad (3.1.10)$$

**Définition 3.10.** *Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé et soit  $E \subset \mathbb{R}$ . Pour tout  $t \in E$ , soit  $X_t: (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  une variable aléatoire. La famille  $(X_t)_{t \in E}$  est appelée processus stochastique d'espace d'états  $E$ .*

1. *Si  $E$  est dénombrable, on dit que le processus  $(X_t)_{t \in E}$  est à temps discret.*
2. *Si  $E$  n'est pas dénombrable, on dit que le processus  $(X_t)_{t \in E}$  est à temps continu.*

Les séries temporelles correspondent à des réalisations de processus à temps discret. Dans la suite, on s'intéressera uniquement au cas discret.

Les processus sont des collections infinies de variables aléatoires.

On peut voir un processus  $(X_t)_{t \in E}$  comme une application de  $\Omega$  dans  $\mathcal{T}(E, \mathbb{R})$  qui est l'ensemble des applications de  $E$  dans  $\mathbb{R}$ . Pour définir correctement une variable aléatoire à valeurs dans  $\mathcal{T}(E, \mathbb{R})$ , il faut mettre une tribu sur cet espace. Si  $E = \mathbb{N}$ , on peut prendre la tribu produit de la tribu borélienne, qui est engendrée par la tribu cylindrique  $\mathcal{C}$ , définie par

$$\mathcal{C} = \{f \in \mathcal{T}(\mathbb{N}, \mathbb{R}), (f(0), \dots, f(k)) \in B_1 \times B_2 \times \dots \times B_{k+1}\}. \quad (3.1.11)$$

**Définition 3.11.** *Soit  $(X_t)_{t \in E}$  un processus stochastique défini sur  $(\Omega, \mathcal{F}, \mathbb{P})$ . Les lois finidimensionnelles de ce processus sont les mesure  $\mu_{t_1, \dots, t_k}$  sur  $\mathcal{B}(\mathbb{R}^k)$  définies pour  $k \in \mathbb{N}^*$ ,  $t_1, \dots, t_k \in E$  et  $B_1, \dots, B_k \in \mathcal{B}(\mathbb{R})$  par*

$$\mu_{t_1, \dots, t_k}(B_1 \times \dots \times B_k) = \mathbb{P}(\omega, \forall \ell \in \llbracket 1, k \rrbracket, X_{t_\ell}(\omega) \in B_\ell). \quad (3.1.12)$$

**Proposition 3.12.** *Si les processus  $(X_t)_{t \in E}$  et  $(Y_t)_{t \in E}$  ont les mêmes lois fini-dimensionnelles, alors ils ont la même loi.*

**Définition 3.13.** *On note*

$$\mathcal{M} := \{\mu_{t_1, \dots, t_k}, k \in \mathbb{N}^*, t_1, \dots, t_k \in E\}, \quad (3.1.13)$$

où pour chaque  $t_1, \dots, t_k$ ,  $\mu_{t_1, \dots, t_k}$  est une mesure sur  $\mathcal{B}(\mathbb{R}^k)$ . La famille de mesures  $\mathcal{M}$  est consistante si pour tout  $k \geq 1$  et tous  $B_1, \dots, B_k \in \mathcal{B}(\mathbb{R})$

1. l'égalité

$$\mu_{t_1, \dots, t_k}(B_1 \times \dots \times B_k) = \mu_{t_{\sigma(1)}, \dots, t_{\sigma(k)}}(B_{\sigma(1)}, \dots, B_{\sigma(k)}) \quad (3.1.14)$$

a lieu pour toute permutation  $\sigma: \llbracket 1, k \rrbracket \rightarrow \llbracket 1, k \rrbracket$  et

2. pour tout  $m \geq k$  et tous  $t_1, \dots, t_m$ ,

$$\mu_{t_1, \dots, t_k, t_{k+1}, \dots, t_m}(B_1 \times \dots \times B_k \times \mathbb{R} \times \dots \times \mathbb{R}) = \mu_{t_1, \dots, t_k}(B_1 \times \dots \times B_k). \quad (3.1.15)$$

**Théorème 3.14.** *Si la famille  $\mathcal{M}$  est consistante, alors il existe un processus stochastique  $(X_t)_{t \in E}$  défini sur  $(\Omega, \mathcal{F}, \mathbb{P})$  tel que pour tous  $k \in \mathbb{N}^*$ ,  $t_1, \dots, t_k \in E$  et  $B_1, \dots, B_k \in \mathcal{B}(\mathbb{R})$  par*

$$\mu_{t_1, \dots, t_k}(B_1 \times \dots \times B_k) = \mathbb{P}(\omega, \forall \ell \in \llbracket 1, k \rrbracket, X_{t_\ell}(\omega) \in B_\ell). \quad (3.1.16)$$

**Exercice 11.** Démontrer que la famille des lois fini-dimensionnelles d'un processus stochastique  $(X_t)_{t \in E}$  forme une famille consistante.

Autrement dit, il est possible de définir un processus stochastique à partir d'une famille de loi consistante.

**Définition 3.15.** *On dit qu'un processus  $(X_t)_{t \in \mathbb{Z}}$  est ergodique en moyenne quadratique si pour chaque  $t$ ,  $X_t$  est de carré intégrable et  $\mathbb{E}[X_t] = \mathbb{E}[X_0]$  et si pour chaque  $t_0 \in \mathbb{Z}$ ,*

$$\lim_{T \rightarrow \infty} \mathbb{E} \left[ \left( \frac{1}{T} \sum_{t=t_0}^{T+t_0-1} X_t - \mathbb{E}[X_0] \right)^2 \right] = 0. \quad (3.1.17)$$

La propriété d'ergodicité en moyenne quadratique garantit la convergence des moyennes empiriques d'un échantillon  $(X_{t_0}, \dots, X_{t_0+T-1})$  de taille  $T$  vers l'espérance commune.

**Exercice 12.** Soit  $(X_t)_{t \in \mathbb{Z}}$  tel que  $\text{Cov}(X_s, X_t) = \sigma^2 \delta_{s,t}$ , et  $\mathbb{E}[X_t] = \mathbb{E}[X_0]$ . Est-ce que  $(X_t)_{t \in \mathbb{Z}}$  est ergodique en moyenne quadratique ?

## 3.2 Exemples de processus stochastiques

**Définition 3.16.** *Un processus stochastique  $(\eta_t)_{t \in \mathbb{Z}}$  est appelé bruit blanc fort si la collection de variables aléatoires  $\{\eta_t, t \in \mathbb{Z}\}$  est indépendante et pour tout  $t \in \mathbb{Z}$ ,  $\mathbb{E}[\eta_t^2] = \sigma^2$  et  $\mathbb{E}[\eta_t] = 0$ .*

**Définition 3.17.** Un processus stochastique  $(\eta_t)_{t \in \mathbb{Z}}$  est appelé bruit blanc faible si pour tout  $t \in \mathbb{Z}$ ,  $\mathbb{E}[\eta_t^2] = \sigma^2$  et  $\mathbb{E}[\eta_t] = 0$ , et pour tous  $s, t \in \mathbb{Z}$  tels que  $s \neq t$ ,  $\text{Cov}(\eta_s, \eta_t) = 0$ .

Dans la suite, "bruit blanc" fera référence au bruit blanc faible.

**Définition 3.18.** On dit que le processus stochastique  $(X_t)_{t \in \mathbb{Z}}$  est stationnaire au second ordre si pour tout  $t \in \mathbb{Z}$ ,  $\mathbb{E}[X_t^2]$  est finie,  $\mathbb{E}[X_t] = \mathbb{E}[X_0]$  et pour tous  $h, t \in \mathbb{Z}$ ,

$$\text{Cov}(X_t, X_{t+h}) = \text{Cov}(X_0, X_h) =: \gamma(h). \quad (3.2.1)$$

Remarquons que si  $(X_t)_{t \in \mathbb{Z}}$  est stationnaire au second ordre, alors en prenant  $h = 0$ ,  $\mathbb{E}[X_t^2] = \mathbb{E}[X_0^2] = \gamma(0)$ . Ainsi, les variables aléatoires  $X_t$ ,  $t \in \mathbb{Z}$ , ont toutes la même variance.

**Définition 3.19.** La fonction  $\gamma(\cdot)$  définie par (3.2.1) est appelée fonction d'auto-covariance.

**Définition 3.20.** On dit que  $(X_t)_{t \in \mathbb{Z}}$  est un processus stationnaire au sens strict si pour tous  $h, k \in \mathbb{N}^*$ , tous  $t_1, \dots, t_k \in \mathbb{Z}$ , les vecteurs  $(X_{t_1}, \dots, X_{t_k})$  et  $(X_{t_1+h}, \dots, X_{t_k+h})$  ont la même loi.

**Exercice 13.** Démontrer que si  $(X_t)_{t \in \mathbb{Z}}$  est stationnaire au sens strict avec  $X_0 \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ , alors  $(X_t)_{t \in \mathbb{Z}}$  est stationnaire au second ordre.

**Exercice 14.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus indépendant, où  $X_{2k}$  prend les valeurs  $-1$  et  $1$  avec probabilité  $1/2$  et  $X_{2k+1}$  suit une loi normale centrée réduite. Étudier la stationnarité au second ordre et au sens large de  $(X_t)_{t \in \mathbb{Z}}$ .

**Définition 3.21.** La marche aléatoire est un processus  $(X_t)_t$  défini par  $X_t = 0$  pour  $t \leq 0$  et  $X_t = \sum_{s=1}^t \varepsilon_s$ , où  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible.

La marche aléatoire n'est pas un processus ergodique en moyenne quadratique ni un processus stationnaire au second ordre, sauf si  $\sigma = 0$ .

Dans le cas où  $\varepsilon_t$  est tel que  $\mathbb{P}(\varepsilon_t = 1) = p \in ]0, 1[$  et  $\mathbb{P}(\varepsilon_t = -1) = 1 - p$ ,  $X_t$  représente le gain d'un joueur après  $T$  tours de pile ou face.

**Exercice 15.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire de moyenne  $\mu \in \mathbb{R}$  et de fonction de covariance  $\gamma(\cdot)$ .

1. On considère le processus  $(Y_t)_{t \in \mathbb{Z}} := ((-1)^t X_t)_{t \in \mathbb{Z}}$ . Donner la valeur de  $\mu$  pour laquelle le processus  $(Y_t)_{t \in \mathbb{Z}}$  est stationnaire au second ordre.
2. En prenant pour  $\mu$  la solution de la question 1), montrer que le processus  $(Z_t)_{t \in \mathbb{Z}} := (X_t + Y_t)_{t \in \mathbb{Z}}$  n'est pas stationnaire. Ceci montre que la somme de deux processus stationnaires n'est pas forcément un processus stationnaire.

**Exercice 16.** Soit  $(\eta_t)_{t \in \mathbb{Z}}$  un bruit blanc faible de variance  $\sigma^2$  et tel que  $\mathbb{E}[\eta_t^4] < \infty$  pour tout  $t \in \mathbb{Z}$ .

1. Le processus  $(X_t)_{t \in \mathbb{Z}}$  défini par  $X_t = \eta_t^2 + \eta_{t-1}/2 - \sigma^2$  est-il un processus stationnaire ?
2. Qu'en est-il si  $(\eta_t)_{t \in \mathbb{Z}}$  est un processus gaussien centré et tel que  $\mathbb{E}[\eta_t \eta_s] = \sigma^2 \delta_{t,s}$  ?

**Définition 3.22.** Un processus  $(X_t)_{t \in \mathbb{Z}}$  est gaussien si pour tout  $k$ , tous  $t_1, \dots, t_k \in \mathbb{Z}$  et tous  $c_1, \dots, c_k \in \mathbb{R}$ , la variable aléatoire  $\sum_{\ell=1}^k c_\ell X_{t_\ell}$  est de loi normale (de variance potentiellement nulle).

**Exercice 17.** Soit  $X$  une variable aléatoire de loi normale centrée et réduite. Soit  $(\varepsilon_t)_{t \in \mathbb{Z}}$  un processus pour lequel les variables aléatoires  $\{\varepsilon_t, ; t \in \mathbb{Z}\}$  sont indépendantes et telles que pour tout  $t \in \mathbb{Z}$ ,  $\mathbb{P}(\varepsilon_t = 1) = \mathbb{P}(\varepsilon_t = -1) = 1/2$ . On suppose enfin que le processus  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est indépendant de  $X$ . On pose enfin  $Y_t = X\varepsilon_t$ .

1. Montrer que le processus  $(Y_t)_{t \in \mathbb{Z}}$  est strictement stationnaire.
2. Pour tout  $t \in \mathbb{Z}$ , donner la loi de  $Y_t$ .
3. Le processus  $(Y_t)_{t \in \mathbb{Z}}$  est-il gaussien ?

**Définition 3.23.** Un processus  $(B_n)_{n \in \mathbb{N}}$  est un mouvement brownien standard si

1.  $B_0 = 0$  presque sûrement,
2. pour tous  $s, t \in \mathbb{N}$  tels que  $t > s$ ,  $B_t - B_s$  est de loi normale centrée et de variance  $t - s$  et
3. pour tous  $s, t \in \mathbb{N}$  tels que  $t > s$ ,  $B_t - B_s$  est indépendant de  $(B_u)_{u=1}^s$ .

**Exercice 18.** 1. Rappeler la définition d'un processus gaussien.

2. Démontrer que le mouvement brownien standard  $(B_n)_{n \in \mathbb{N}}$  est gaussien et vérifie  $\text{Cov}(B_s, B_t) = \min\{s, t\}$ .

**Définition 3.24.** On dit qu'une suite de réels  $(\Phi_t)_{t \in \mathbb{Z}}$  est absolument sommable si  $\sum_{t \in \mathbb{Z}} |\Phi_t| < \infty$ .

**Proposition 3.25.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stochastique défini sur l'espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$ . On suppose que  $\sup_{t \in \mathbb{Z}} \mathbb{E}[|X_t|] < \infty$ . Soit  $(\Phi_t)_{t \in \mathbb{Z}}$  une suite absolument sommable. Alors pour chaque  $t \in \mathbb{Z}$ , la suite de processus  $(Y_{n,t})_{n \geq 1}$  définie par

$$Y_{n,t} = \sum_{i=-n}^n \phi_i B^i X_t = \sum_{i=-n}^n \phi_i X_{t+i} \quad (3.2.2)$$

converge presque sûrement, quand  $n$  tend vers l'infini, vers

$$Y_t := \sum_{i \in \mathbb{Z}} \phi_i X_{t+i}. \quad (3.2.3)$$

Si on suppose de plus que  $\sup_{t \in \mathbb{Z}} \mathbb{E}[X_t^2] < \infty$ , alors la convergence a lieu dans  $\mathbb{L}^2$ .

**Définition 3.26.** Le processus  $(Y_t)_{t \in \mathbb{Z}}$  défini par (3.2.3) est appelé moyenne mobile infinie.

### 3.3 Espérance linéaire et innovation

**Définition 3.27.** Soit  $I$  un ensemble au plus dénombrable et soit  $(X_i)_{i \in I}$  une famille de variables aléatoires telles que pour tout  $i \in I$ ,  $X_i \in \mathbb{L}^2$ . On note  $\overline{\text{vect}} \{X_i, i \in I\}$  le plus petit (pour l'inclusion) sous-espace vectoriel de  $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$  contenant toutes les combinaisons linéaires finies de variables aléatoires prises dans l'ensemble  $\{X_i, i \in I\}$ .

**Définition 3.28.** L'espérance linéaire d'une variable aléatoire  $Y \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$  sur l'ensemble  $(X_i)_{i \in I}$  est notée  $\mathbb{E}\mathbb{L}(Y \mid (X_i)_{i \in I})$  est la projection orthogonale de  $Y$  sur le sous-espace vectoriel fermé  $\overline{\text{vect}} \{X_i, i \in I\}$ .

On a donc

$$\|Y - \mathbb{E}\mathbb{L}(Y \mid (X_i)_{i \in I})\|_2 = \inf \{\|Y - Z\|_2, Z \in \overline{\text{vect}} \{X_i, i \in I\}\}. \quad (3.3.1)$$

**Exercice 19.** Soient  $X$  et  $Y$  deux variables aléatoires de carré intégrable. Déterminer  $\mathbb{E}\mathbb{L}(Y \mid X)$  et  $\mathbb{E}\mathbb{L}(Y \mid 1, X)$ .

Si  $I = (i_\ell)_{\ell \geq 1}$  est infini dénombrable, alors

$$\mathbb{E}\mathbb{L}(Y \mid (X_i)_{i \in I}) = \lim_{N \rightarrow \infty} \mathbb{E}\mathbb{L}(Y \mid (X_{i_\ell})_{\ell=1}^N), \quad (3.3.2)$$

où la limite est prise dans  $\mathbb{L}^2$ . Par le Théorème de la projection orthogonale, la variable aléatoire  $U$  définie par

$$U = Y - \mathbb{E}\mathbb{L}(Y \mid (X_i)_{i \in I}) \quad (3.3.3)$$

appartient à l'orthogonal de  $\overline{\text{vect}} \{X_i, i \in I\}$ .

Si il existe  $i \in I$  tel que  $X_i$  est constante, alors  $U$  est orthogonale à une constante et donc nécessairement centrée.

Soient  $Y$  et  $X_i, i \in I$  des variables aléatoires de carré intégrable. On note  $Y' := Y - \mathbb{E}[Y]$  et  $X'_i = X_i - \mathbb{E}[X_i]$ . On a alors

$$\mathbb{E}\mathbb{L}(Y \mid 1, (Y_i)_{i \geq 1}) = \mathbb{E}[Y] + \mathbb{E}\mathbb{L}(Y' \mid (X'_i)_{i \in I}) \quad (3.3.4)$$

Par conséquent, si les variables aléatoires  $Y$  et  $X_i$  sont centrées, on a

$$\mathbb{E}\mathbb{L}(Y \mid 1, (Y_i)_{i \geq 1}) = \mathbb{E}\mathbb{L}(Y \mid (X_i)_{i \in I}). \quad (3.3.5)$$

Pour éviter l'introduction d'une constante, on supposera donc les variables aléatoires centrées.

**Exercice 20.** On considère le processus  $(X_t)_{t \in \mathbb{Z}}$  où pour tout  $t \in \mathbb{Z}$ ,  $\mathbb{E}[X_t] = \mu \neq 0$  et  $\mathbb{E}[X_t^2] < \infty$ . Pour  $t \in \{2, 3, \dots\}$ , on définit le processus

$$Y_t = \mathbb{E}\mathbb{L}(X_t \mid \overline{\text{vect}} \{X_1, \dots, X_{t-1}\}).$$

Pour  $t \in \mathbb{N}^*$ , on introduit le vecteur aléatoire  $\mathbb{X}_t := (X_1, \dots, X_t)^\top$  et on suppose que la matrice  $S_t := \mathbb{E}[\mathbb{X}_t \mathbb{X}_t^\top]$  est inversible pour tout  $t \in \mathbb{N}^*$ .

1. Déterminer l'expression du vecteur  $\theta \in \mathbb{R}^{t-1}$  tel que  $Y_t = \theta^\top \mathbb{X}_{t-1}$ .
2. On suppose que  $(X_t)_{t \in \mathbb{Z}}$  est un processus stationnaire d'espérance  $\mu \neq 0$  et de fonction d'autocovariance

$$\gamma_X(h) = \text{Cov}(X_t, X_{t+h}) = \mathbb{E}[X_t X_{t+h}] - \mu^2.$$

pour tout  $h \in \mathbb{Z}$ . Calculer l'espérance de  $Y_t$  en fonction de  $\gamma_X(\cdot)$ . Le processus  $(Y_t)_{t \in \mathbb{Z}}$  est-il stationnaire ?

Il ne faut pas confondre l'espérance linéaire avec l'espérance conditionnelle. Si on pose  $F_1 = \overline{\text{vect}} \{X_i, i \in I\}$  et  $F_2 = \mathbb{L}^2(\Omega, \sigma(Y_i, i \in I), \mathbb{P})$  alors  $F_1 \subset F_2$  et l'inclusion est stricte en général. Dans le cas simple où  $I$  ne contient qu'un élément noté  $X \in \mathbb{L}^4$ ,  $\mathbb{E}\mathbb{L}(X^2 | X) = cX$  pour une certaine constante  $c$ , alors que  $\mathbb{E}[X^2 | X] = X^2$ .

En revanche, comment souvent, le cas gaussien est particulier.

**Proposition 3.29.** *Si  $I = \llbracket 1, n \rrbracket$  est fini, et  $(Y, X_1, \dots, X_n)$  est gaussien centré, alors*

$$\mathbb{E}\mathbb{L}(Y | (X_i)_{i=1}^n) = \mathbb{E}[Y | (X_i)_{i=1}^n]. \quad (3.3.6)$$

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus défini sur l'espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$  et tel que pour tout  $t \in \mathbb{Z}$ ,  $X_t \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ . On pose

$$\mathcal{H}_t^X = \overline{\text{vect}} \{X_s, s \in \mathbb{Z}, s \leq t\}. \quad (3.3.7)$$

L'espace  $\mathcal{H}_t^X$  est appelé histoire du processus  $(X_t)_{t \in \mathbb{Z}}$  jusqu'à l'instant  $t$ .

Remarquons que  $\mathcal{H}_t^X \subset \mathcal{H}_{t+1}^X$ . On note

$$\mathcal{H}_{-\infty}^X = \bigcap_{t \in \mathbb{Z}} \mathcal{H}_t^X \quad (3.3.8)$$

le passé lointain du processus  $(X_t)_{t \in \mathbb{Z}}$ .

**Définition 3.30.** *Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus défini sur l'espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$  et tel que pour tout  $t \in \mathbb{Z}$ ,  $X_t \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$  et  $\mathbb{E}[X_t] = 0$ . L'innovation de ce processus est le processus  $(\varepsilon_t)_{t \in \mathbb{Z}}$  donné par*

$$\varepsilon_t = X_t - \mathbb{E}\mathbb{L}(X_t | \mathcal{H}_{t-1}^X). \quad (3.3.9)$$

**Exercice 21.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire centré. Montrer que l'innovation  $(\varepsilon_t)_{t \in \mathbb{Z}}$  de  $(X_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible.

<https://plmlatex.math.cnrs.fr/project>

### 3.4 Processus défini par une équation de récurrence

On traitera le cas processus  $(X_t)_{t \in \mathbb{Z}}$  défini par une équation les liant à un processus  $(Y_t)_{t \in \mathbb{Z}}$ . Le processus  $(X_t)_{t \in \mathbb{Z}}$  est observé alors que  $(Y_t)_{t \in \mathbb{Z}}$  ne l'est pas.

On suppose que  $(X_t)_{t \in \mathbb{Z}}$  vérifie une relation de récurrence de la forme

$$P(B)X_t = Y_t, \quad (3.4.1)$$

où  $B$  est l'opérateur de retard et  $P$  un polynôme à coefficients réels.

Par exemple, si  $P(z) = 1 - z/2$ , on a  $X_t - X_{t-1}/2 = Y_t$ . La première question qui se pose est l'existence d'un tel processus. De plus, on aimerait avoir une expression de  $X_t$  en fonction de  $(Y_t)_{t \in \mathbb{Z}}$ .

**Définition 3.31.** Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé et

$$V = \left\{ (X_t)_{t \in \mathbb{Z}}, X_t : \Omega \rightarrow \mathbb{R}, \sup_{t \in \mathbb{Z}} \mathbb{E} [|X_t|] < \infty \right\}, \quad (3.4.2)$$

l'espace des processus à temps discret bornés dans  $\mathbb{L}^1$ .

**Exercice 22.** Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé et

$$V = \left\{ (X_t)_{t \in \mathbb{Z}}, X_t : \Omega \rightarrow \mathbb{R}, \sup_{t \in \mathbb{Z}} \mathbb{E} [|X_t|] < \infty \right\},$$

l'espace des processus à temps discret bornés dans  $\mathbb{L}^1$ .

1. Soit  $(X_t)_{t \in \mathbb{Z}}$  telle que pour chaque  $t$ ,  $X_t$  suit une loi de Poisson de paramètre  $\lambda_t > 0$ . Donner une condition nécessaire et suffisante sur  $(\lambda_t)_{t \in \mathbb{Z}}$  pour que la suite  $(X_t)_{t \in \mathbb{Z}}$  appartienne à  $V$ .
2. Soit  $(X_t)_{t \in \mathbb{Z}}$  telle que pour chaque  $t$ ,  $X_t$  suit une loi exponentielle de paramètre  $\lambda_t > 0$ . Donner une condition nécessaire et suffisante sur  $(\lambda_t)_{t \in \mathbb{Z}}$  pour que la suite  $(X_t)_{t \in \mathbb{Z}}$  appartienne à  $V$ .
3. Soit  $(X_t)_{t \in \mathbb{Z}}$  telle que pour chaque  $t$ ,  $X_t$  suit une loi normale d'espérance  $\mu_t$  et de variance  $\sigma_t^2 > 0$ . Donner une condition nécessaire et suffisante sur  $(\mu_t, \sigma_t)_{t \in \mathbb{Z}}$  pour que la suite  $(X_t)_{t \in \mathbb{Z}}$  appartienne à  $V$ .

**Lemme 3.32.** On pose

$$\|(X_t)_{t \in \mathbb{Z}}\|_V := \sup_{t \in \mathbb{Z}} \mathbb{E} [|X_t|]. \quad (3.4.3)$$

Alors  $V$  muni de  $\|\cdot\|_V$  est un espace de Banach.

L'intérêt de disposer d'un espace de Banach réside dans le fait que pour établir la convergence d'une série  $\sum_{i=1}^{\infty} y_i$ , avec  $y_i \in V$ , il suffit de montrer que  $\sum_{i=1}^{\infty} \|y_i\|_V < \infty$ .

**Exercice 23.** En utilisant le fait que  $(\mathbb{L}^1, \|\cdot\|_1)$  est un espace de Banach, démontrer que  $V = \{(X_t)_{t \in \mathbb{Z}}, X_t : \Omega \rightarrow \mathbb{R}, \sup_{t \in \mathbb{Z}} \mathbb{E} [|X_t|] < \infty\}$  muni de la norme  $\|(X_t)_{t \in \mathbb{Z}}\|_V := \sup_{t \in \mathbb{Z}} \mathbb{E} [|X_t|]$  est un espace de Banach.

**Lemme 3.33.** Soit  $B : V \rightarrow V$  défini par  $B((X_t)_{t \in \mathbb{Z}}) = (X_{t-1})_{t \in \mathbb{Z}}$ . Alors  $B$  est linéaire et continu. De plus,  $\|B\|_{\mathcal{L}(V)} := \sup_{\|(X_t)_{t \in \mathbb{Z}}\|_V = 1} \|B((X_t)_{t \in \mathbb{Z}})\|_V = 1$ .

**Exercice 24.** Soit  $B : V \rightarrow V$  défini par  $B((X_t)_{t \in \mathbb{Z}}) = (X_{t-1})_{t \in \mathbb{Z}}$ .

1. Démontrer que  $B$  est linéaire.
2. Démontrer que  $\|B\|_{\mathcal{L}(V)} := \sup_{\|(X_t)_{t \in \mathbb{Z}}\|_V = 1} \|B((X_t)_{t \in \mathbb{Z}})\|_V = 1$ .

**Lemme 3.34.** Soit  $\lambda \in \mathbb{C}$  tel que  $|\lambda| \neq 1$ . Soit  $P(B) = I - \lambda B : V \rightarrow V$ . Alors  $P(B)$  est inversible, d'inverse  $(P(B))^{-1}$  défini de la manière suivante.

- Si  $|\lambda| < 1$ ,

$$(P(B))^{-1} = \sum_{j=0}^{\infty} \lambda^j B^j, \quad (3.4.4)$$

- si  $|\lambda| > 1$ ,

$$(P(B))^{-1} = - \sum_{j=-\infty}^{-1} \lambda^j B^j. \quad (3.4.5)$$

**Lemme 3.35.** Soit  $\lambda \in \mathbb{C}$  tel que  $|\lambda| \neq 1$ . Soit  $P(B) = (I - \lambda B)(I - \bar{\lambda}B) : V \rightarrow V$ . Alors  $P(B)$  est inversible, d'inverse  $(P(B))^{-1}$  défini de la manière suivante.

- Si  $|\lambda| < 1$ ,

$$(P(B))^{-1} = \sum_{j=0}^{\infty} \sum_{\ell=0}^j \bar{\lambda}^{\ell} \lambda^{\ell-j} B^j, \quad (3.4.6)$$

- si  $|\lambda| > 1$ ,

$$(P(B))^{-1} = - \sum_{j=-\infty}^0 \sum_{k=j+1}^{-1} \bar{\lambda}^j \lambda^{k-j} B^j. \quad (3.4.7)$$

**Proposition 3.36.** Soit  $P$  un polynôme à coefficients réels tel que  $P(0) = 1$  et vérifiant  $P(z) \neq 0$  pour tout  $z \in \mathbb{C}$  tel que  $|z| = 1$ . Alors  $P(B) : V \rightarrow V$  est inversible. Plus précisément, il existe une famille de réels absolument sommable  $(a_j)_{j \in \mathbb{Z}}$  telle que

$$(P(B))^{-1} = \sum_{j \in \mathbb{Z}} a_j B^j. \quad (3.4.8)$$

Si on suppose de plus que  $P(\cdot)$  n'a aucune racine dans le disque unité fermé, alors  $P^{-1}(\cdot)$  admet la représentation

$$(P(B))^{-1} = \sum_{j=0}^{\infty} a_j B^j. \quad (3.4.9)$$

**Proposition 3.37.** Soit  $(Y_t)_{t \in \mathbb{Z}}$  un processus dont les moments d'ordre 1 sont uniformément bornés, c'est-à-dire

$$\sup_{t \in \mathbb{Z}} \mathbb{E}[|Y_t|] < \infty. \quad (3.4.10)$$

1. On suppose que  $P(\cdot)$  est un polynôme n'ayant aucune racine sur le cercle unité de  $\mathbb{C}$ . Alors l'équation  $P(B)X_t = Y_t$  admet une unique solution  $(X_t)_{t \in \mathbb{Z}}$  dont l'expression est donnée par

$$X_t = (P(B))^{-1} Y_t = \sum_{j \in \mathbb{Z}} a_j Y_{t-j}, \quad t \in \mathbb{Z}, \quad (3.4.11)$$

où les coefficients  $(a_j)_{j \in \mathbb{Z}}$  sont absolument sommables.

2. On suppose que  $P(\cdot)$  est un polynôme n'ayant aucune racine sur le disque unité de  $\mathbb{C}$ . Alors l'équation  $P(B)X_t = Y_t$  admet une unique solution  $(X_t)_{t \in \mathbb{Z}}$  dont l'expression est donnée par

$$X_t = (P(B))^{-1} Y_t = \sum_{j=0}^{\infty} a_j Y_{t-j}, \quad t \in \mathbb{Z}, \quad (3.4.12)$$

où les coefficients  $(a_j)_{j \geq 0}$  sont absolument sommables.

Dans le cas 1, le processus  $(X_t)_{t \in \mathbb{Z}}$  est inversible et dans le cas 2, le processus  $(X_t)_{t \in \mathbb{Z}}$  est causal. Il ne dépend que des valeurs passées du processus  $(Y_t)_{t \in \mathbb{Z}}$ .

### 3.5 Processus stationnaires au second ordre : innovation, auto-covariance, prédiction linéaire

**Exercice 25.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire au second ordre, centré, de fonction d'auto-covariance  $\gamma_X(\cdot)$  et soit  $(a_j)_{j \in \mathbb{Z}}$  une famille de réels absolument sommable. Le processus  $(Y_t)_{t \in \mathbb{Z}}$  défini par

$$Y_t = \sum_{j \in \mathbb{Z}} a_j X_{t-j}. \quad (3.5.1)$$

Démontrer que  $(Y_t)_{t \in \mathbb{Z}}$  est stationnaire au second ordre et donner l'expression de sa fonction d'auto-covariance en fonction de celle de  $(X_t)_{t \in \mathbb{Z}}$  et de  $(a_j)_{j \in \mathbb{Z}}$ .

**Proposition 3.38.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire au second ordre, centré, de fonction d'auto-covariance  $\gamma_X(\cdot)$  et soit  $(a_j)_{j \in \mathbb{Z}}$  une famille de réels absolument sommable. Le processus  $(Y_t)_{t \in \mathbb{Z}}$  défini par

$$Y_t = \sum_{j \in \mathbb{Z}} a_j X_{t-j} \quad (3.5.2)$$

est stationnaire, de fonction d'auto-covariance

$$\gamma_Y(h) = \sum_{i,j \in \mathbb{Z}} a_i a_{j+h} \gamma_X(i-j). \quad (3.5.3)$$

**Définition 3.39.** Une fonction  $K: \mathbb{Z} \rightarrow \mathbb{R}$  est de type positif si pour tout entier  $n \geq 1$ , tous  $a_1, \dots, a_n \in \mathbb{R}$  et tous  $t_1, \dots, t_n \in \mathbb{Z}$ ,

$$\sum_{k,\ell=1}^n a_k a_\ell K(t_k - t_\ell) \geq 0. \quad (3.5.4)$$

**Exercice 26.** Démontrer que si  $K: \mathbb{Z} \rightarrow \mathbb{R}$  est la fonction d'auto-covariance d'un processus stationnaire au second ordre, alors  $K$  est paire et de type positif.

**Proposition 3.40.** Une fonction  $K: \mathbb{Z} \rightarrow \mathbb{R}$  est la fonction d'auto-covariance d'un processus stationnaire au second ordre si et seulement si elle est paire et de type positif.

**Définition 3.41.** La fonction d'auto-corrélation d'un processus stationnaire au second ordre  $(X_t)_{t \in \mathbb{Z}}$  de variance  $\sigma^2$  est la fonction  $\rho: \mathbb{Z} \rightarrow \mathbb{R}$  définie par

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{\text{Cov}(X_t, X_{t+h})}{\sigma^2}. \quad (3.5.5)$$

La quantité  $\rho(h)$  est le coefficient de corrélation linéaire entre  $X_t$  et  $X_{t+h}$ . Si  $\rho(h)$  est proche de 1,  $X_t$  et  $X_{t+h}$  ont une forte liaison linéaire.

**Exercice 27.** Démontrer que pour tout  $h \in \mathbb{Z}$ ,  $\rho(-h) = \rho(h)$  et  $|\rho(h)| \leq 1$ .

On cherche à mesurer la dépendance directe entre  $X_t$  et  $X_{t+h}$ .

**Définition 3.42.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire au second ordre. Pour  $(s, t) \in \mathbb{Z}^2$  tel que  $s \leq t$ , on note  $H_s^t = \{\sum_{i=s}^t c_i X_i, (c_i)_{s \leq i \leq t} \in \mathbb{R}\}$ . Pour  $h \geq 2$ , on note

$$U_h := X_t - \mathbb{E}\mathbb{L}(X_t \mid (X_i)_{t-h+1 \leq i \leq t-1}), \quad V_h := X_{t-h} - \mathbb{E}\mathbb{L}(X_{t-h} \mid (X_i)_{t-h+1 \leq i \leq t-1}). \quad (3.5.6)$$

La fonction d'autocorrélation partielle du processus stationnaire  $(X_t)_{t \in \mathbb{Z}}$  est définie pour  $h \geq 2$  par

$$\tau(h) = \frac{\text{Cov}(U_h, V_h)}{(\text{Var}(U_h) \text{Var}(V_h))^{1/2}}. \quad (3.5.7)$$

La variable aléatoire  $U_h$  est la variable aléatoire  $X_t$  à laquelle on a retiré toute l'information intermédiaire strictement comprise entre les instants  $t-h$  et  $t$ . De même, la variable aléatoire  $V_h$  est la variable aléatoire  $X_{t-h}$  à laquelle on a retiré toute l'information intermédiaire strictement comprise entre les instants  $t-h$  et  $t$ .

**Exercice 28.** Expliquer pourquoi  $\tau(h)$  défini par (3.5.7) ne dépend pas de  $t$ .

On peut définir  $\tau(h)$  pour  $h \leq -2$  par  $\tau(h) = \tau(-h)$ . On pose également

$$\tau(0) = 0 \quad \tau(1) = \tau(-1) = \frac{\gamma(1)}{\gamma(0)}. \quad (3.5.8)$$

**Proposition 3.43.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire au second ordre centré. Alors pour tout  $t \in \mathbb{Z}$  et tout  $h \geq 1$ ,

$$\mathbb{E}\mathbb{L}(X_t \mid H_{t-h}^{t-1}) = \mathbb{E}\mathbb{L}(X_t \mid (X_i)_{t-h \leq i \leq t-1}) = \sum_{\ell=1}^h \tau(\ell) X_{t-\ell}. \quad (3.5.9)$$

Pour  $h \in \mathbb{N}^*$  et  $t \in \mathbb{Z}$ , on note  $\Gamma^{(h)}$  la matrice de covariance du vecteur  $\mathbb{X} = (X_{t-1}, \dots, X_{t-h})^\top$ . Autrement dit,  $\Gamma_{i,j}^{(h)} = \gamma(|i-j|) / \gamma(0)$ . On note également par  $\tilde{\rho}_h = (\rho(1), \dots, \rho(h))^\top$  le vecteur des  $h$  premières valeurs de la fonction d'auto-corrélation.

**Proposition 3.44.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire au second ordre centré. Soit  $\Lambda = (\lambda_1, \dots, \lambda_h)$  tel que  $\mathbb{E}\mathbb{L}(X_t \mid H_{t-h}^{t-1}) = \Lambda^\top \mathbb{X}$ . Alors  $\tilde{\rho}_h = \Gamma^{(h)} \Lambda$ .

On cherche à présent à exprimer la meilleure prédiction linéaire d'un processus stationnaire  $(X_t)_{t \in \mathbb{Z}}$ . On dispose d'observations  $X_1, \dots, X_T$ . La meilleure prédiction linéaire à l'horizon  $h \in \mathbb{N}^*$  est donnée par la projection orthogonale de  $X_{T+h}$  sur l'espace vectoriel engendré par  $X_1, \dots, X_T$ , noté  $H_T$ .

**Proposition 3.45.** *Soit  $h \in \mathbb{N}^*$  et soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire au second ordre et centré de fonction d'autocorrélation  $\rho(\cdot)$ . Soit  $\Gamma_T$  la matrice de corrélation du vecteur aléatoire  $\mathbb{X} = (X_1, \dots, X_T)^\top$ . Soit  $\tilde{\rho}_h := (\rho(T+h-1), \dots, \rho(h))^\top$ . Si  $\Gamma_T$  est inversible, alors les coefficients  $\tilde{\Psi}_T = (\Psi_{1,h}, \dots, \Psi_{T,h})^\top$  de la régression linéaire*

$$\mathbb{E}\mathbb{L}(X_{T+h} \mid H_T) = \sum_{\ell=1}^T \Psi_{\ell,h} X_\ell \quad (3.5.10)$$

sont donnés par  $\tilde{\Psi}_T = \Gamma_T^{-1} \tilde{\rho}_T$ . De plus,

$$\mathbb{E}[(X_{T+h} - \mathbb{E}\mathbb{L}(X_{T+h} \mid H_T))^2] = \gamma(0) \left(1 - \tilde{\Psi}_T^\top \tilde{\rho}_T\right). \quad (3.5.11)$$

La prédiction linéaire demande l'inversion d'une matrice de taille  $T \times T$ , ce qui est dans la pratique coûteux en termes de calcul. L'algorithme des innovations permet de palier ce problème.

On se base sur la remarque suivante : l'espace vectoriel des combinaisons linéaires des variables aléatoires  $X_1, \dots, X_t$  est le même que l'espace vectoriel constitué des combinaisons linéaires des variables aléatoires  $X_i - \mathbb{E}\mathbb{L}(X_i \mid H_{i-1})$ , avec  $H_{i-1} = \text{vect}(X_1, \dots, X_{i-1})$  et  $\mathbb{E}\mathbb{L}(X_1 \mid H_0) = \mathbb{E}[X_1]$ . Par conséquent, on peut écrire

$$\mathbb{E}[X_{t+1} \mid H_t] = \sum_{i=1}^t \theta_i^{(t)} (X_{t+1-i} - \mathbb{E}\mathbb{L}(X_{t+1-i} \mid H_{-it})). \quad (3.5.12)$$

L'algorithme des innovations permet de calculer récursivement les coefficients  $\theta_i^{(t)}$  ainsi que le terme d'erreur

$$v_t = \mathbb{E}[(X_{t+1} - \mathbb{E}[X_{t+1} \mid H_t])^2]. \quad (3.5.13)$$

**Proposition 3.46.** *Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus centré défini sur un espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$  et tel que  $\mathbb{E}[X_t^2] < \infty$  pour tout  $t \in \mathbb{Z}$ . On suppose que pour tout  $t \geq 1$ , la matrice de covariance du vecteur  $(X_1, \dots, X_t)$  est inversible. Alors les coefficients  $(\theta_i^{(t)})$  dans (3.5.12) et les erreurs moyennes  $v_t$  dans (3.5.13) par*

$$v_0 = \mathbb{E}[X_1^2] \quad (3.5.14)$$

$$\theta_t^{(t)} = v_0^{-1} \mathbb{E}[X_{t+1} X_1] \quad (3.5.15)$$

$$\theta_{t-i}^{(t)} = v_i^{-1} \left( \mathbb{E}[X_{t+1} X_{i+1}] - \sum_{j=0}^{i-1} \theta_{j-i}^{(i)} \theta_{t-i}^{(t)} v_j \right) \quad (3.5.16)$$

$$v_t = \mathbb{E}[X_{t+1}^2] - \sum_{j=0}^{t-1} (\theta_{t-j}^{(t)})^2 v_j. \quad (3.5.17)$$

En notant  $S_{i,j} = \mathbb{E}[X_i X_j]$ , on voit que l'on peut calculer récursivement les  $v_t$  et  $\theta_{t-i}^{(t)}$ . En effet,  $v_0 = S_{1,1}$ ,  $\theta_1^{(1)} = S_{1,2}/S_{1,1}$  et  $v_1 = S_{2,2} - \left(\theta_1^{(1)}\right)^2 S_{1,1}$  puis

$$\theta_2^{(2)} = S_{1,3}/S_{1,1}, \quad \theta_1^{(2)} = \frac{1}{v_1} \left( S_{2,3} - \theta_1^{(1)} \theta_2^{(2)} S_{1,2} \right) \quad (3.5.18)$$

et

$$v_2 = S_{3,3} - \left(\theta_2^{(2)}\right)^2 v_0 - \left(\theta_1^{(2)}\right)^2 v_1. \quad (3.5.19)$$

On peut également mettre en place l'algorithme des innovations pour faire de la prédiction à l'horizon  $h \geq 2$ , c'est-à-dire de déterminer  $\mathbb{E}\mathbb{L}(X_{T+h} | H_T)$  au lieu de  $\mathbb{E}\mathbb{L}(X_{T+1} | H_T)$ . Par la Proposition 3.9,

$$\mathbb{E}\mathbb{L}(X_{T+h} | H_T) = \mathbb{E}\mathbb{L}(\mathbb{E}\mathbb{L}(X_{T+h} | H_{T+h-1}) | H_T). \quad (3.5.20)$$

De plus, en utilisant (3.5.12) avec  $t$  remplacé par  $T+h-1$ , on obtient

$$\mathbb{E}\mathbb{L}(X_{T+h} | H_T) = \sum_{u=1}^{T+h-1} \theta_u^{(T+h-1)} \mathbb{E}\mathbb{L}(X_{T+h-u} - \mathbb{E}\mathbb{L}(X_{T+h-u} | H_{T+h-u-1}) | H_T). \quad (3.5.21)$$

Si  $u \leq h-1$ ,  $X_{T+h-u} - \mathbb{E}\mathbb{L}(X_{T+h-u} | H_{T+h-u-1})$  est orthogonale à  $H_{T+h-u-1}$  donc à  $H_T$  et si  $u \geq h$ , alors  $X_{T+h-u} - \mathbb{E}\mathbb{L}(X_{T+h-u} | H_{T+h-u-1})$  est un élément de  $H_T$ , d'où

$$\mathbb{E}\mathbb{L}(X_{T+h} | H_T) = \sum_{u=h}^{T+h-1} \theta_u^{(T+h-1)} (X_{T+h-u} - \mathbb{E}\mathbb{L}(X_{T+h-u} | H_{T+h-u-1})). \quad (3.5.22)$$

**Exercice 29.** Soient  $\{C_j, j = 1, \dots, J\}$  et  $\{D_j, j = 1, \dots, J\}$  des variables aléatoires centrées telles que pour tout  $(j, k) \in \{1, \dots, J\}^2$ ,  $\mathbb{E}[C_j C_k] = \mathbb{E}[D_j D_k] = \sigma_j^2 \delta_{j,k}$  et  $\mathbb{E}[C_j D_k] = 0$ . On considère le processus harmonique défini pour tout  $t \in \mathbb{Z}$  par :

$$X_t = \sum_{j=1}^J (C_j \cos(t\lambda_j) - D_j \sin(t\lambda_j)),$$

avec pour tout  $j \in \{1, \dots, J\}$ ,  $\lambda_j \in [-\pi, \pi]$ . Montrer que le processus  $(X_t)_{t \in \mathbb{Z}}$  est stationnaire et calculer sa fonction d'autocovariance. *Astuce : utiliser la relation  $\cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b)$ .*

### 3.6 Estimation de la moyenne et de la fonction d'auto-corrélation

Étant données des réalisations  $x_1, \dots, x_T$  de variables aléatoires  $X_1, \dots, X_T$ , on considère l'estimateur empirique de la moyenne

$$\hat{\mu}_T = \frac{1}{T} \sum_{t=1}^T X_t. \quad (3.6.1)$$

**Exercice 30.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire au second ordre. Démontrer que  $\hat{\mu}_T = \frac{1}{T} \sum_{t=1}^T X_t$  est un estimateur sans biais de  $\mu$  et calculer sa variance en fonction de  $(\gamma_X(h))_{h \in \mathbb{Z}}$ .

**Exercice 31.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire au second ordre tel que  $(\gamma_X(h))_{h \in \mathbb{Z}}$  soit absolument sommable. Déterminer  $\lim_{T \rightarrow \infty} T \mathbb{E} [(\hat{\mu}_T - \mu)^2]$  en fonction de  $(\gamma_X(h))_{h \in \mathbb{Z}}$ .

Ceci implique qu'un processus stationnaire au second ordre pour lequel  $(\gamma_X(h))_{h \in \mathbb{Z}}$  est absolument sommable est ergodique en moyenne quadratique.

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire au second ordre. Alors pour chaque  $i \in \mathbb{Z}$ , la variable aléatoire  $(X_i - \mu)(X_{i+h} - \mu)$  a pour espérance  $\gamma_X(h)$ . En se basant sur cette remarque, on considère l'estimateur de  $\gamma(h)$  pour  $h$  tel que  $|h| < T$ , défini par

$$\hat{\gamma}_T(h) = \frac{1}{T} \sum_{t=1}^{T-|h|} (X_t - \hat{\mu}_T)(X_{t+|h|} - \hat{\mu}_T) \quad (3.6.2)$$

(on notera que la normalisation est  $T$  et non  $T - |h|$ ). Pour la fonction de corrélation, on prend l'estimateur

$$\hat{\rho}_T(h) = \frac{\hat{\gamma}_T(h)}{\hat{\gamma}_T(0)}. \quad (3.6.3)$$

On définit pour  $p \in \llbracket 1, T \rrbracket$  l'estimateur  $\hat{\Gamma}_p$  de la matrice de corrélation du vecteur  $\mathbb{X} = (X_1, \dots, X_p)^\top$ , défini par

$$[\hat{\Gamma}_p]_{i,j} = \hat{\rho}_T(|i-j|) = \frac{\hat{\gamma}_T(|i-j|)}{\hat{\gamma}_T(0)}. \quad (3.6.4)$$

**Proposition 3.47.** Si  $\hat{\gamma}_T(0) \neq 0$ , alors la matrice  $\hat{\Gamma}_p$  est définie positive.

Pour  $h \in \llbracket 1, T-1 \rrbracket$ , on estime  $\tau(h)$  par  $\hat{\tau}_T(h)$ , la  $h$ -ième composante du vecteur

$$\hat{\Gamma}_h^{-1}(\hat{\rho}_T(1), \dots, \hat{\rho}_T(h))^\top.$$

### 3.7 Fonction génératrice de l'auto-covariance

**Définition 3.48.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire au second ordre, de fonction d'auto-covariance  $\gamma_X(\cdot)$ . On suppose qu'il existe  $\delta$  tel que pour tout  $z \in C_\delta = \{z \in \mathbb{C}, 1 - \delta < |z| < 1 + \delta\}$ ,

$$\sum_{h \in \mathbb{Z}} |\gamma_X(h)| |z|^h < \infty. \quad (3.7.1)$$

La fonction génératrice de l'auto-covariance est donnée par

$$G_X(z) = \sum_{h \in \mathbb{Z}} \gamma_X(h) z^h. \quad (3.7.2)$$

Si  $(X_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible, alors  $G_X(z) = \sigma^2$  pour tout  $z \in \mathbb{C}$ .

**Proposition 3.49.** Soit  $(\eta_t)_{t \in \mathbb{Z}}$  un bruit blanc faible de variance  $\sigma^2$  et soit  $(a_i)_{i \in \mathbb{Z}}$  une famille de réels telle qu'il existe  $\delta > 0$  pour lequel la série  $\sum_{i \in \mathbb{Z}} |a_i| |z|^i$  converge pour tout  $z \in C_\delta$ . On pose  $\Phi(z) = \sum_{i \in \mathbb{Z}} a_i z^i, z \in C_\delta$ . Soit  $(X_t)_{t \in \mathbb{Z}}$  le processus défini par  $X_t = \Phi(B) \eta_t$ .

Alors la fonction génératrice de l'auto-covariance de  $(X_t)_{t \in \mathbb{Z}}$  est donnée par  $G_X(z) = \sigma \Phi(z) \Phi(1/z)$ , où  $z$  est tel que  $z, 1/z \in C_\delta$ .

## 4 Les processus moyennes mobiles

### 4.1 Définition et principales caractéristiques

**Définition 4.1.** Soit  $q \in \mathbb{N}^*$ . On appelle moyenne mobile d'ordre  $q$  (noté MA ( $q$ )) tout processus admettant la représentation

$$X_t = Q(B) \eta_t = \eta_t - \sum_{k=1}^q \theta_k \eta_{t-k}, \quad (4.1.1)$$

où  $Q(\cdot)$  est un polynôme à coefficients réels d'ordre  $q$  défini pour  $z \in \mathbb{C}$  par  $Q(z) = 1 - \sum_{k=1}^q \theta_k z^k$ ,  $\theta_q \neq 0$  et  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible de variance  $\sigma^2$ .

*Remarque 4.2.* Si  $(X_t)_{t \in \mathbb{Z}}$  est un processus MA ( $q$ ), alors  $\mathbb{E}[X_t] = 0$ .

*Remarque 4.3.* Si  $(X_t)_{t \in \mathbb{Z}}$  est un processus MA ( $q$ ), alors  $(X_t)_{t \in \mathbb{Z}}$  est stationnaire au second ordre. Notons qu'il n'y a pas besoin de faire d'hypothèse supplémentaire sur le polynôme  $Q(\cdot)$ .

**Proposition 4.4.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus MA ( $q$ ) de polynôme  $Q(\cdot)$ . On suppose que toutes les racines de  $Q(\cdot)$  sont de module strictement supérieur à 1. Alors le bruit blanc d'innovation  $\varepsilon_t = X_t - \mathbb{E}[X_t | \mathcal{H}_{t-1}^X]$  est égal à  $\eta_t$ .

**Proposition 4.5.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus MA ( $q$ ) de polynôme  $Q(\cdot)$ . On suppose que  $Q(\cdot)$  n'admet aucune racine de module 1. Alors il existe un polynôme  $\tilde{Q}(\cdot)$  dont toutes les racines sont de module strictement supérieur à 1 et tel que  $X_t = \tilde{Q}(B) \varepsilon_t$ , où  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible.

**Définition 4.6.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus MA ( $q$ ) défini par  $X_t = Q(B) \eta_t$ , où  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc. Si toutes les racines de  $Q$  sont de module différent de 1, le polynôme  $\tilde{Q}(\cdot)$  comme dans la Proposition 4.5 est appelé polynôme canonique.

**Proposition 4.7.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus MA ( $q$ ) défini par  $X_t = Q(B) \eta_t$ , où  $Q(z) = 1 - \sum_{\ell=1}^q \theta_\ell z^\ell$  et  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible. Alors la covariance de  $(X_t)_{t \in \mathbb{Z}}$  est donnée par

$$\gamma(h) = \mathbb{E}[\eta_0^2] \sum_{i=0}^{q-h} \theta_i \theta_{i+h} \mathbf{1}_{\{0 \leq h \leq q\}}, \quad (4.1.2)$$

où  $\theta_0 = -1$ .

Remarquons que si  $h > q$ , alors  $\gamma(h) = 0$ .

**Proposition 4.8.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire au second ordre. On suppose qu'il existe  $q \geq 1$  tel que si  $h > q$ , alors  $\gamma_X(h) = 0$ . Alors  $(X_t)_{t \in \mathbb{Z}}$  est un processus MA ( $q$ ).

**Corollaire 4.9.** Soient  $(X_t)_{t \in \mathbb{Z}}$  et  $(Y_t)_{t \in \mathbb{Z}}$  des processus stationnaires au second ordre et centrés. On suppose que  $(X_t)_{t \in \mathbb{Z}}$  est un processus MA ( $q_1$ ), que  $(Y_t)_{t \in \mathbb{Z}}$  est un processus MA ( $q_2$ ) et que pour tous  $s, t \in \mathbb{Z}$ ,  $\mathbb{E}[X_s Y_t] = 0$ . Alors le processus  $(Z_t)_{t \in \mathbb{Z}}$  défini par  $Z_t = X_t + Y_t$  est un processus MA ( $\max\{q_1, q_2\}$ ).

**Exercice 32.** On considère le processus  $(X_t)_{t \in \mathbb{Z}}$  défini par :

$$X_t = \eta_t - 2\eta_{t-1}, \quad t \in \mathbb{Z},$$

où  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible de variance  $\sigma_\eta^2$ . Déterminer le processus d'innovation du processus  $(X_t)_{t \in \mathbb{Z}}$  et calculer sa variance en fonction de  $\sigma_\eta^2$ .

**Exercice 33.** 1. Rappeler l'énoncé du théorème limite central pour une suite de variables aléatoires i.i.d.

2. Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus moyenne mobile d'ordre 1, de la forme  $X_t = \eta_t - \theta\eta_{t-1}$ , où  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc fort. Établir un théorème limite central pour  $S_T := \sum_{t=1}^T X_t$ .

## 4.2 Estimation d'un modèle moyenne mobile

On cherche à déterminer si un processus est un processus MA( $q$ ). Si on avait la fonction d'auto-covariance à disposition, il suffirait de regarder si elle s'annule à partir d'un certain rang. Dans la pratique, on ne connaît pas la fonction d'auto-covariance ni celle d'auto-corrélation. On l'estime par

$$\hat{\rho}_T(h) = \frac{\sum_{t=1}^{T-|h|} (X_t - \hat{\mu}_T)(X_{t+|h|} - \hat{\mu}_T)}{\sum_{t=1}^T (X_t - \hat{\mu}_T)^2}, \quad (4.2.1)$$

où

$$\hat{\mu}_T = \frac{1}{T} \sum_{t=1}^T X_t. \quad (4.2.2)$$

**Proposition 4.10.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus MA( $q$ ), c'est-à-dire que  $X_t = \eta_t - \sum_{\ell=1}^q \theta_\ell \eta_\ell$ , où  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc fort. Pour tout  $h > q$ , la suite  $\left(\sqrt{T} \hat{\rho}_T(h)\right)_{T \geq |h|}$  converge en loi vers une loi normale centrée de variance

$$\sigma^2(q) = \sum_{k=1}^{\infty} (\rho(k-h))^2 = \sum_{k=h-q}^{h+q} (\rho(k-h))^2 = 1 + 2\rho(1)^2 + \dots + 2\rho(q)^2. \quad (4.2.3)$$

Ceci permet de construire les intervalles de confiance suivants : pour tout  $h > q$ ,  $\mathbb{P}(\hat{\rho}_T(h) \in I_{q,\alpha}) = 1 - \alpha$ , avec

$$I_{q,\alpha} = \left[ -\Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \left( \frac{1 + 2\rho(1)^2 + \dots + 2\rho(q)^2}{T} \right)^{1/2}, \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \left( \frac{1 + 2\rho(1)^2 + \dots + 2\rho(q)^2}{T} \right)^{1/2} \right], \quad (4.2.4)$$

où  $\Phi$  désigne la fonction de répartition d'une loi normale centrée réduite et  $\Phi^{-1}$  sa réciproque.

En pratique, on cherche à savoir si des observations  $X_1, \dots, X_T$  sont issues d'un processus MA( $q$ ). On se fixe une erreur (typiquement,  $\alpha = 0.05$ ). On calcule l'intervalle de confiance estimé

$$\hat{I}_{1,\alpha} = \left[ -\Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \left( \frac{1 + 2\hat{\rho}_T(1)^2}{T} \right)^{1/2}, \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \left( \frac{1 + 2\hat{\rho}_T(1)^2}{T} \right)^{1/2} \right]. \quad (4.2.5)$$

Si pour tout  $h > 1$ , on trouve que  $\widehat{\rho}_T(h) \in \widehat{I}_{1,\alpha}$ , on accepte que les observations provienne d'un processus MA (1). Sinon, on recommence avec  $q = 2$ . En posant

$$\widehat{I}_{q,\alpha} = \left[ -\Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \left( \frac{1 + 2\widehat{\rho}_T^2(1) + \cdots + 2\widehat{\rho}_T^2(q)}{T} \right)^{1/2}, \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \left( \frac{1 + 2\widehat{\rho}_T^2(1) + \cdots + 2\widehat{\rho}_T^2(q)}{T} \right)^{1/2} \right], \quad (4.2.6)$$

l'ordre du processus est donné par

$$\widehat{q} = \inf \left\{ q \geq 1, \widehat{\rho}_T(h) \in \widehat{I}_{q,\alpha} \text{ pour tout } q > h \right\}. \quad (4.2.7)$$

Si on trouve des estimations en dehors des intervalles de confiance (en pratique pour  $h > 20$ ), on rejette l'hypothèse d'un modèle MA.

Pour estimer les coefficients, on fait appel à l'algorithme des innovations.

## 5 Les processus autorégressifs

### 5.1 Définition et principales caractéristiques

**Définition 5.1.** Soit  $p \in \mathbb{N}^*$ . On dit que  $(X_t)_{t \in \mathbb{Z}}$  est un processus auto-régressif d'ordre  $p$  (noté AR( $p$ )) si  $(X_t)_{t \in \mathbb{Z}}$  est stationnaire au second ordre et est solution d'une équation de récurrence de la forme

$$P(B) X_t = X_t - \sum_{k=1}^p \Phi_k X_{t-k} = \eta_t, \quad (5.1.1)$$

où  $P(\cdot)$  est un polynôme à coefficients réels et de degré  $p$  et  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible de variance  $\sigma^2$ .

On notera qu'un processus vérifiant l'équation de récurrence  $P(B) X_t = \eta_t$  n'est pas nécessairement stationnaire.

Par la Proposition 3.37 :

- Si  $P(\cdot)$  n'a aucune racine sur le cercle unité, alors  $(X_t)_{t \in \mathbb{Z}}$  est inversible et s'écrit sous la forme

$$X_t = \sum_{i \in \mathbb{Z}} a_i \eta_{t-i}, \quad (5.1.2)$$

où  $(a_i)_{i \in \mathbb{Z}}$  est absolument sommable.

- Si  $P(\cdot)$  n'a aucune racine sur le disque unité fermé, alors  $(X_t)_{t \in \mathbb{Z}}$  est causal et s'écrit sous la forme

$$X_t = \sum_{i=0}^{\infty} a_i \eta_{t-i}, \quad (5.1.3)$$

où  $(a_i)_{i \geq 0}$  est absolument sommable.

Remarquons que si  $(X_t)_{t \in \mathbb{Z}}$  est inversible, alors ce processus est nécessairement stationnaire au second ordre et centré.

**Proposition 5.2.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus AR( $p$ ) défini par l'équation  $P(B)X_t = \eta_t$ . Si  $(X_t)_{t \in \mathbb{Z}}$  est causal, alors son bruit blanc d'innovation est le processus  $(\eta_t)_{t \in \mathbb{Z}}$ .

Comme pour les processus MA, on peut écrire un processus AR en fonction de son bruit blanc d'innovation.

**Proposition 5.3.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus AR( $p$ ) défini par  $P(B)X_t = \eta_t$ . On suppose que  $P(\cdot)$  n'a aucune racine sur le cercle unité. Alors il existe un polynôme  $\tilde{P}(\cdot)$  n'ayant aucune racine dans le disque unité fermé et tel que  $\tilde{P}(B)X_t = \varepsilon_t$ , où  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc.

**Exercice 34.** Soit  $P(\cdot)$  un polynôme d'ordre  $p \in \mathbb{N}^*$  tel que  $P(z) \neq 0$  pour tout  $z \in \mathbb{C}$  tel que  $|z| = 1$ . Montrer que si le processus  $(X_t)_{t \in \mathbb{Z}}$  est solution de l'équation  $P(B)X_t = \eta_t$ ,  $t \in \mathbb{Z}$ , il existe un polynôme  $\tilde{P}(\cdot)$  avec  $\tilde{P}(z) \neq 0$  pour tout  $z$  tel que  $|z| \leq 1$  et tel que  $(X_t)_{t \in \mathbb{Z}}$  soit solution de l'équation  $\tilde{P}(B)X_t = \varepsilon_t$ ,  $t \in \mathbb{Z}$  où  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc.

L'écriture  $\tilde{P}(B)X_t = \varepsilon_t$ , où  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc et les racines de  $\tilde{P}$  sont toutes de module  $> 1$  est appelée forme canonique de  $(X_t)_{t \in \mathbb{Z}}$ .

**Proposition 5.4.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus AR( $p$ ) défini par  $P(B)X_t = \eta_t$ , où  $P(\cdot)$  a toutes ses racines en dehors du disque unité.

1. La fonction d'auto-covariance est donnée pour tout  $h \in \mathbb{N}^*$  par

$$\gamma_X(h) = \mathbb{E}[\eta_0^2] \sum_{i=0}^{\infty} a_i a_{i+h}. \quad (5.1.4)$$

Pour  $h \in \mathbb{N}$ , on a

$$\gamma_X(h) = \sum_{i=1}^p \Phi_i \gamma(i-h). \quad (5.1.5)$$

2. La fonction d'auto-corrélation partielle  $\tau(\cdot)$  est telle que

$$\tau(p) = \Phi_p \text{ et } \tau(h) = 0 \text{ pour } h > p. \quad (5.1.6)$$

**Proposition 5.5** (Équations de Yule-Walker). Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus AR( $p$ ) défini par  $P(B)X_t = \eta_t$ , où  $P(\cdot)$  n'a aucune racine dans le disque unité. On a

$$\mathbb{E}[\eta_0^2] = \gamma(0) - \sum_{i=1}^p \Phi_i \gamma(i), \quad (5.1.7)$$

et pour  $h \geq p$ ,  $(\rho(1), \dots, \rho(h))^\top = \Gamma_h (\Phi_1, \dots, \Phi_h)^\top$ , où  $\Phi_i = 0$  si  $i > p$  et  $\Gamma_h$  est la matrice de corrélation du vecteur  $(X_1, \dots, X_h)^\top$ .

**Exercice 35.** On considère le processus  $(X_t)_{t \in \mathbb{Z}}$  solution de l'équation de récurrence :

$$X_t = \eta_t + \frac{8}{15}X_{t-1} - \frac{1}{15}X_{t-2},$$

où  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible.

1. Montrer que  $(X_t)_{t \in \mathbb{Z}}$  est un processus auto régressif pouvant s'écrire sous la forme

$$X_t = \sum_{i=0}^{\infty} \Phi_i^* \eta_{t-i}, \quad t \in \mathbb{Z}.$$

Préciser les valeurs de  $\Phi_0^*$ ,  $\Phi_1^*$  et  $\Phi_2^*$ .

2. À l'aide des équations de Yule-Walker, calculer les valeurs  $\rho(1)$  et  $\rho(2)$  où  $\rho(\cdot)$  est la fonction d'autocorrélation.
3. Donner la valeur de  $\tau(2)$  où  $\tau(\cdot)$  est la fonction d'autocorrélation partielle.

**Exercice 36.** 1. Soit  $(\eta_t)_{t \in \mathbb{Z}}$  un bruit blanc faible. Montrer que le processus  $(X_t)_{t \in \mathbb{Z}}$  défini implicitement par l'équation de récurrence :

$$X_t - \frac{1}{2}X_{t-1} = \eta_t, \quad t \in \mathbb{Z},$$

est un processus stationnaire. Calculer la valeur de la variance de  $X_t$ . Quel est le nom du processus ainsi défini ?

2. Soit  $(\eta_t)_{t \in \mathbb{Z}}$  un bruit blanc faible. On considère le processus  $(X_t)_{t \in \mathbb{Z}}$  défini implicitement par  $X_0 = \eta_0$  et pour  $|t| \geq 1$ , par l'équation de récurrence :

$$X_t - \frac{1}{2}X_{t-1} = \eta_t.$$

Ce processus est-il stationnaire ?

**Exercice 37.** Soit le processus  $(X_t)_{t \in \mathbb{Z}}$  solution de l'équation

$$X_t - \frac{9}{2}X_{t-1} + 15X_{t-2} - \frac{13}{2}X_{t-3} = \eta_t, \quad t \in \mathbb{Z},$$

où  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible.

1. Montrer que  $(X_t)_{t \in \mathbb{Z}}$  est un processus autorégressif dont vous préciserez l'ordre. (*Aide : remarquer que 2 est racine du polynôme  $1 - 9z/2 + 15z^2 - 13z^3/2$* ).
2. Donner la représentation canonique du processus  $(X_t)_{t \in \mathbb{Z}}$ .

**Exercice 38.** Soit le processus  $(X_t)_{t \in \mathbb{Z}}$  défini implicitement par l'équation de récurrence,

$$X_t = \Phi X_{t-1} + \eta_t, \quad t \in \mathbb{Z},$$

où  $(\eta_t)_{t \in \mathbb{N}}$  est un bruit blanc faible et  $|\Phi| < 1$ . Montrer que ce processus est stationnaire et que pour tout  $h \in \mathbb{N}^*$ , la fonction de covariance est donnée par  $\gamma(h) = \gamma(0)\Phi^{|h|}$ .

## 5.2 Estimation d'un modèle autorégressif

On cherche à estimer les coefficients d'un processus auto-régressif. Tout d'abord, il faut estimer l'ordre. Le point clé est que dans l'équation

$$\mathbb{E}\mathbb{L}(X_t | X_{t-1}, \dots, X_{t-h}) = \sum_{j=1}^h \Phi_j X_{t-j}, \quad (5.2.1)$$

si  $(X_t)_{t \in \mathbb{Z}}$  est un processus AR( $p$ ), alors  $\Phi_h = 0$  pour  $j > p$ . Or on sait que les coefficients  $\Phi_j$  peuvent s'exprimer à l'aide de la fonction d'auto-corrélation partielle.

On pose

$$\hat{\rho}_T(h) = \frac{\sum_{t=1}^{T-|h|} (X_t - \hat{\mu}_T)(X_{t+|h|} - \hat{\mu}_T)}{\sum_{t=1}^T (X_t - \hat{\mu}_T)^2}, \quad (5.2.2)$$

où

$$\hat{\mu}_T = \frac{1}{T} \sum_{t=1}^T X_t. \quad (5.2.3)$$

Un estimateur de  $\Phi^{(h)} = (\Phi_1, \dots, \Phi_h)^\top$  est donné par  $\hat{\Phi}_T^{(h)} = \left( \hat{\Phi}_1, \dots, \hat{\Phi}_h \right)^\top = \widehat{\Gamma}_h^{-1} (\hat{\rho}_T(1), \dots, \hat{\rho}_T(h))^\top$ .

**Proposition 5.6.** *Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus AR( $p$ ) causal défini par  $P(B)X_t = \eta_t$ , où  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc fort. Alors pour tout  $h > p$ , la suite  $\left( \sqrt{T} \hat{\Phi}_T^{(h)} \right)_{T \geq 1}$  converge vers une loi normale centrée réduite.*

On pose alors

$$\hat{p} = \inf \left\{ r \geq 1, \forall h \geq r, \left| \hat{\Phi}_h \right| \leq Q_N(1 - \alpha/2) / \sqrt{T} \right\}, \quad (5.2.4)$$

où  $Q_N(u)$  désigne le quantile d'une loi normale centrée réduite. On prend en pratique  $\alpha = 0.05$ .

Pour estimer les paramètres, on utilise les équations de Yule-Walker.

## 5.3 Prédiction d'un processus AR( $p$ )

**Proposition 5.7.** *Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus AR( $p$ ) sous sa forme canonique. La meilleure prédiction de  $X_{T+h}$  à partir de  $X_1, \dots, X_T$  est donnée par*

$$\mathbb{E}\mathbb{L}(X_{T+h} | X_1, \dots, X_T) = \sum_{i=1}^p \Phi_i \mathbb{E}\mathbb{L}(X_{T+h-i} | X_1, \dots, X_T). \quad (5.3.1)$$

On peut estimer à l'horizon  $h$  de manière récursive. On estime  $\mathbb{E}\mathbb{L}(X_{T+1} | X_1, \dots, X_T)$  par  $\hat{X}_{T+1} = \sum_{i=1}^p \hat{\Phi}_i X_{T+1-i}$ . On se base sur cette estimation de  $X_{T+1}$  pour en fournir une de  $\mathbb{E}\mathbb{L}(X_{T+2} | X_1, \dots, X_T)$  :

$$\hat{X}_{T+2} = \hat{\Phi}_1 \hat{X}_{T+1} + \sum_{i=2}^p \hat{\Phi}_i X_{T+1-i}. \quad (5.3.2)$$

On peut également estimer les résidus. Pour un AR( $p$ ) causal :

$$\eta_t = X_t - \mathbb{E}L(X_t | \mathcal{H}_{t-1}^X) = X_t - \sum_{i=1}^p \Phi_i X_{t-i} \quad (5.3.3)$$

et les résidus sont estimés par

$$\hat{\eta}_t = X_t - \sum_{i=1}^p \hat{\Phi}_i X_{t-i}, \quad t \in \llbracket p+1, T \rrbracket. \quad (5.3.4)$$

En cas d'adéquation au modèle AR( $p$ ), les observations  $\{\hat{\eta}_t, t \in \llbracket p+1, T \rrbracket\}$  doivent être issues d'un bruit blanc faible.

## 6 Processus ARMA

### 6.1 Définition

**Définition 6.1.** Soient  $(p, q) \in \mathbb{N}$ . Un processus autorégressif moyenne mobile d'ordres  $(p, q)$  (noté ARMA( $p, q$ )) est un processus  $(X_t)_{t \in \mathbb{Z}}$  stationnaire au second ordre, centré, solution d'une équation de récurrence de la forme

$$P(B)X_t = Q(B)\eta_t, \quad t \in \mathbb{Z}, \quad (6.1.1)$$

où  $P: z \mapsto 1 - \sum_{i=1}^p \Phi_i z^i$ ,  $\Phi_p \neq 0$ ,  $Q: z \mapsto 1 - \sum_{k=1}^q \theta_k z^k$ ,  $\theta_q \neq 0$  et  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible de variance  $\sigma^2$ .

*Remarque 6.2.* Un processus solution de (6.1.1) n'est pas nécessairement stationnaire.

*Remarque 6.3.* La classe des processus ARMA contient celle des processus MA ainsi que celle des processus AR. Plus précisément :

1. Un processus ARMA( $p, 0$ ) est un AR( $p$ ).
2. Un processus ARMA( $0, q$ ) est un MA( $q$ ).
3. Un processus ARMA( $0, 0$ ) est un bruit blanc faible.

On cherche dans un premier temps à simplifier l'équation de récurrence (6.1.1). On suppose que les polynômes  $P(\cdot)$  et  $Q(\cdot)$  ont une racine commune  $\mu$  de module différent de 1. On peut alors exprimer  $P(\cdot)$  sous la forme  $P(z) = (1 - z/\mu)P_1(z)$  et  $Q(z) = (1 - z/\mu)Q_1(z)$ , où  $P_1(0) = Q_1(0) = 1$ . Ainsi,  $(X_t)_{t \in \mathbb{Z}}$  vérifie l'équation de récurrence

$$\left(I - \frac{1}{\mu}B\right)P_1(B)X_t = \left(I - \frac{1}{\mu}B\right)Q_1(B)\eta_t \quad (6.1.2)$$

Comme l'opérateur  $I - \mu^{-1}B: V \rightarrow V$  est inversible,  $(X_t)_{t \in \mathbb{Z}}$  est solution de l'équation de récurrence

$$P_1(B)X_t = Q_1(B)\eta_t. \quad (6.1.3)$$

On peut ainsi « éliminer » les racines communes de  $P(\cdot)$  et  $Q(\cdot)$ , ce qui nous mène à la définition suivante.

**Définition 6.4.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus ARMA  $(p, q)$  solution de l'équation de récurrence (6.1.1). On dit que le processus  $(X_t)_{t \in \mathbb{Z}}$  est donné dans sa représentation minimale si les racines communes de  $P(\cdot)$  et  $Q(\cdot)$  sont toutes de module différent de 1.

## 6.2 Représentation d'un processus ARMA $(p, q)$ en MA $(\infty)$

**Proposition 6.5.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus ARMA  $(p, q)$  défini par (6.1.1).

1. Si  $P(\cdot)$  n'a aucune racine sur le cercle unité de  $\mathbb{C}$ , alors le processus  $(X_t)_{t \in \mathbb{Z}}$  est inversible et s'écrit sous la forme

$$X_t = \sum_{i \in \mathbb{Z}} \Phi_i^* \eta_{t-i}, \quad t \in \mathbb{Z}, \quad (6.2.1)$$

où  $(\Phi_i^*)_{i \in \mathbb{Z}}$  est absolument sommable.

2. Si  $P(\cdot)$  n'a aucune racine sur le disque unité fermé de  $\mathbb{C}$ , alors le processus  $(X_t)_{t \in \mathbb{Z}}$  est causal et s'écrit sous la forme

$$X_t = \sum_{i=0}^{\infty} \Phi_i^* \eta_{t-i}, \quad t \in \mathbb{Z}, \quad (6.2.2)$$

où  $(\Phi_i^*)_{i=0}^{\infty}$  est absolument sommable. De plus,

$$\phi_0 = 1 \text{ et } \Phi_i^* = -\theta_i + \sum_{k=1}^i \Phi_k \Phi_{i-k}^*. \quad (6.2.3)$$

## 6.3 Représentation d'un ARMA $(p, q)$ en AR $(\infty)$

**Proposition 6.6.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus ARMA  $(p, q)$  défini par (6.1.1).

1. Si  $Q(\cdot)$  n'a aucune racine sur le cercle unité, alors  $(\eta_t)_{t \in \mathbb{Z}}$  est inversible et s'écrit sous la forme

$$\eta_t = \sum_{i \in \mathbb{Z}} \theta_i^* X_{t-i}, \quad t \in \mathbb{Z}, \quad (6.3.1)$$

où  $(\theta_i^*)_{i \in \mathbb{Z}}$  est absolument sommable.

2. Si  $Q(\cdot)$  n'a aucune racine sur le cercle unité, alors  $(\eta_t)_{t \in \mathbb{Z}}$  est causal et s'écrit sous la forme

$$\eta_t = \sum_{i=0}^{\infty} \theta_i^* X_{t-i}, \quad t \in \mathbb{Z}, \quad (6.3.2)$$

où  $(\theta_i^*)_{i=0}^{\infty}$  est absolument sommable. De plus,

$$\theta_0^* = 1 \text{ et } \theta_i^* = -\Phi_i + \sum_{k=1}^i \theta_k \theta_{i-k}^*. \quad (6.3.3)$$

La démonstration serait un copié-collé de celle de la Proposition 6.5 en échangeant les rôles de  $P(\cdot)$  et  $Q(\cdot)$  ainsi que ceux de  $(\eta_t)_{t \in \mathbb{Z}}$  et  $(X_t)_{t \in \mathbb{Z}}$ .

## 6.4 Innovation d'un processus ARMA

Comme ce fut le cas pour les processus MA et AR, le processus  $(\eta_t)_{t \in \mathbb{Z}}$  n'est pas nécessairement le bruit blanc d'innovation du processus  $(X_t)_{t \in \mathbb{Z}}$ .

**Proposition 6.7.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus ARMA  $(p, q)$  défini par (6.1.1). Si toutes les racines de  $P(\cdot)$  ainsi que toutes celles de  $Q(\cdot)$  se trouvent à l'extérieur du disque unité fermé de  $\mathbb{C}$ , alors  $(X_t)_{t \in \mathbb{Z}}$  est un processus ARMA  $(p, q)$  de bruit blanc d'innovation  $(\eta_t)_{t \in \mathbb{Z}}$ .

**Exercice 39.** Soit le processus ARMA(1, 1)  $(X_t)_{t \in \mathbb{Z}}$  solution de l'équation

$$X_t - aX_{t-1} = \eta_t - b\eta_{t-1}, \quad t \in \mathbb{Z},$$

où  $|a| < 1$  et  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible de variance  $\sigma^2$ .

1. Si  $a = b$ , identifier le processus  $(X_t)_{t \in \mathbb{Z}}$  et calculer la valeur de  $\gamma(h)$  pour tout  $h \in \mathbb{N}$  où  $\gamma(\cdot)$  est la fonction d'autocovariance du processus  $(X_t)_{t \in \mathbb{Z}}$ .
2. Dans le cas  $a \neq b$ , calculer de deux façons différentes la variance de  $X_t$ .

**Exercice 40.** Soit le processus  $(X_t)_{t \in \mathbb{Z}}$  défini par l'équation  $X_t - \Phi X_{t-1} = \eta_t - \theta \eta_{t-1}$  où  $|\Phi| < 1$ ,  $|\theta| < 1$  et  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc de variance  $\sigma^2$ . Calculer en fonction de  $\Phi$ ,  $\theta$  et  $\sigma^2$  la fonction d'autocovariance du processus  $(X_t)_{t \in \mathbb{Z}}$ .

**Exercice 41.** Soit le processus  $(X_t)_{t \in \mathbb{Z}}$  solution de l'équation de récurrence :

$$X_t - 5X_{t-1} + 6X_{t-2} = \eta_t - 2\eta_{t-1}, \quad t \in \mathbb{Z},$$

où  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible.

1. Montrer que  $(X_t)_{t \in \mathbb{Z}}$  est un processus stationnaire.
2. Écrire la représentation canonique et minimale du processus  $(X_t)_{t \in \mathbb{Z}}$ . Identifier le processus  $(X_t)_{t \in \mathbb{Z}}$ .
3. On note  $(\varepsilon_t)$  le bruit blanc d'innovation du processus  $(X_t)_{t \in \mathbb{Z}}$ . Donner l'expression des coefficients  $\Psi_i$  de l'écriture

$$X_t = \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i}.$$

**Exercice 42.** On considère le polynôme  $P(\cdot)$  défini pour tout  $z \in \mathbb{C}$  par

$$P(z) = 2z^2 - \frac{9}{2}z + 1.$$

Soit un processus  $(X_t)_{t \in \mathbb{Z}}$  solution de l'équation de récurrence

$$P(B)X_t = \eta_t, \quad t \in \mathbb{Z}, \tag{6.4.1}$$

où  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc de variance  $\sigma^2$ .

1. Expliquer pourquoi il existe un unique processus  $(X_t)_{t \in \mathbb{Z}}$  solution de l'équation de récurrence (6.4.1).  
De quel type est le processus  $(X_t)_{t \in \mathbb{Z}}$  ?

2. On note  $(\varepsilon_t)_{t \in \mathbb{Z}}$  l'innovation du processus  $(X_t)_{t \in \mathbb{Z}}$ . Calculer les coefficients  $\{\varphi_i, i \in \mathbb{Z}\}$  tels que pour tout  $t \in \mathbb{Z}$ ,

$$\varepsilon_t = \sum_{i \in \mathbb{Z}} \varphi_i \eta_{t-i}.$$

Vérifier que la famille  $\{\varphi_i, i \in \mathbb{Z}\}$  est absolument sommable et que pour tout  $h > 1$  et  $t \in \mathbb{Z}$ ,  $\mathbb{E}[\varepsilon_t \varepsilon_{t+h}] = 0$ .

3) Montrer que pour tout  $t \in \mathbb{Z}$ ,

$$X_t = \sum_{i \in \mathbb{N}} \lambda_i \varepsilon_{t-i},$$

avec pour tout  $i \in \mathbb{N}$ ,

$$\lambda_i = \frac{1}{4^i} \sum_{j=0}^i 2^j.$$

## 7 Processus ARIMA (Auto Regressive Integrated Moving Average)

On peut modéliser des séries chronologiques ayant une tendance polynomiale  $T(t) = \sum_{i=0}^d a_i t^i$ . En différentiant cette série, on peut alors supprimer la tendance et ainsi modéliser la série par un ARMA.

**Définition 7.1.** Soit  $(p, d, q) \in \mathbb{N}^3$ . Un processus  $(X_t)_{t \in \mathbb{Z}}$  est un processus ARIMA d'ordres  $(p, d, q)$  si le processus différentié  $Y_t = (I - B)^d X_t$  est un processus ARMA  $(p, q)$ . On a donc

$$P(B)(I - B)^d X_t = Q(B) X_t, \quad (7.0.1)$$

où  $P(0) = Q(0) = 1$  et  $P$  et  $Q$  sont des polynômes de degrés respectifs  $p$  et  $q$ .

Le processus  $(X_t)_{t \in \mathbb{Z}}$  n'est pas défini de manière unique (par exemple, pour  $d = 1$ ,  $X_t + C$  est aussi une solution).

*Exemple 7.2.* Soit  $(X_t)_{t \in \mathbb{Z}}$  tel que  $X_t - X_{t-1}$ , où  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible. Alors  $(\eta_t)_{t \in \mathbb{Z}}$  n'est pas un processus stationnaire au second ordre (pourquoi, au fait ?) mais est un processus ARIMA  $(0, 1, 0)$ .

*Exemple 7.3.* Soient  $(a_0, \dots, a_{d-1}) \in \mathbb{R}^d$  et  $(\eta_t)_{t \in \mathbb{Z}}$  un bruit blanc faible et  $(X_t)_{t \in \mathbb{Z}}$  le processus défini par

$$X_t = \sum_{i=0}^{d-1} a_i t^i + \eta_t. \quad (7.0.2)$$

Ce n'est pas un processus stationnaire mais  $Y_t = (I - B)^d X_t$  est un MA  $(d)$  donc  $(X_t)_{t \in \mathbb{Z}}$  est un processus ARIMA  $(0, d, d)$ .

*Exemple 7.4.* Soit  $(Y_t)_{t \in \mathbb{Z}}$  un processus ARMA  $(p, q)$  et  $Z$  une variable aléatoire quelconque. On pose

$$X_t = \begin{cases} Z + \sum_{i=1}^t Y_i, & \text{si } t > 0, \\ Z, & \text{si } t = 0, \\ Z - \sum_{i=t+1}^0 Y_i, & \text{si } t < 0. \end{cases} \quad (7.0.3)$$

Alors  $(X_t)_{t \in \mathbb{Z}}$  est un processus ARIMA  $(p, 1, q)$ .

L'estimation d'un processus ARIMA se fait de la manière suivante. On dispose d'observations  $(x_t)_{t=1-d}^T$  que l'on suppose provenir de réalisations d'un processus ARIMA  $(p, d, q)$ , avec  $d$  inconnu.

**Exercice 43.** Soit  $(X_t)_{t \in \mathbb{Z}}$  le processus défini par

$$X_t = X_{t-1} + 6\eta_t - 5\eta_{t-1} + \eta_{t-2},$$

où  $(\eta_t)_{t \in \mathbb{Z}}$  est un bruit blanc faible de variance  $\sigma^2$ .

1. Déterminer  $\mathbb{E}[(X_{t+h} - X_t)^2]$ .
2. Déterminer si  $(X_t)_{t \in \mathbb{Z}}$  est stationnaire au second ordre.
3. Dédire une série temporelle stationnaire au second ordre  $(Y_t)_{t \in \mathbb{Z}}$  à partir de  $(X_t)_{t \in \mathbb{Z}}$ .
4. Donner la meilleure prédiction au sens des moindres carrés de  $Y_4$  à partir de  $Y_1, Y_2$  et  $Y_3$ .
5. En déduire une prédiction de  $X_4$  à l'aide de  $X_0, X_1, X_2$  et  $X_3$ .