

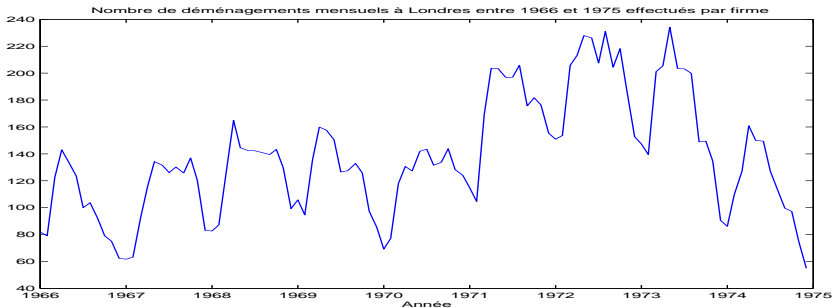
# Analyse des séries financières

Master 2 M.O.

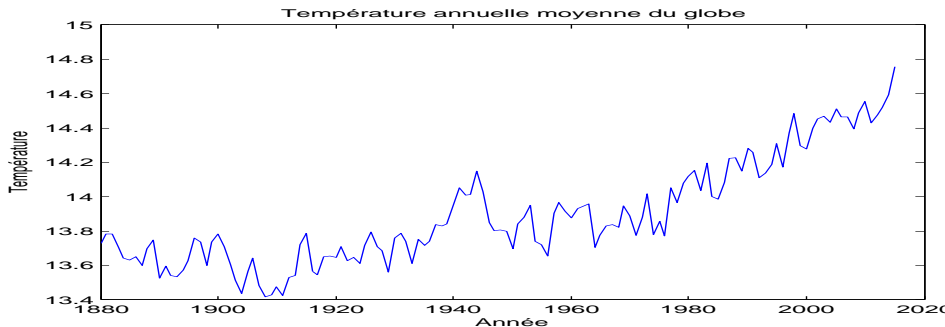


Année 2024-2025

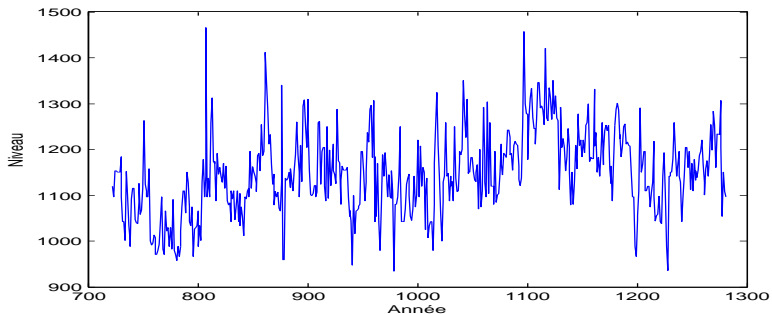
# Un exemple de série chronologique



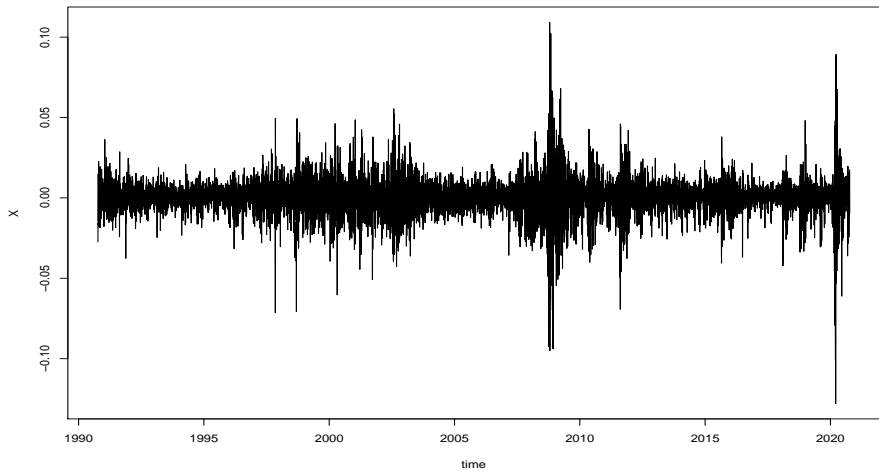
# Un autre exemple de série chronologique



# Un autre autre exemple de série chronologique



# Un exemple de série financière



# But du cours

But du cours :

- 1 Définir une série temporelle et donner des propriétés ;
- 2 Définir des modèles pour séries financières + propriétés ;
- 3 Analyse statistique (estimation, test, sélection de modèle) ;
- 4 Prédiction pour séries financières.

# Plan du cours

# Processus stochastique

## Définition

Soit  $(\Omega, \mathcal{A}, P)$  un espace de probabilité.

- $X = (X_t)_{t \in T}$  processus aléatoire (stochastique) réel sur  $T$  si  $\forall t \in T$ ,  $X_t$  v.a. sur  $(\Omega, \mathcal{A})$ .
- Pour  $\omega \in \Omega$ ,  $(X_t(\omega))_{t \in T}$  est appelé une trajectoire de  $X$ .
- $(X_t)_{t \in T}$  processus du second ordre si  $\forall t \in T$ ,  $\mathbb{E}[X_t^2] < \infty$ .

$$\Rightarrow \begin{cases} \text{fonction espérance : } m(t) = \mathbb{E}[X_t] \\ \text{fonction variance : } \sigma^2(t) = \mathbb{E}[(X_t - \mathbb{E}[X_t])^2] \\ \text{fonction covariance : } r(s, t) = \mathbb{E}[(X_s - \mathbb{E}[X_s])(X_t - \mathbb{E}[X_t])] \\ \text{fonction corrélation : } \rho(s, t) = \frac{r(s, t)}{\sigma(s)\sigma(t)} \in [-1, 1] \end{cases}$$

**Remarque :**  $\rho(s, t) \in [-1, 1]$  par l'Inégalité de Cauchy-Schwarz.



# Exemples de processus stochastiques

Exemples de processus stochastiques :

- Fonctions réelles ;
- Suites de variables indépendantes identiquement distribuées ;
- Marches aléatoires ;
- Chaînes de Markov ;
- Mouvement brownien et équations de diffusion.

**Remarque** : On peut aussi considérer des processus à valeurs dans  $\mathbb{R}^d$  (multidimensionnels), à indice dans  $\mathbb{R}^p$  (Champs),...

# Série chronologique

## Définition

*Série chronologique (temporelle) : processus aléatoire indicé par  $\mathbb{Z}$  ou  $\mathbb{N}$  (si  $T = \mathbb{R}$  ou  $[0, 1]$ , on parle de processus à temps continu).*

**Conséquence** : Une série chronologique : suite de v.a. non forcément i.i.d.

## Propriété

*Si  $U$  et  $V$  sont 2 v.a.i. alors  $\text{cov}(U, V) = 0$ . Réciproque fausse en général.*

## Définition

- *Bruit blanc (fort) : suite de v.a.i.i.d. centrées d'ordre 2.*
- *Bruit blanc faible : suite de v.a.i.d. centrées d'ordre 2 non corrélées.*

**Remarque** : Parfois juste  $\mathbf{E}[|X|^\alpha] < \infty$  avec  $1 \leq \alpha < 2$ .

# Processus gaussien

## Définition

*Processus gaussien*  $X = (X_t)_{t \in T} : \forall k \in \mathbb{N}^*, \forall (t_1, \dots, t_n) \in T^n, (X_{t_1}, \dots, X_{t_n})$  vecteur gaussien.

**Remarque :**  $(X_{t_1}, \dots, X_{t_n})$  vecteur gaussien  $\iff \forall (u_1, \dots, u_n) \in \mathbb{R}^n, u_1 X_{t_1} + \dots + u_n X_{t_n}$  v.a. gaussienne.

## Propriété

- Soit  $(X, Y)$  un vecteur gaussien. Alors :  
 $(X \text{ indépendant de } Y) \iff (\text{cov}(X, Y) = 0)$ .
- Un processus gaussien est intégralement défini par ses fonctions espérance  $m(\cdot)$  et covariance  $r(\cdot, \cdot)$ .
- Réciproquement, à chaque  $m(\cdot)$  et  $r(\cdot, \cdot)$  t.q.  $\forall k \in \mathbb{N}^*, (t_1, \dots, t_k) \in T^k, (r(t_i, t_j))_{1 \leq i, j \leq k}$  matrice  $\geq 0$ , on définit un unique processus gaussien.

# Plan du cours

# Stationnarité

## Définition

Une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  est (strictement) stationnaire lorsque  $\forall n \in \mathbb{N}^*, \forall (t_1, \dots, t_n) \in \mathbb{Z}^n, \forall c \in \mathbb{Z},$


$$(X_{t_1}, \dots, X_{t_n}) \stackrel{\mathcal{L}}{\sim} (X_{t_1+c}, \dots, X_{t_n+c}).$$

$\implies$  Si  $(X_t)$  stationnaire,  $(X_t)$  suite identiquement distribuée

**Contre-ex.** :  $(\varepsilon_t)$  v.a.i.i.d.  $\mathcal{N}(0, 1)$ ,  $X_t = \frac{1}{\sqrt{t}} \sum_{k=1}^t \varepsilon_k$  : i.i.d. non stationnaire !

**Remarque** : On pourra montrer que pour tout  $(u_1, \dots, u_n) \in \mathbb{R}^n,$

$$\phi_{(X_{t_1}, \dots, X_{t_n})}(u_1, \dots, u_n) = \mathbb{E}(e^{i \sum_{j=1}^n u_j X_{t_j}}) = \phi_{(X_{t_1+c}, \dots, X_{t_n+c})}(u_1, \dots, u_n).$$

$\implies$  Utilisation de la fonction caractéristique de la somme de 2 v.a. indépendantes qui vaut le produit des fonctions caractéristiques. 

# Exemples de séries stationnaires

## Exemple de séries stationnaires :

- Une suite de v.a.i.i.d. est stationnaire.

### Démonstration.

On a pour tout  $(u_1, \dots, u_n) \in \mathbb{R}^n$ ,

$$\phi_{(X_{t_1}, \dots, X_{t_n})}(u_1, \dots, u_n) = \prod_{i=1}^n \phi_X(u_i) = \phi_{(X_{t_1+c}, \dots, X_{t_n+c})}(u_1, \dots, u_n)$$



### Propriété

$$(X_t)_{t \in \mathbb{Z}} \text{ stationnaire} \iff \forall n \in \mathbb{N}^*, c \in \mathbb{Z}, (X_1, \dots, X_n) \stackrel{\mathcal{L}}{\sim} (X_{1+c}, \dots, X_{n+c})$$

### Démonstration.

$\implies$  On prend  $(t_1, \dots, t_n) = (1, \dots, n)$ .

$\impliedby$  Soit  $n \in \mathbb{N}^*$  et  $(t_1, \dots, t_n) \in \mathbb{Z}^n$ . On note  $\ell = 1 + \max_{1 \leq i \leq n} (t_i) - \min_{1 \leq i \leq n} (t_i)$ . Alors

$(X_1, \dots, X_\ell) \stackrel{\mathcal{L}}{\sim} (X_{1+c}, \dots, X_{\ell+c}) \stackrel{\mathcal{L}}{\sim} (X_{1+c+c'}, \dots, X_{\ell+c+c'})$ . Pour  $c = \max_{1 \leq i \leq n} (t_i) - \ell$ , on

obtient  $(X_{\min_{1 \leq i \leq n} (t_i)}, \dots, X_{\max_{1 \leq i \leq n} (t_i)}) \stackrel{\mathcal{L}}{\sim} (X_{\min_{1 \leq i \leq n} (t_i)+c'}, \dots, X_{\max_{1 \leq i \leq n} (t_i)+c'})$  pour tout

$c' \in \mathbb{Z}$ . Ceci est donc aussi vrai pour tout sous-vecteur inclus dans

$(X_{\min_{1 \leq i \leq n} (t_i)}, \dots, X_{\max_{1 \leq i \leq n} (t_i)})$ , donc pour  $(X_{t_1}, \dots, X_{t_n})$ .



## Exemples de séries stationnaires (2)

- Une chaîne de Markov homogène sous mesure invariante est stationnaire.

### Démonstration.

Pour tout  $k \in \mathbb{N}$ ,  $\mathbb{P}(X_{k+1} = x_j | (X_0, \dots, X_k)) = \mathbb{P}(X_{k+1} = x_j | X_k = x_i) = Q_{i,j}$  et  $X_1 \stackrel{\mathcal{L}}{\sim} \mu$  mesure invariante, soit  $\mu Q = \mu$  (cas d'un espace d'état  $\{x_1, \dots, x_m\}$  fini). Pour  $n \in \mathbb{N}^*$ , et tout  $(i_1, \dots, i_n) \in \{1, \dots, m\}^n$ ,

$$\begin{aligned}\mathbb{P}((X_1, \dots, X_n) = (x_{i_1}, \dots, x_{i_n})) &= \mathbb{P}(X_n = x_{i_n} | X_{n-1} = x_{i_{n-1}}) \times \dots \times \mathbb{P}(X_2 = x_{i_2} | X_1 = x_{i_1}) \\ &\quad \times \mathbb{P}(X_1 = x_{i_1}) \\ &= Q(i_{n-1}, i_n) \times \dots \times Q(i_1, i_2) \mu(x_{i_1}).\end{aligned}$$

Par le même raisonnement, comme la loi de  $X_{1+c}$  est aussi  $\mu$  puisque  $\mu Q = \mu$ , on a aussi  $\mathbb{P}((X_{1+c}, \dots, X_{n+c}) = (x_{i_1}, \dots, x_{i_n})) = Q(i_{n-1}, i_n) \times \dots \times Q(i_1, i_2) \mu(x_{i_1})$ , donc  $(X_{1+c}, \dots, X_{n+c}) \stackrel{\mathcal{L}}{\sim} (X_1, \dots, X_n)$ . □

## Exemples de séries stationnaires (3)

- Si  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc et  $X_t = \varepsilon_t + \alpha \varepsilon_{t-1}$  pour tout  $t \in \mathbb{Z}$ , où  $\alpha \in \mathbb{R}$ , alors  $(X_t)_{t \in \mathbb{Z}}$  est stationnaire.

### Démonstration.

Soit  $n \in \mathbb{N}^*$  et  $(t_1, \dots, t_n) \in \mathbb{Z}^n$ . Comme  $(\varepsilon_t)$  est stationnaire, pour tout  $c \in \mathbb{Z}$ ,

$$(\varepsilon_{t_1}, \varepsilon_{t_1-1}, \varepsilon_{t_2}, \varepsilon_{t_2-1}, \dots, \varepsilon_{t_n}, \varepsilon_{t_n-1}) \stackrel{\mathcal{L}}{\sim} (\varepsilon_{t_1+c}, \varepsilon_{t_1+c-1}, \varepsilon_{t_2+c}, \varepsilon_{t_2+c-1}, \dots, \varepsilon_{t_n+c}, \varepsilon_{t_n+c-1}).$$

Si deux vecteurs aléatoires  $U$  et  $V$  vérifient  $U \stackrel{\mathcal{L}}{\sim} V$  alors pour  $g : \mathbb{R}^p \rightarrow \mathbb{R}^q$  fonction mesurable  $g(U) \stackrel{\mathcal{L}}{\sim} g(V)$ . Ici avec  $g : (x_1, \dots, x_{2n}) \in \mathbb{R}^{2n} \mapsto (x_1 + \alpha x_2, x_3 + \alpha x_4, \dots, x_{2n-1} + \alpha x_{2n})$  on a  $g(\varepsilon_{t_1}, \varepsilon_{t_1-1}, \varepsilon_{t_2}, \varepsilon_{t_2-1}, \dots, \varepsilon_{t_n}, \varepsilon_{t_n-1}) = (X_{t_1}, \dots, X_{t_n})$  et donc  $(X_{t_1}, \dots, X_{t_n}) \stackrel{\mathcal{L}}{\sim} (X_{t_1+c}, \dots, X_{t_n+c})$  □

**Remarque :** Le résultat serait aussi valable pour tout  $(\varepsilon_t)$  stationnaire.



## Exercice

**Exercice :**  $(\varepsilon_t)_{t \in \mathbb{Z}}$  bruit blanc et  $X_t = \varepsilon_t \varepsilon_{t-1}$  pour tout  $t \in \mathbb{Z}$ . Montrer que  $(X_t)_{t \in \mathbb{Z}}$  est stationnaire.  $(X_t)$  bruit blanc faible ? Bruit blanc ?

### Démonstration.

On reprend la preuve précédente mais avec  $g : (x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mapsto (x_1 x_2, x_2 x_3, \dots, x_n x_{n+1})$ . On a bien  $(X_t)$  suite de v.a.i.d. et  $\mathbb{E}(X_t^2) = \mathbb{E}(\varepsilon_t^2 \varepsilon_{t-1}^2) = \mathbb{E}(\varepsilon_t^2) \mathbb{E}(\varepsilon_{t-1}^2) < \infty$ . On a  $\mathbb{E}(X_t) = \mathbb{E}(\varepsilon_t \varepsilon_{t-1}) = \mathbb{E}(\varepsilon_t) \mathbb{E}(\varepsilon_{t-1})$  par indépendance, d'où  $\mathbb{E}(X_t) = 0$ .  
 $\text{cov}(X_t, X_s) = \mathbb{E}(\varepsilon_t \varepsilon_{t-1} \varepsilon_s \varepsilon_{s-1}) = 0$  directement si  $|s - t| > 1$ . Si  $|t - s| = 1$ , par exemple  $s = t + 1$ ,  $\text{cov}(X_t, X_s) = \mathbb{E}(\varepsilon_t^2 \varepsilon_{t-1} \varepsilon_{t+1}) = 0$  également. Donc  $(X_t)$  bruit blanc faible.

$(X_t)$  bruit blanc si  $X_t$  indépendant de  $X_{t-1}$ . Or  $\text{cov}(|X_t|, |X_{t-1}|) = \mathbb{E}(|\varepsilon_0|)^2 (\mathbb{E}(\varepsilon_0^2) - \mathbb{E}(|\varepsilon_0|)^2)$ . Mais d'après l'Inégalité de Cauchy-Schwarz,  $\mathbb{E}(1 \times |U|)^2 \leq \mathbb{E}(U^2)$  avec égalité si  $|U|$  colinéaire à 1. Donc, sauf si la loi de  $\varepsilon$  est de la forme  $1/2 \delta_a + 1/2 \delta_{-a}$  avec  $a \in [0, \infty[$  alors  $\text{cov}(|X_t|, |X_{t-1}|) > 0$ , donc non indépendance, donc  $(X_t)$  n'est pas un bruit blanc. Si la loi de  $\varepsilon$  est  $1/2 \delta_a + 1/2 \delta_{-a}$ , on vérifie alors que c'est un bruit blanc. □

# Stationnarité : propriétés

## Propriété

Si  $(X_t)$  série chronologique stationnaire et d'ordre 2

- les fonctions  $m(t) = \mathbb{E}[X_t]$  et  $\sigma^2(t) = \text{var}(X_t)$  sont constantes
- $\exists h : \mathbb{N} \rightarrow \mathbb{R}$  telle que 
$$\begin{cases} r(s, t) = \text{cov}(X_s, X_t) = h(|t - s|) \\ \rho(s, t) = \text{cor}(X_s, X_t) = \frac{h(|t-s|)}{h(0)} \end{cases} .$$

## Démonstration.

- Stationnarité  $\implies$  i.d. Donc  $m(t) = m(0)$  et  $\sigma^2(t) = \sigma^2(0)$  pour tout  $t \in \mathbb{Z}$ .
- Comme  $(X_t)$  stationnaire, pour tout  $(s, t) \in \mathbb{Z}^2$ ,  $(X_s, X_t) \stackrel{\mathcal{L}}{\sim} (X_0, X_{t-s})$  avec  $c = -s$ , d'où  $r(s, t) = \text{cov}(X_0, X_{t-s})$ , fonction de  $t - s$ . Et  $(X_s, X_t) \stackrel{\mathcal{L}}{\sim} (X_{s-t}, X_0)$ , d'où  $r(s, t) = \text{cov}(X_0, X_{s-t})$ , même fonction mais de  $s - t$ . La fonction est donc paire et  $h$  existe. Finalement,  $\sigma^2(s) = r(s, s) = h(0)$ , d'où le résultat pour la corrélation. □

# Stationnarité : propriétés

## Conséquences :

- Si  $m(t)$  non constante, alors  $(X_t)$  admet une tendance et/ou une saisonnalité et est non stationnaire.
- Une marche aléatoire, définie par  $X_t = \sum_{k=1}^t \varepsilon_k$  pour  $t \in \mathbb{N}^*$  et  $(\varepsilon_t)$  bruit blanc, n'est pas stationnaire car  $\sigma^2(t) = t\sigma_\varepsilon^2$ .

## Définition

Soit  $(X_t)_{t \in \mathbb{Z}}$  série chronologique du second ordre. On dit que  $(X_t)$  est stationnaire d'ordre 2 lorsque :

- 1 son espérance  $m(t)$  est constante ;
- 2 sa covariance  $r(s, t) = \text{cov}(X_s, X_t)$  est une fonction de  $|t - s|$ .

# Stationnarité : propriétés (2)

## Propriété

- (Stationnarité)  $\implies$  (Stationnarité d'ordre 2), mais réciproque fausse.
- Si  $(X_t)$  série gaussienne, (Stationnarité)  $\iff$  (Stationnarité d'ordre 2)

## Démonstration.

- Soit la suite de v.a. indépendantes  $(X_n)$  à valeurs dans  $\{-n, 0, n\}$  telle que  $\mathbf{P}(X_n = n) = \mathbf{P}(X_n = -n) = \frac{1}{2n^2}$  et  $\mathbf{P}(X_n = 0) = 1 - \frac{1}{n^2}$ . Alors  $m(n) = 0$ ,  $r(s, t) = 1$  si  $s = t$  et 0 sinon, soit  $r(s, t) = h(|s - t|)$  avec  $h(0) = 1$  et 0 sinon :  $(X_n)$  stationnaire d'ordre 2 mais pas stationnaire (support de la loi dépend de  $n$ )!
- Vrai car la loi de tout vecteur  $(X_1, \dots, X_n)$  dépend uniquement de  $m(\cdot)$  et de  $\Gamma = (r(i, j))_{i, j}$ .  $\square$

**Exercice :**  $A$  et  $B$  v.a.i. centrées de même variance  $\sigma^2$  et pour  $t \in \mathbb{Z}$ ,  $X_t = A \cos\left(\frac{2\pi t}{3}\right) + B \sin\left(\frac{2\pi t}{3}\right)$ .  $(X_t)$  stationnaire d'ordre 2 ? Stationnaire ?

## Démonstration.

$\mathbf{E}(X_t) = 0$  pour  $t \in \mathbb{Z}$ .  $r(s, t) = \text{cov}\left(A \cos\left(\frac{2\pi t}{3}\right) + B \sin\left(\frac{2\pi t}{3}\right), A \cos\left(\frac{2\pi s}{3}\right) + B \sin\left(\frac{2\pi s}{3}\right)\right) = \sigma^2 \left( \cos\left(\frac{2\pi t}{3}\right) \cos\left(\frac{2\pi s}{3}\right) + \sin\left(\frac{2\pi t}{3}\right) \sin\left(\frac{2\pi s}{3}\right) \right) = \sigma^2 \cos\left(\frac{2\pi(t-s)}{3}\right) = h(|t-s|)$ . Pas statio.  $\square$

## Proposition

Soit  $(X_t)$  série stationnaire (ou stationnaire d'ordre 2). Si  $I \subset \mathbb{Z}$  et  $(a_i)_{i \in I}$  famille de réels, soit  $(Y_t)$  la combinaison (filtre) linéaire de  $X$  définie par  $Y_t = \sum_{i \in I} a_i X_{t-i}$  pour  $t \in \mathbb{Z}$ . Alors

- 1 *I fini* : si  $X$  dans  $\mathbb{L}^p$  ( $p \geq 1$ ) alors  $Y$  dans  $\mathbb{L}^p$  et si  $X$  est stationnaire (resp. du second ordre) alors  $Y$  est stationnaire (resp. du second ordre).
- 2 *I infini*,  $\sum_{i \in I} |a_i| < \infty$  et  $\sup_{t \in \mathbb{Z}} \mathbf{E}(|X_t|) < \infty \implies \sup_{t \in \mathbb{Z}} \mathbf{E}(|Y_t|) < \infty$  et  $Y$  existe p.s., si  $X$  stationnaire (centré) alors  $Y$  stationnaire (centré).
- 3 *I infini*,  $\sum_{i \in I} |a_i| < \infty$  et  $\sup_{t \in \mathbb{Z}} \mathbf{E}(|X_t|^2) < \infty \implies Y$  existe dans  $\mathbb{L}^2$ , si  $X$  stationnaire (second ordre, gaussien) alors  $Y$  stationnaire (second ordre, gaussien).

# Combinaisons linéaires : preuves

## Démonstration.

(1) Inégalité triangulaire :  $\|Y_t\|_p \leq \sum_{i \in I} |a_i| \|X_{t-i}\|_p < \infty$  pour tout  $t \in \mathbb{Z}$ . Stationarité : avec  $I = \{-m, \dots, m\}$ , si  $X$  stationnaire, alors pour tout  $n \in \mathbb{N}^*$ ,  $(t_1, \dots, t_n) \in \mathbb{Z}^n$  et tout  $c \in \mathbb{Z}$ ,

$$\begin{aligned} & (X_{t_1-m}, X_{t_1-m+1}, \dots, X_{t_1+m}, X_{t_2-m}, \dots, X_{t_n+m}) \\ & \stackrel{\mathcal{L}}{\sim} (X_{t_1-m+c}, X_{t_1-m+1+c}, \dots, X_{t_1+m+c}, X_{t_2-m+c}, \dots, X_{t_n+m+c}). \end{aligned}$$

Avec  $g : \mathbb{R}^{(2m+1)n} \rightarrow \mathbb{R}^n$  telle que

$g(X_{t_1-m}, X_{t_1-m+1}, \dots, X_{t_1+m}, X_{t_2-m}, \dots, X_{t_n+m}) = (\sum_{i=-m}^m a_i X_{t_1-i}, \dots, \sum_{i=-m}^m a_i X_{t_n-i})$ , la fonction  $g$  étant continue donc mesurable, on a

$$\begin{aligned} & g(X_{t_1-m}, X_{t_1-m+1}, \dots, X_{t_1+m}, X_{t_2-m}, \dots, X_{t_n+m}) \\ & \stackrel{\mathcal{L}}{\sim} g(X_{t_1-m+c}, X_{t_1-m+1+c}, \dots, X_{t_1+m+c}, X_{t_2-m+c}, \dots, X_{t_n+m+c}). \end{aligned}$$

donc  $(Y_{t_1}, \dots, Y_{t_n})$  a la même loi que  $(Y_{t_1+c}, \dots, Y_{t_n+c})$  :  $(Y_t)$  est bien stationnaire.

Si  $X$  stationnaire d'ordre 2, on a  $\mathbf{E}(Y_t) = \sum_{i \in I} a_i \mathbf{E}(X_{t-i}) = \text{constante}$ . Pour la covariance, on a  $\text{cov}(Y_s, Y_t) = \sum_{i \in I} \sum_{i' \in I} a_i a_{i'} \text{cov}(X_{t-i}, X_{s-i'}) = \sum_{i \in I} \sum_{i' \in I} a_i a_{i'} r(t-s-i+i')$  en notant  $r(k) = \text{cov}(X_0, X_k)$ . Donc  $\text{cov}(Y_s, Y_t)$  est bien une fonction de  $t-s$ . De plus, on peut intervertir  $i$  et  $i'$  et du fait de la parité de  $r$  on voit bien que  $\text{cov}(Y_s, Y_t)$  est une fonction de  $|t-s|$ .

(2) On a  $\mathbf{E}(|Y_t|) \leq \mathbf{E}(\sum_{i \in I} |a_i| |X_{t-i}|) \leq (\sum_{i \in I} |a_i|) \sup_{j \in \mathbb{Z}} \mathbf{E}(|X_j|)$  d'après le Théorème de Fubini, donc  $\sup_{t \in \mathbb{Z}} \mathbf{E}(|Y_t|) < \infty$ . Pour tout  $\varepsilon > 0$ , avec l'Inégalité de Markov, il existe  $M_0 > 0$  tel que  $\forall M \geq M_0$ ,  $\mathbf{P}(|Y_t| \geq M) \leq \varepsilon$ , donc  $\mathbf{P}(|Y_t| < \infty) = 1$ . □

## Combinaisons linéaires : preuves (2)

### Démonstration.

(2) Stationnarité : on note  $Y_t^{(m)} = \sum_{i=-m}^m a_i X_{t-i}$ , et  $(Y_t^{(m)})$  stationnaire d'après (1). En reprenant la preuve précédente, pour tout  $t$ ,  $\mathbf{E}[|Y_t - Y_t^{(m)}|] \leq (\sum_{|i|>m} |a_i|) \sup_{j \in \mathbb{Z}} \mathbf{E}(|X_j|) \rightarrow 0$  pour  $m \rightarrow \infty$ . De même,  $(Y_{t_1}^{(m)}, \dots, Y_{t_n}^{(m)})$  tend dans  $\mathbb{L}^1$  vers  $(Y_{t_1}, \dots, Y_{t_n})$  lorsque  $m \rightarrow \infty$ , donc on a également convergence en loi. Comme  $(Y_{t_1}^{(m)}, \dots, Y_{t_n}^{(m)})$  a la même loi que  $(Y_{t_1+c}^{(m)}, \dots, Y_{t_n+c}^{(m)})$ , cette égalité a également lieu à la limite :  $(Y_t)$  est bien stationnaire.

Si  $X_t$  est centrée, d'après le Théorème de convergence dominée,  $\mathbf{E}(Y_t) = \sum_{i \in I} a_i \mathbf{E}(X_{t-i}) = 0$  pour tout  $t \in \mathbb{Z}$ .

(3)  $\mathbf{E}[(Y_t - Y_t^{(m)})^2] = \mathbf{E}(\sum_{|i|>m} \sum_{|i'|>m} a_i a_{i'} X_{t-i} X_{t-i'}) \leq (\sum_{|i|>m} |a_i|)^2 \sup_{j \in \mathbb{Z}} \mathbf{E}(|X_j^2|)$  d'après le Théorème de Fubini, donc  $\mathbf{E}(Y_t - Y_t^{(m)})^2 \rightarrow 0$  ( $m \rightarrow \infty$ ) car  $\sum_{i \in I} |a_i| < \infty$  :  $(Y_t)$  existe dans  $\mathbb{L}^2$ .

La stationnarité stricte s'obtient comme dans le cas précédent et la stationnarité d'ordre 2 utilise le point (1) : on a  $(Y_t^{(m)})_t$  qui est stationnaire d'ordre 2 et comme on a convergence dans  $\mathbb{L}^2$  donc en loi, on a bien convergence de l'espérance et de la covariance des  $(Y_t^{(m)})$  :  $(Y_t)$  est bien aussi stationnaire d'ordre 2.

Si  $(X_t)$  est gaussien, il est clair que  $(Y_t^{(m)})$  est gaussien comme combinaison linéaire finie issue d'un vecteur gaussien. La limite en loi d'une suite de v.a. gaussienne est gaussienne si la suite des moments d'ordre 1 et 2 converge, ce qui est le cas.

# Processus linéaires

## Définition

Soit  $I \subset \mathbb{Z}$  et  $(a_i)_{i \in I}$  famille de réels. Pour  $(\varepsilon_t)_{t \in \mathbb{Z}}$  un bruit blanc,  $(Y_t)_{t \in \mathbb{Z}}$  tel que  $Y_t = \sum_{i \in I} a_i \varepsilon_{t-i}$  pour  $t \in \mathbb{Z}$  est appelé un processus linéaire.

## Proposition

Si  $Y$  processus linéaire avec  $Y_t = \sum_{i \in I} a_i \varepsilon_{t-i}$  où  $(\varepsilon_t)_{t \in \mathbb{Z}}$  bruit blanc et  $\sum_{i \in I} |a_i|^2 < \infty$  alors  $\sup_{t \in \mathbb{Z}} \mathbf{E}(Y_t^2) < \infty$  et  $(Y_t)_{t \in \mathbb{Z}}$  stationnaire.

## Démonstration.

Dans le cas d'un processus linéaire, on peut obtenir l'existence et la stationnarité sous la condition  $\sum_{i \in I} a_i^2 < \infty$  qui est plus faible que la condition  $\sum_{i \in I} |a_i| < \infty$ . En effet, on a :

$\mathbf{E}((Y_t - Y_t^{(m)})^2) = \mathbf{E}(\sum_{|i| > m} \sum_{|i'| > m} a_i a_{i'} \varepsilon_{t-i} \varepsilon_{t-i'}) = \sum_{|i| > m} a_i^2 \mathbf{E}[\varepsilon_0^2] \rightarrow 0$  ( $m \rightarrow \infty$ ) dès que  $(\varepsilon_t)$  est un bruit blanc. La preuve de la stationnarité est immédiate (voir (3)). □



# Filtre de bruit blanc faible

## Proposition

Soit  $(\varepsilon_t)_{t \in \mathbb{Z}}$  un bruit blanc faible (v.a.i.d. centrées non corrélées). Si  $Y = (Y_t)$  tel que  $Y_t = \sum_{i \in \mathbb{Z}} a_i \varepsilon_{t-i}$  où et  $\sum_{i \in \mathbb{Z}} |a_i|^2 < \infty$  alors  $\sup_{t \in \mathbb{Z}} \mathbf{E}(Y_t^2) < \infty$ , alors  $(Y_t)_{t \in \mathbb{Z}}$  stationnaire d'ordre 2.

## Démonstration.

On reprend la preuve précédente, et on a encore  $\mathbf{E}[(Y_t - Y_t^{(m)})^2] = \sum_{|i| > m} a_i^2 \mathbf{E}[\varepsilon_0^2] \rightarrow 0$  ( $m \rightarrow \infty$ ) car  $(\varepsilon_t)$  est un bruit blanc faible. Donc  $\sup_{t \in \mathbb{Z}} \mathbf{E}[Y_t^2] < \infty$ . On en déduit que  $\mathbf{E}[|Y_t|] < \infty$  et ainsi  $\mathbf{E}[Y_t]$  existe. De plus pour tout  $m \geq 0$ , avec  $Y^{(m)} = \sum_{i \leq m} a_i \varepsilon_{t-i}$ , on a  $\mathbf{E}[Y_t^{(m)}] = 0$  pour tout  $t$  et  $m$ . Or d'après le Théorème de convergence dominée de Lebesgue,  $\lim_{m \rightarrow \infty} \mathbf{E}[Y_t^{(m)}] = \mathbf{E}[\lim_{m \rightarrow \infty} Y_t^{(m)}]$ . D'où  $\mathbf{E}[Y_t] = \lim_{m \rightarrow \infty} \mathbf{E}[Y_t^{(m)}] = 0$ . Par ailleurs, on a d'après Cauchy-Schwarz  $\text{cov}(Y_s, Y_t) = \mathbf{E}[\varepsilon_0^2] \sum_{i \in I} a_i a_{i+s-t}$  qui est une fonction dépendant de  $|t - s|$  et qui existe. □

# Plan du cours

# Autocorrélation

## Définition

Soit  $(X_k)_{k \in \mathbb{Z}}$  série stationnaire du second ordre.

- $r(k) = \text{cov}(X_0, X_k)$  pour  $k \in \mathbb{Z}$  autocovariance de  $X$ ,  $r(0)$  variance.
- $\rho(k) = r(k)/r(0)$  autocorrélation de  $X$  et  $-1 \leq \rho(k) \leq 1$  pour  $k \in \mathbb{Z}$ .

**Conséquence** : Si  $\rho(k) \neq 0$  pour un  $k \neq 0$  alors  $(X_k)$  suite de v.a. non indépendantes.

## Propriété

Pour  $(s, t) \in \mathbb{Z}^2$ ,  $\text{cov}(X_s, X_t) = r(|t - s|)$  et la matrice de variance-covariance de  $(X_1, \dots, X_n)$  est  $(r(|j - i|))_{1 \leq i, j \leq n}$ .

# Densité spectrale

## Définition

$X = (X_k)_{k \in \mathbb{Z}}$  série stationnaire du second ordre. S'il existe  $f : [-\pi, \pi[ \rightarrow \mathbb{C}$  telle que  $\forall k \in \mathbb{Z}, r(k) = \int_{-\pi}^{\pi} e^{ik\lambda} f(\lambda) d\lambda$ ,  $X$  admet une densité spectrale  $f$ .

**Exemple** : Densité spectrale bruit blanc faible :  $f(\lambda) = \frac{\sigma^2}{2\pi}$  pour  $\lambda \in [-\pi, \pi]$ .

## Propriété

Soit  $X = (X_k)_{k \in \mathbb{Z}}$  série stationnaire d'ordre 2.

- 1  $r(\cdot)$  telle que  $\sum |r(k)|^2 < \infty$   
 $\iff f(\lambda) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} r(k) e^{-ik\lambda}$  p.p. sur  $[-\pi, \pi]$  et  $f \in \mathbb{L}^2([-\pi, \pi])$
- 2  $r(\cdot)$  telle que  $\sum |r(k)| < \infty$   
 $\iff f(\lambda) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} r(k) e^{-ik\lambda}$  sur  $[-\pi, \pi]$  et  $f$  continue sur  $[-\pi, \pi]$

## Démonstration.

1/ Théorème de Carleson et 2/ résultat classique sur les séries de Fourier ( $r(\cdot)$  coefficients de Fourier). □

## Moments empiriques

Si  $(X_1, \dots, X_n)$  trajectoire observée, comment estimer  $r(k)$  et  $\rho(k)$  ?

### Définition

Si  $(X_1, \dots, X_n)$  trajectoire observée de  $(X_t)$  série chronologique de 2nd ordre

- La moyenne empirique est :  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  ;

- La variance empirique est :  $\overline{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$  ;

Autre estimateur (biaisé) :  $\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2$  ;

- L'autocovariance empirique est :  $\widehat{r}(k) = \frac{1}{n} \sum_{i=1}^{n-|k|} (X_i - \bar{X}_n)(X_{i+k} - \bar{X}_n)$  ;

- L'autocorrélation empirique (correlogram) est :  $\widehat{\rho}(k) = \frac{\widehat{r}(k)}{\widehat{r}(0)} = \frac{\widehat{r}(k)}{\widehat{\sigma}_n^2}$ .

# Convergence pour une suite de v.a.i.i.d.

## Propriété

Si  $(X_k)_{k \in \mathbb{Z}}$  est une suite de v.a.i.i.d. avec  $\mathbb{E}[X_0^2] < \infty$ ,

①  $\bar{X}_n \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} m = \mathbb{E}[X_0]$  et  $\overline{\sigma_n^2}$  ou  $\widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2 = \text{var}(X_0)$  ;

②  $\widehat{r}(0) = \widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2$  et  $\widehat{r}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0$  pour tout  $k \neq 0$  ;

③  $\widehat{\rho}(0) = 1$  et  $\widehat{\rho}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0$  pour tout  $k \neq 0$ .

## Attention !

- Si  $(X_k)$  série stationnaire avec v.a. non indépendantes, les convergences peuvent être vraies ... ou fausses (exemple  $X_k = X_0$  pour tout  $k$ ) !
- Si  $(X_k)$  série non stationnaire, les convergences peuvent être vraies ... ou fausses, même si indépendance !

## Démonstration.

- 1 Evident par la LFGN pour  $\bar{X}_n$ . Pour  $\widehat{\sigma}_n^2$ , on écrit  $\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2$ . Le premier terme tend vers  $\mathbb{E}[X_0^2]$  par LFGN, le second terme vers  $m^2$  par ce qui est précède et le fait que la fonction  $x \rightarrow x^2$  est continue. On utilise alors le Lemme de Slutsky pour achever la preuve.
- 2 Evident pour  $\widehat{r}(0)$ . Pour  $\widehat{r}(k)$ , moins simple car  $X_1 X_{1+k}$  non indépendant de  $X_{k+1} X_{2k+1}$ , voir la preuve un peu plus loin...
- 3 Facile par le Lemme de Slutsky et le fait que la fonction  $x \rightarrow 1/x$  est continue sur  $]0, \infty[$ .



Peut-on obtenir des convergences si  $(X_n)_{n \in \mathbb{Z}}$  stationnaire sans indépendance ?

# Convergence et dépendance

## Proposition

$X = (X_k)_{k \in \mathbb{Z}}$  série stationnaire d'ordre 2 avec  $\sum_{k \geq 0} |r_X(k)| < \infty$ . Alors :

$$\text{var}(\sqrt{n}(\bar{X}_n - \mathbb{E}(X_0))) \xrightarrow{n \rightarrow +\infty} \sum_{k \in \mathbb{Z}} r_X(k) \quad \implies \quad \bar{X}_n \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \mathbb{E}(X_0).$$

## Démonstration.

On a  $\text{var}(\sqrt{n}(\bar{X}_n - m)) = n \text{var}(\bar{X}_n) = \frac{1}{n} \sum_{j=1}^n \sum_{i=1}^n r_X(|j-i|)$  où  $r_X$  est l'autocovariance de  $X$ .  
Alors  $n \text{var}(\bar{X}_n) = \frac{1}{n} \sum_{k=-n}^n (n-|k|) r_X(|k|)$  avec  $\sum_{k \in \mathbb{Z}} |r_X(k)| < \infty$ . Par suite

$$\begin{aligned} \frac{1}{n} \sum_{k=-n}^n (n-|k|) r_X(|k|) &= \sum_{k=-n}^n r_X(|k|) - \frac{1}{n} \sum_{k=-n}^n |k| r_X(|k|). \\ \text{Mais } \frac{1}{n} \sum_{k=-n}^n |k| r_X(|k|) &= \frac{1}{n} \sum_{k=-[\sqrt{n}]}^{[\sqrt{n}]} |k| r_X(|k|) + \frac{2}{n} \sum_{k=[\sqrt{n}]+1}^n k r_X(|k|) \\ \implies \frac{1}{n} \sum_{k=-n}^n |k| |r_X(|k|)| &\leq \frac{1}{\sqrt{n}} \sum_{k=-[\sqrt{n}]}^{[\sqrt{n}]} |r_X(|k|)| + 2 \sum_{k=[\sqrt{n}]+1}^n |r_X(|k|)| \end{aligned}$$

En conséquence,  $\frac{1}{n} \sum_{k=-n}^n (n-|k|) r_X(|k|) \xrightarrow[n \rightarrow +\infty]{} 0$  car  $\sum_{k \in \mathbb{Z}} |r_X(k)| < \infty$   
 $\xrightarrow[n \rightarrow +\infty]{} \sum_{k \in \mathbb{Z}} r_X(k) = 2\pi f(0)$ . □



## Convergence et dépendance (2)

**Conséquence** : Si  $X = (X_k)_{k \in \mathbb{Z}}$  processus gaussien stationnaire d'espérance  $m$  avec  $\sum |r(k)| < \infty$ , alors  $\sqrt{n}(\bar{X}_n - m) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \sum_{k \in \mathbb{Z}} r_X(k))$ .

Le cas gaussien permet donc d'avoir un TLC sous des conditions faibles. Cas général, plus difficile... Voici un cadre possible :

### Définition

$X = (X_t)_{t \in \mathbb{Z}}$  processus stationnaire. On dit que  $X$  est  $M$ -dépendant où  $M^* \in \mathbb{N}$ , si  $\forall t \in \mathbb{Z}$ , alors  $(X_k)_{k \leq t}$  est indépendant de  $(X_{k+M+1})_{k \geq t}$ .

### Théorème

$X = (X_t)_{t \in \mathbb{Z}}$  processus d'ordre 2 stationnaire d'autocovariance  $r_X$ , et  $M$ -dépendant, avec  $M \in \mathbb{N}^*$ . Alors

$$\sqrt{n}(\bar{X}_n - \mathbb{E}(X_0)) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \gamma^2) \quad \text{avec} \quad \gamma^2 = \sum_{k=-M}^M r_X(k).$$

## Convergence et dépendance (3)

### Démonstration.

Pour simplifier les notations, on commence par supposer que  $\mathbf{E}(X_0) = 0$  (on retrouve le TLC pour  $\mathbf{E}(X_0)$  quelconque en retirant cette espérance).

La preuve se fait en découpant  $n$  en  $r$  morceaux, blocs, de taille  $p$ , avec  $r$  et  $p$  qui tendent vers l'infini (donc  $n = rp$ ) en considérant donc  $p > M$ .

Pour  $i = 1, \dots, r$ , soit  $Z_i = \frac{1}{\sqrt{p}} \sum_{k=1}^{p-M} X_{(i-1)p+k}$  et  $U_i^{(p)} = \frac{1}{\sqrt{p}} \sum_{k=0}^{M-1} X_{ip-k}$ , d'où

$$\sqrt{n} \bar{X}_i = \frac{1}{\sqrt{rp}} \sum_{i=1}^n X_i = \sqrt{r} \bar{Z}_r + \sqrt{r} \bar{U}_r^{(p)} \text{ où } \bar{Z}_r = \frac{1}{r} \sum_{i=1}^r Z_i \text{ et } \bar{U}_r^{(p)} = \frac{1}{r} \sum_{i=1}^r U_i^{(p)}.$$

Les v.a.  $Z_i$  sont i.i.d. du fait de la stationnarité de  $X$  et de la  $M$ -dépendance. A  $p$  fixé, on peut appliquer un TLC classique aux variables  $Z_i$  et on obtient que  $\sqrt{r} \bar{Z}_r \xrightarrow[r \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \gamma_p^2)$  avec

$$\gamma_p^2 = \frac{1}{p} \sum_{i=-(p-M)}^{p-M} ((p-M) - |i|) r X(i) \xrightarrow[p \rightarrow +\infty]{} \gamma^2.$$

Les v.a.  $U_i^{(p)}$  sont également i.i.d. avec  $\text{var}(U_i^{(p)}) \leq \frac{M\gamma^2}{p}$  et ainsi  $\text{var}(\sqrt{r} \bar{U}_r^{(p)}) \leq \frac{M\gamma^2}{p}$ . On a

donc  $\text{var}(\sqrt{r} \bar{U}_r^{(p)}) \leq \frac{M\gamma^2}{p}$  qui tend vers 0 quand  $p \rightarrow \infty$ , d'où  $\sqrt{r} \bar{U}_r^{(p)}$  tend en probabilité vers 0

quand  $p \rightarrow \infty$ . Et  $\sqrt{r} \bar{Z}_r \xrightarrow[r \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \gamma_p^2)$  qui tend en loi vers  $\mathcal{N}(0, \gamma_p^2)$  quand  $p \rightarrow \infty$ . On

conclut en utilisant  $\bar{X}_i = \sqrt{r} \bar{Z}_r + \sqrt{r} \bar{U}_r^{(p)}$  et le Théorème de Slutsky. □

## Convergence et dépendance (4)

### Théorème (TLC pour les autovariances de v.a.i.i.d.)

$(X_t)_{t \in \mathbb{Z}}$  suite de v.a.i.i.d. de variance  $\sigma_X^2 < \infty$ . Alors pour  $K_{\max} \in \mathbb{N}^*$  fixé,

$$\sqrt{n} (\hat{r}_X(k))_{1 \leq k \leq K_{\max}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, \sigma_X^4 I_{K_{\max}}).$$

### Démonstration.

On commence par supposer que  $(X_t)$  est centrée et on définit  $\tilde{r}_X(p) = \frac{1}{n} \sum_{k=1}^{n-p} X_{k+p} X_k$ . Il est clair que les  $(X_t)$  étant indépendantes, alors les  $Y_t = X_t X_{t+p}$  forment une suite stationnaire  $p$ -dépendante. De plus  $\mathbf{E}(Y_t) = 0$  et  $\text{var}(Y_t) = \sigma_X^4$  pour tout  $|p| \geq 1$ . On va appliquer le TLC pour le processus  $K_{\max}$ -dépendant  $Z_t = u_1 X_t X_{t+1} + \dots + u_{K_{\max}} X_t X_{t+K_{\max}}$ , où  $(u_1, \dots, u_{K_{\max}})$  est un vecteur quelconque de  $\mathbb{R}^{K_{\max}}$ . Les  $(Z_t)$  forment une suite stationnaire  $K_{\max}$ -dépendante avec  $\mathbf{E}(Z_t) = 0$  et  $\text{var}(Z_t) = \sigma_X^4 (u_1^2 + \dots + u_{K_{\max}}^2)$  car  $\text{cov}(X_t X_{t+i}, X_t X_{t+j}) = 0$  sauf pour  $i = j$  où l'on a  $\sigma_X^4$ . On applique le TLC pour les processus  $K_{\max}$ -dépendants et on a :

$$\sqrt{n - K_{\max}} \left( \frac{1}{n - K_{\max}} \sum_{i=1}^{n - K_{\max}} Z_t - 0 \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \sigma_X^4 (u_1^2 + \dots + u_{K_{\max}}^2)).$$

Comme  $K_{\max}$  ne dépend pas de  $n$  on peut en déduire le même TLC en remplaçant  $n - K_{\max}$  par  $n$  par le Lemme de Slutsky. Il est clair que ce résultat s'étend à une série stationnaire d'espérance  $m$  en considérant dans ce cas  $\tilde{r}_X(p) = \frac{1}{n} \sum_{k=1}^{n-p} (X_{k+p} - m)(X_k - m)$ . □

# Suite de la preuve

## Démonstration.

Il nous reste à considérer pour  $1 \leq p \leq K_{\max}$ ,

$$\begin{aligned}\tilde{r}_X(p) - \hat{r}_X(p) &= \frac{1}{n} \sum_{k=1}^{n-p} (X_{k+p} - m)(X_k - m) - (X_{k+p} - \bar{X}_n)(X_k - \bar{X}_n) \\ &= -\frac{1}{n} (m - \bar{X}_n) \sum_{k=1}^{n-p} (X_{k+p} - m) + (X_k - m) - (m - \bar{X}_n)^2\end{aligned}$$

$$\mathbf{E}[|\tilde{r}_X(p) - \hat{r}_X(p)|] \leq \frac{n-p}{n} \mathbf{E}[(m - \bar{X}_n)^2] + 2 \sqrt{\mathbf{E}[(m - \bar{X}_n)^2]} \sqrt{\frac{1}{n^2} \sum_{k=1}^{n-p} \sigma_X^2},$$

en utilisant l'inégalité de Cauchy-Schwarz. Or  $\mathbf{E}[(m - \bar{X}_n)^2] = \sigma_X^2/n$  et après calculs on en déduit donc que  $\sqrt{n} \mathbf{E}[|\tilde{r}_X(p) - \hat{r}_X(p)|] \xrightarrow{n \rightarrow +\infty} 0$  et par l'inégalité de Markov

$$\sqrt{n} \mathbf{E}[|\tilde{r}_X(p) - \hat{r}_X(p)|] \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0.$$

En utilisant tout ce qui précède, on en déduit que :

$$\sqrt{n} (u_1 \hat{r}_X(1) + \dots + u_{K_{\max}} \hat{r}_X(K_{\max})) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \sigma_X^4 (u_1^2 + \dots + u_{K_{\max}}^2)).$$

De ceci on en déduit le résultat final vectoriel. □

## Convergence et dépendance (5)

### Théorème (TLC pour les autocorrélations de v.a.i.i.d.)

$(X_t)_{t \in \mathbb{Z}}$  suite de v.a.i.i.d. de variance  $\sigma_X^2 < \infty$ . Alors pour  $K_{\max} \in \mathbb{N}^*$  fixé,  
$$\sqrt{n} (\hat{\rho}_X(k))_{1 \leq k \leq K_{\max}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, I_{K_{\max}}).$$

### Démonstration.

Ce TLC se déduit du précédent par le Lemme de Slutsky en utilisant

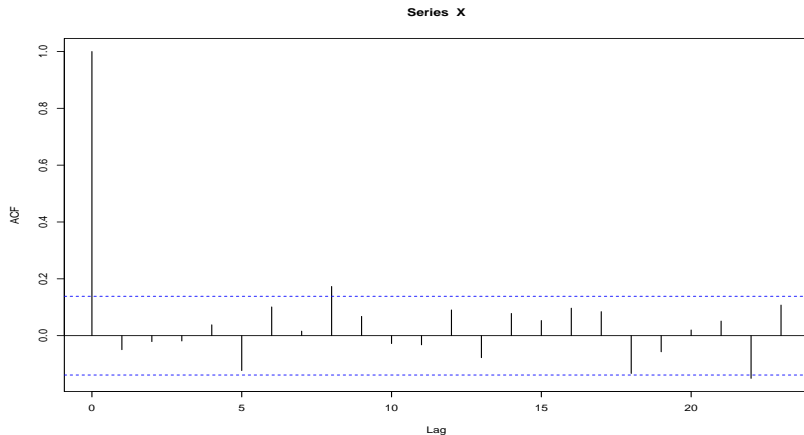
$$\sqrt{n} (\hat{\rho}_X(k))_{1 \leq k \leq K_{\max}} = \sqrt{n} (\hat{r}_X(k)/\hat{r}_X(0))_{1 \leq k \leq K_{\max}} \text{ et } \hat{r}_X(0) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma_X^2. \quad \square$$

$\implies$  Pour tout  $k \geq 1$ , intervalle de confiance à 95% pour  $\hat{\rho}_X(k)$  :

$$\left[ -\frac{1.96}{\sqrt{n}}, \frac{1.96}{\sqrt{n}} \right]$$

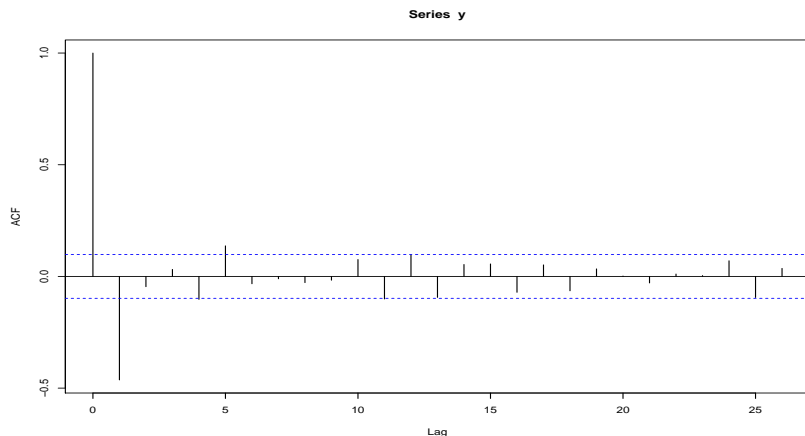
$\implies$  ACF (correlogram) d'une série temporelle.

# ACF



ACF (Auto-Correlogram Function) d'une suite  $(\varepsilon_k)$  de v.a.i.i.d. pour  $n = 200$

# ACF



ACF d'une suite de v.a.  $(X_k)$  telle que  $X_k = \varepsilon_k - \varepsilon_{k-1}$  pour  $n = 400$

Exercice : Montrer que  $\rho(1) = -1/2...$

## Vers un test de "blancheur"

ACF intéressant pour visualiser une dépendance mais difficile de conclure.

### Théorème (Test portemanteau)

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  série stationnaire du second ordre. On désire tester :

$$\begin{cases} H_0 : (X_t) \text{ forme une suite de variables indépendantes} \\ H_1 : (X_t) \text{ ne forme pas une suite de variables indépendantes} \end{cases}$$

On définit deux statistique de test :

- La statistique de Box-Pierce :  $\hat{T}_{BP} = n \sum_{k=1}^{K_{\max}} \hat{\rho}_X^2(k)$  ;
- La statistique de Ljung-Box :  $\hat{T}_{LB} = n(n+2) \sum_{k=1}^{K_{\max}} \frac{\hat{\rho}_X^2(k)}{n-k}$ .

Dans les deux cas, sous  $H_0$ , on a

$$\begin{matrix} \hat{T}_{BP} \\ \hat{T}_{LB} \end{matrix} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2(K_{\max}).$$



# Test portemanteau

## Démonstration.

$\hat{T}_{BP} = \left\| \sqrt{n} (\hat{\rho}_X(k))_{1 \leq k \leq K_{\max}} \right\|^2$  et comme  $\sqrt{n} (\hat{\rho}_X(k))_{1 \leq k \leq K_{\max}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, I_{K_{\max}})$  sous  $H_0$  et  $x \rightarrow \|x\|^2$  continue alors  $\hat{T}_{BP} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2(K_{\max})$  qui est la loi de  $\|Z\|^2$  où  $Z$  est un vecteur gaussien centré réduit de taille  $K_{\max}$ . Pour  $\hat{T}_{LB}$  on utilise le Lemme de Slutsky. □

## Choix de $K_{\max}$ :

- Théoriquement il faut  $K_{\max} = o(\sqrt{n})$ .
- Empiriquement, on pourra utiliser  $K_{\max} = 5$  si  $n$  ordre de la centaine,  $K_{\max} = 10$  pour ordre du millier

**Attention !** Si une tendance ou saisonnalité est présente, aucune conclusion n'est possible !

⇒ Il faut estimer et retirer une éventuelle tendance ou saisonnalité

# Décomposition de Cramèr-Wald

Un résultat théorique puissant, avec cependant peu d'intérêt en pratique :

## Théorème (Décomposition de Cramèr-Wald)

$X = (X_t)_{t \in \mathbb{Z}}$  série stationnaire d'ordre 2 tel que  $\int_{-\pi}^{\pi} |\log(f(\lambda))| d\lambda < \infty$ .  
Alors il existe un unique bruit blanc faible  $(\varepsilon_t)_{t \in \mathbb{Z}}$  et  $(a_i)_{i \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$  telle que  $a_0 \geq 0$  et  $\sum_{k \in \mathbb{N}} |a_k|^2 < \infty$ , vérifiant dans  $\mathbb{L}^2$

$$X_t = \mathbb{E}[X_0] + \sum_{i=0}^{\infty} a_i \varepsilon_{t-i} \quad \text{pour tout } t \in \mathbb{Z}.$$

## Démonstration.

Trop dur, nécessite des connaissances poussées en algèbre. □

**Remarque** : Bruit blanc faible et non bruit blanc, sinon toute série stationnaire serait un processus linéaire !

# Plan du cours

# Processus ARMA

## Définition (Yule, 1927 et Slutsky, 1921)

Un processus ARMA( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

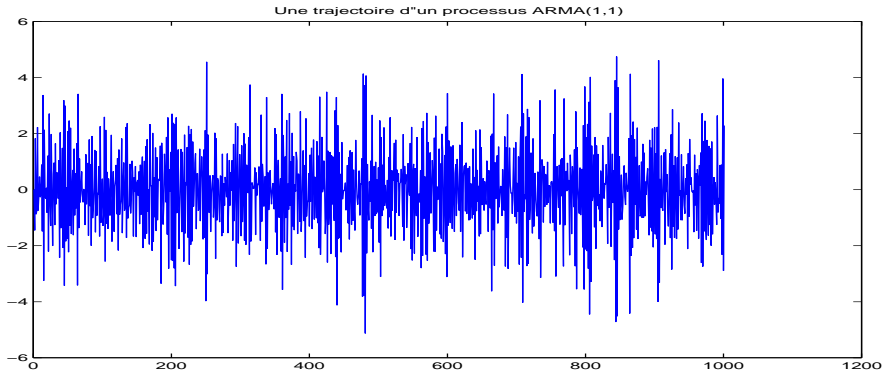
où :

- $(\varepsilon_t)$  bruit blanc de variance  $\sigma_\varepsilon^2$  ;
- $(a_1, \dots, a_p, b_1, \dots, b_q) \in \mathbb{R}^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

## Cas particuliers :

- Si  $q = 0$ , un ARMA( $p, q$ ) est un AR( $p$ ) (Auto-Regressive) ;
- Si  $p = 0$ , un ARMA( $p, q$ ) est un MA( $q$ ) (Moving Average) ;

# Une trajectoire de processus ARMA(1,1)



## Processus ARMA : trois exemples

**Exemple 1** : Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ .  
Alors :

$$\begin{aligned} X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^2(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &\quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= \alpha^{n+1} X_{t-n-1} + \sum_{k=0}^n \alpha^k \varepsilon_{t-k}, \quad n \geq 0. \end{aligned}$$

Donc pour tout  $t \in \mathbb{Z}$ , s'il existe une solution d'ordre 2 stationnaire

$$\mathbb{E}\left[\left(X_t - \sum_{k=0}^n \alpha^k \varepsilon_{t-k}\right)^2\right] = |\alpha|^{2n+1} \mathbb{E}[X_{t-n-1}^2] \xrightarrow{n \rightarrow +\infty} 0$$

De plus, on a vu que  $(\sum_{i=0}^{\infty} a_i \varepsilon_{t-i})_t$  stationnaire si  $\sum_{i=0}^{\infty} a_i^2 < \infty$

$$\Rightarrow \sum_{k=0}^n \alpha^k \varepsilon_{t-k} \xrightarrow[n \rightarrow +\infty]{\mathbf{L}^2} \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k}, \text{ s\u00e9rie stationnaire.}$$

# Unicité de la solution

## Propriété

Soit  $\alpha \in ]-1, 1[$  et  $(\varepsilon_t)_{t \in \mathbb{Z}}$  un bruit blanc. Alors  $X_t = \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k}$  est p.s. l'unique solution de l'équation

$$X_{t+1} = \alpha X_t + \varepsilon_{t+1} \quad \text{pour tout } t \in \mathbb{Z}. \quad (1)$$

## Démonstration.

On note  $X_t^\infty = \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k}$  solution stationnaire de (??). Soit  $X_t$  une autre solution de (??). Alors pour tout  $t \in \mathbb{Z}$ , on a  $(X_t - X_t^\infty) = \alpha (X_{t-1} - X_{t-1}^\infty)$ . D'où pour tout  $t \in \mathbb{Z}$ ,

$(X_t - X_t^\infty) = \alpha^n (X_{t-n} - X_{t-n}^\infty)$  pour tout  $n \in \mathbb{N}$ . On en déduit que

$$\mathbf{E}[(X_t - X_t^\infty)^2] \leq 2 \alpha^{2n} (\mathbf{E}[X_{t-n}^2] + \mathbf{E}[(X_{t-n}^\infty)^2]) = 2 \alpha^{2n} (\mathbf{E}[X_0^2] + \mathbf{E}[(X_0^\infty)^2]) \xrightarrow{n \rightarrow +\infty} 0 \quad (\text{car les}$$

2 solutions sont stationnaires). Donc pour tout  $t \in \mathbb{Z}$ ,  $X_t = X_t^\infty$  p.s. □

**Remarque :** La série solution est appelée processus AR(1) causal car  $X_t \in \sigma(\varepsilon_t, \varepsilon_{t-1}, \dots)$ .

## Processus ARMA : trois exemples (2)

Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned} X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-1} (\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= \alpha^{-n} X_{t+n} - \sum_{k=1}^n (\alpha^{-1})^k \varepsilon_{t+k} \quad \text{pour } n \geq 0 \end{aligned}$$

Comme précédemment, on montre que

- $X_t = - \sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k}$  est solution et stationnaire car  $|\alpha^{-1}| < 1$ .
- $X_t = - \sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k}$  est solution unique p.s. mais non causale !



## Processus ARMA : trois exemples (3)

Soit le processus tel que  $X_t = \alpha X_{t-1} + \varepsilon_t$  pour  $t \in \mathbb{Z}$  et  $|\alpha| = 1$ .

**Exercice :** Montrer qu'il n'existe pas de solution stationnaire.

### Démonstration.

Considérons  $\alpha = 1$ . Alors par itération  $X_t = X_{t-1} + \varepsilon_t = X_{t-n} + \sum_{i=0}^{n-1} \varepsilon_{t-i}$  pour tout  $n \in \mathbb{N}^*$ . Mais si  $(X_t)$  stationnaire, alors en notant  $r_X$  l'autocovariance de  $X$ , on a

$\mathbb{E}[(X_t - X_{t-n})^2] = 2(r_X(0) - r_X(n)) = \text{var}\left(\sum_{i=0}^{n-1} \varepsilon_{t-i}\right) = n\sigma_\varepsilon^2$ . Donc

$1 - \rho_X(n) = n \frac{\sigma_\varepsilon^2}{2r_X(0)} \xrightarrow{n \rightarrow +\infty} \infty$ , ce qui est impossible pour une série stationnaire car sinon on devrait avoir  $0 \leq 1 - \rho_X(n) \leq 2$ .

Pour  $\alpha = -1$ ,  $X_t = -X_{t-1} + \varepsilon_t = X_{t-2n} + \sum_{i=0}^{2n-1} (-1)^i \varepsilon_{t-i}$ . On peut alors reprendre la preuve utilisée pour le cas  $\alpha = 1$ . □

## Processus ARMA : un autre exemple

Soit le processus tel que  $X_t = X_{t-1} + \varepsilon_t - \varepsilon_{t-1}$  pour  $t \in \mathbb{Z}$ .

**Exercice :** Montrer qu'il existe p.s. une infinité de solutions stationnaires à cette équation.

### Démonstration.

Par itération  $X_t = X_{t-2} + \varepsilon_{t-1} - \varepsilon_{t-2} + \varepsilon_t - \varepsilon_{t-1} = X_{t-2} + \varepsilon_t - \varepsilon_{t-2} = X_{t-n} + \varepsilon_t - \varepsilon_{t-n}$ .

Si on pose  $Y_t = X_t - \varepsilon_t$  alors on a  $Y_t = Y_{t-n}$  pour tout  $n \in \mathbb{N}$ .

Soit  $Z$  une variable aléatoire quelconque sur  $(\Omega, \mathcal{A})$ . Alors  $Y_t = Z$  pour tout  $t \in \mathbb{Z}$  solution p.s. de l'équation précédente.

D'où  $X_t = Z + \varepsilon_t$  est p.s. une solution de l'équation initiale et  $(X_t)$  est bien stationnaire. □

# Processus ARMA : conditions de stationnarité

## Propriété (Unicité d'un processus linéaire)

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus linéaire tel que  $X_t = \sum_{i \in \mathbb{Z}} a_i \varepsilon_{t-i} = 0$  pour tout  $t \in \mathbb{Z}$ . Alors si  $\sigma_\varepsilon^2 > 0$ , on a  $a_i = 0$  pour tout  $i \in \mathbb{Z}$ .

## Démonstration.

On sait que  $\sum_{i \in \mathbb{Z}} a_i^2 < \infty$ . On a  $\text{var}(X_t) = \sum_{i \in \mathbb{Z}} a_i^2 \text{var}(\varepsilon_{t-i})$  car  $(\varepsilon_t)$  est un bruit blanc. D'où  $\text{var}(X_t) = \sigma_\varepsilon^2 \sum_{i \in \mathbb{Z}} a_i^2$ . Comme  $X_t = 0$  pour tout  $t \in \mathbb{Z}$ , on doit avoir  $\sigma_\varepsilon^2 \sum_{i \in \mathbb{Z}} a_i^2 = 0$ , d'où  $a_i = 0$  pour tout  $i \in \mathbb{Z}$ . □

**Conséquence** : Deux processus linéaires indépendants égaux ont les mêmes coefficients (à une constante multiplicative près).

## Processus ARMA : conditions de stationnarité (2)

### Définition

On appelle opérateur retard  $B$  sur  $\mathbb{R}^{\mathbb{Z}}$  l'application linéaire telle que :

$$B : (v_t)_{t \in \mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}} \mapsto (v'_t)_{t \in \mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}} \quad \text{avec} \quad v'_t = v_{t-1} \quad \text{pour tout } t \in \mathbb{Z}.$$

### Définition

A partir de cette définition de  $B$ , on note également :

- $B^p = B \circ B \circ \dots \circ B$  pour  $p \in \mathbb{N}^*$  et  $B^0 = I_d$  ;
- $B^{-1}$  l'opérateur telle que  $B^{-1}B = I_d$  et  $B^{-p} = B^{-1}B^{-1} \dots \circ B^{-1}$  ;
- Pour  $P(x) = a_0 + a_1x + \dots + a_px^p$ ,  $P(B) = a_0I_d + a_1B + \dots + a_pB^p$  ;
- Si  $\sum_{i \in \mathbb{N}} a_i^2 < \infty$ , alors (Cauchy-Schwarz),

$$\sum_{k \in \mathbb{N}} a_k B^k : (v_t)_{t \in \mathbb{Z}} \in \ell^2(\mathbb{R}) \mapsto \left( \sum_{k \in \mathbb{N}} a_k v_{t-k} \right)_t \in \ell^2(\mathbb{R}).$$

## Processus ARMA : conditions de stationnarité (3)

Intérêt de l'opérateur  $B$  : on a

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q} \quad \forall t \in \mathbb{Z}$$

$$\iff P(B)(X) = Q(B)(\varepsilon) \text{ avec } \begin{cases} P(x) &= 1 + a_1 x + \cdots + a_p x^p \\ Q(x) &= 1 + b_1 x + \cdots + b_q x^q \end{cases}.$$

### Propriété

Si  $\alpha \in ]-1, 1[$  et  $P(x) = 1 - \alpha x$ , alors il existe  $P^{-1}$  tel que  $P_O P^{-1} = P_O^{-1} P = I_d$  avec  $P^{-1}(x) = \sum_{k=0}^{\infty} \alpha^k x^k$  et  $(\alpha^k) \in \ell^2(\mathbb{R})$ .  
De même si  $|\alpha| > 1$ , avec  $P^{-1}(B) = -\sum_{k=1}^{\infty} \alpha^{-k} B^{-k}$ .

### Propriété

Si  $P(x) = 1 + a_1 x + \cdots + a_p x^p$  et  $x_0 \neq \pm 1$  racine réelle de  $P$  telle que  $P(x) = (x - x_0) R(x)$ , alors  $-(1 - B/x_0)^{-1} P(B) = x_0 R(B)$ .

### Démonstration.

On peut écrire que  $P(B) = -(1 - B/x_0) x_0 R(B)$  et  $(1 - B/x_0)^{-1}$  existe d'après ce qui précède. □

## Processus ARMA : conditions de stationnarité (4)

### Propriété

Si  $X$  vérifie  $P(B)(X) = Q(B)(\varepsilon)$  équation ARMA( $p, q$ ), si  $x_0 \neq \pm 1$  racine commune de  $P$  et  $Q$ ,  $P(x) = (x - x_0)R_P(x)$  et  $Q(x) = (x - x_0)R_Q(x)$ , alors  $X$  vérifie l'équation ARMA( $p - 1, q - 1$ ),  $R_P(B)(X) = R_Q(B)(\varepsilon)$ .

### Démonstration.

Immédiat en composant par  $(1 - B/x_0)^{-1}$ . □

### Propriété

Si  $P(x) = 1 + a_1x + \dots + a_px^p$  est scindé dans  $\mathbb{R}$  avec des racines  $\neq \pm 1$ , sans racine commune avec  $Q$ , alors  $P(B)(X) = Q(B)(\varepsilon)$  définit un processus ARMA( $p, q$ ) stationnaire.

## Processus ARMA : conditions de stationnarité (5)

### Proposition

Si  $X$  vérifie  $P(B)(X) = Q(B)(\varepsilon)$ , les racines de  $P(x) = 1 + a_1x + \dots + a_px^p$  sont en dehors du disque unité, soit  $P(z) \neq 0$  pour  $|z| \leq 1$ , et sans racine commune avec  $Q$  alors  $(X_t)$  est un processus ARMA( $p, q$ ) et s'écrit p.s. comme un unique processus linéaire causal : pour  $t \in \mathbb{Z}$

$$X_t = \sum_{i=0}^{\infty} u_i \varepsilon_{t-i} \quad \text{avec} \quad |u_n| \leq C \rho^n \quad \text{pour } n \in \mathbb{N}, \quad 0 < C \text{ et } 0 < \rho < 1.$$

### Démonstration.

Supposons que  $X_t = \sum_{i=0}^{\infty} u_i \varepsilon_{t-i}$  pour tout  $t \in \mathbb{Z}$  avec  $\sum_{i=0}^{\infty} u_i^2 < \infty$ . Alors  $X_{t-k} = \sum_{i=0}^{\infty} u_i \varepsilon_{t-k-i} = \sum_{i=k}^{\infty} u_{i-k} \varepsilon_{t-i}$  pour  $k \geq 0$ , d'où

$$\begin{aligned} X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} &= \sum_{i=p}^{\infty} (u_i + a_1 u_{i-1} + \dots + a_p u_{i-p}) \varepsilon_{t-i} \\ &+ (u_{p-1} + a_1 u_{p-2} + \dots + a_{p-1} u_0) \varepsilon_{t-(p-1)} + \dots + (u_1 + a_1 u_0) \varepsilon_{t-1} + u_0 \varepsilon_t \end{aligned}$$



## Processus ARMA : conditions de stationnarité (6)

### Suite de preuve.

si  $X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$  pour tout  $t \in \mathbb{Z}$ , d'après l'unicité d'écriture d'un processus linéaire, on a

$$\begin{cases} u_0 & = & 1 \\ u_1 + a_1 u_0 & = & b_1 \\ \vdots & \vdots & \vdots \\ u_i + a_1 u_{i-1} + \dots + a_p u_{i-p} & = & 0 \quad \text{pour tout } i \geq \max(p, q) \end{cases}$$

L'équation  $u_i + a_1 u_{i-1} + \dots + a_p u_{i-p} = 0$  pour tout  $i \geq \max(p, q)$  définit une suite récurrente linéaire d'ordre  $p$ . Son polynôme caractéristique est

$\chi(X) = X^p + a_1 X^{p-1} + \dots + a_p = X^p P(1/X)$ . Comme  $P$  a ses racines en dehors du disque unité, toutes les racines de  $\chi$  sont de modules  $< 1$ . Or la solution générale de la suite linéaire s'écrit  $u_n = R_1(n)x_1^n + \dots + R_m(n)x_m^n$ , les  $R_i$  étant des polynômes et les  $x_i$  les racines de  $\chi$ . On en déduit qu'il existe  $C > 0$  et  $0 < \rho < 1$  tel que  $u_n \leq C \rho^n$ . Par conséquent,  $\sum_{i=0}^{\infty} u_i^2 < \infty$ . Enfin, le système étant triangulaire, il admet une unique solution  $(u_n)$ .

Pour montrer que  $X_t = X_t^\infty = \sum_{i=0}^{\infty} u_i \varepsilon_{t-i}$  est p.s. l'unique solution stationnaire, on considère  $(X_t)$  une autre solution stationnaire et alors :

$$(X_t - X_t^\infty) + a_1(X_{t-1} - X_{t-1}^\infty) + \dots + a_p(X_{t-p} - X_{t-p}^\infty) = 0 \quad \text{pour tout } t \in \mathbb{Z}.$$

Avec ce qui précède  $|X_t - X_t^\infty| \leq C((X_{t-n-k} - X_{t-n-k}^\infty)_{1 \leq k \leq p}) \rho^n$  pour tout  $n \in \mathbb{N}$  d'où  $X_t = X_t^\infty$  pour tout  $t \in \mathbb{Z}$ . □



## Processus ARMA : conditions de stationnarité (7)

**Remarque** : On peut montrer plus généralement que si  $P$  a ses racines en dehors du cercle unité, alors il existe p.s. une unique solution sous forme de processus linéaire. Mais pour que cette solution soit causale (essentiel dans ce cours), il faut que  $P(z) \neq 0$  pour  $|z| \leq 1$ .

**Remarque** : Les conditions de stationnarité ne portent que sur les  $a_i$  (partie AR) et non sur les  $b_i$  (partie MA) quand il n'y a pas de racines communes entre  $P$  et  $Q$ .

# Propriétés des processus ARMA (1)

## Propriété

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \omega_0 + \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

Alors :

- 1 Si  $\omega_0 = 0$ , on a  $\mathbf{E}(X_t) = 0$  ;
- 2 Si  $\omega_0 \neq 0$ , alors  $(X'_t)$  tel que  $X'_t = X_t - \frac{\omega_0}{1+a_1+\cdots+a_p}$  vérifie

$$X'_t + a_1 X'_{t-1} + \cdots + a_p X'_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

$$\implies \mathbf{E}(X_t) = \frac{\omega_0}{1 + a_1 + \cdots + a_p} : \text{ARMA}(p, q) \text{ décentré}$$

## Propriété

Un processus ARMA( $p, q$ ) causal de bruit blanc  $(\varepsilon_t)$  gaussien est gaussien.

# Propriétés des processus ARMA (2)

## Proposition

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

Alors si  $r_X$  est l'autocovariance de  $X$ ,  $r_X$  vérifie l'équation de Yule-Walker :

$$\begin{aligned} r_X(k) + a_1 r_X(k-1) + \cdots + a_p r_X(k-p) &= 0 \quad \text{pour tout } k \geq q+1 \\ \implies |r_X(n)| &\leq C \rho^n \quad \text{avec } 0 < \rho < 1 \text{ pour tout } n \in \mathbb{N} \end{aligned}$$

**Cas particulier :** Si  $(X_t)$  est un processus MA( $q$ ),

- 1  $u_j = b_j$  pour  $j = 1, \dots, q$ , et  $\alpha_j = 0$  si  $j \geq q+1$ ;
- 2  $r_X(k) = 0$  pour  $k \geq q+1$  et avec  $b_0 = 1$ , pour  $0 \leq k \leq q$ ,

$$r_X(k) = \text{cov}\left(\sum_{j=0}^q b_j \varepsilon_{-j}, \sum_{\ell=0}^q b_\ell \varepsilon_{k-\ell}\right) = \sum_{j=0}^q \sum_{\ell=0}^q b_j b_\ell \mathbf{E}[\varepsilon_{-j} \varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

# Propriétés des processus ARMA (3)

## Preuve de l'équation de Yule-Walker.

On a pour tout  $k \in \mathbb{N}$ ,

$$\begin{aligned} X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} &= \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q} \\ \Rightarrow X_{t-k} X_t + a_1 X_{t-k} X_{t-1} + \dots + a_p X_{t-k} X_{t-p} &= X_{t-k} (\varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}) \\ \mathbf{E}[X_{t-k} X_t] + a_1 \mathbf{E}[X_{t-k} X_{t-1}] + \dots + a_p \mathbf{E}[X_{t-k} X_{t-p}] &= \mathbf{E}[X_{t-k} (\varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q})] \\ \Rightarrow r_X(k) + a_1 r_X(k-1) + \dots + a_p r_X(k-p) &= 0 \end{aligned}$$

dès que  $k \geq q+1$  car  $(X_t)$  est causal donc  $X_{t-k}$  indépendant de  $\varepsilon_t$ , de  $\varepsilon_{t-1}$  et de  $\varepsilon_{t-q}$ . Enfin, comme les racines de  $P$  sont en dehors du disque unité, on a bien  $|r_X(n)| \leq C \rho^n$ .  $\square$

**Exercice :**  $(X_t)$  processus ARMA(1, 1) causal  $X_t = \alpha X_{t-1} + \varepsilon_t + \beta \varepsilon_{t-1}$  avec pour  $t \in \mathbb{Z}$ . Déterminer  $r_X(n)$  pour  $n \in \mathbb{N}$  en fonction de  $\sigma_\varepsilon^2$ ,  $\alpha$  et  $\beta$ .

## Démonstration.

D'après Yule-Walker,  $r_X(k) = \alpha r_X(k-1)$  pour  $k \geq 2$ . D'où pour  $k \geq 1$ ,  $r_X(k) = \alpha^{k-1} r_X(1)$ . Comme  $X_t = \alpha X_{t-1} + \varepsilon_t + \beta \varepsilon_{t-1}$ ,  $\text{var}(X_t - \alpha X_{t-1}) = \text{var}(\varepsilon_t + \beta \varepsilon_{t-1})$  soit  $r_X(0)(1 + \alpha^2) - 2\alpha r_X(1) = \sigma_\varepsilon^2(1 + \beta^2)$ . Enfin,  $\text{cov}(X_t, \varepsilon_t) = \sigma_\varepsilon^2$ , d'où  $r_X(1) = \alpha r_X(0) + \beta \sigma_\varepsilon^2$ , d'où  $r_X(0) = \sigma_\varepsilon^2 \frac{1+2\alpha\beta+\beta^2}{1-\alpha^2}$  et  $r_X(1) = \sigma_\varepsilon^2 \frac{(1+\alpha\beta)(\alpha+\beta)}{1-\alpha^2}$ .  $\square$

# Processus ARMA : corrélogramme

## Proposition

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus ARMA( $p, q$ ) stationnaire causal. Soit  $\rho(\cdot)$  son autocorrélation et  $\widehat{\rho}(\cdot)$  son autocorrélation empirique calculée sur  $(X_1, \dots, X_n)$ . Alors :

$$\widehat{\rho}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(k) \quad \text{pour tout } k \in \mathbb{N}$$

## Démonstration.

Voir le chapitre 3. □

**Conséquence** : l'ACF converge vers les autocorrélations théoriques

# Plan du cours

# Processus GARCH

Définition (Engle, 1982 et Bollerssev, 1986)

Un processus GARCH( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t = \varepsilon_t \sigma_t \quad \text{et} \quad \sigma_t^2 = a_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2 + b_1 \sigma_{t-1}^2 + \dots + b_q \sigma_{t-q}^2$$

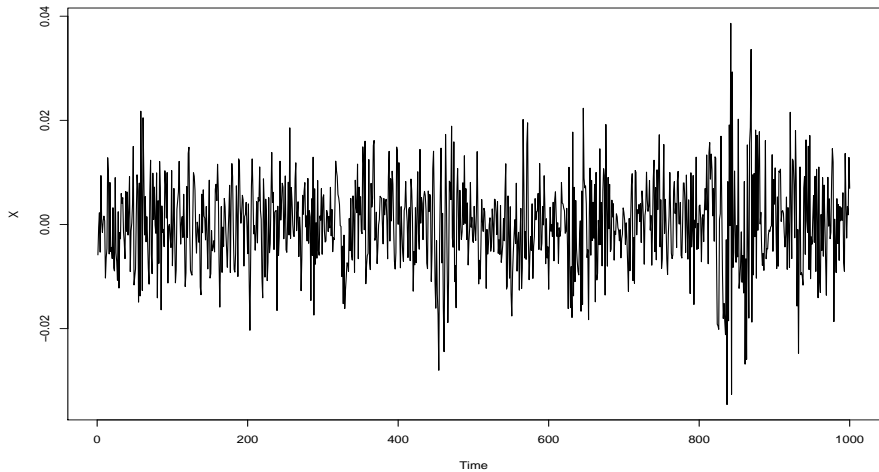
où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $a_0 > 0, (a_1, \dots, a_p, b_1, \dots, b_q) \in [0, \infty)^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

Cas particulier :

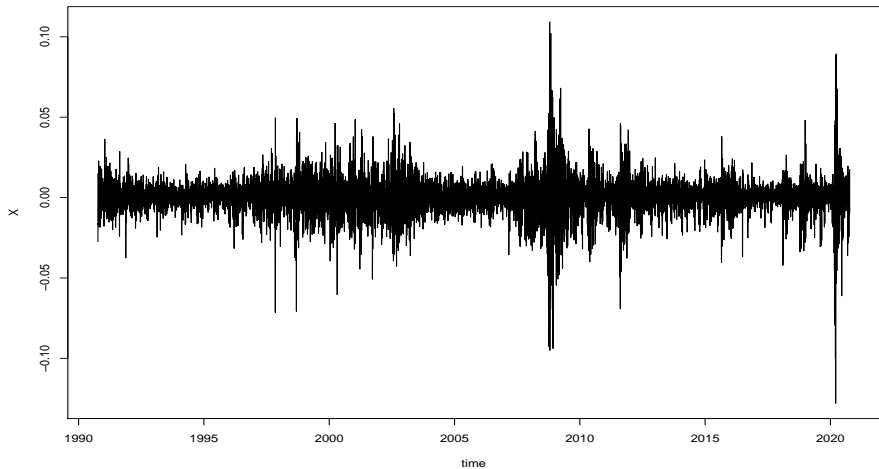
- Si  $q = 0$ , un GARCH( $p, q$ ) est un ARCH( $p$ ) (Conditionnaly Heteroskedastic Auto-Regressive) ;

# Trajectoire d'un processus GARCH(1,1)

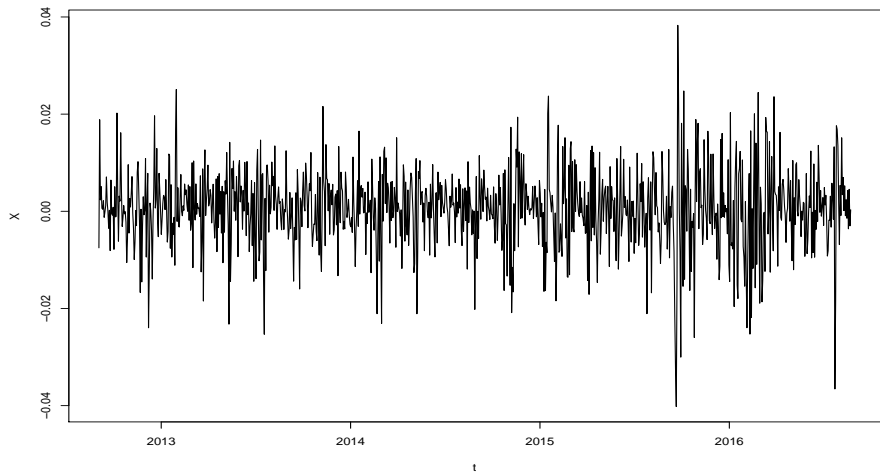




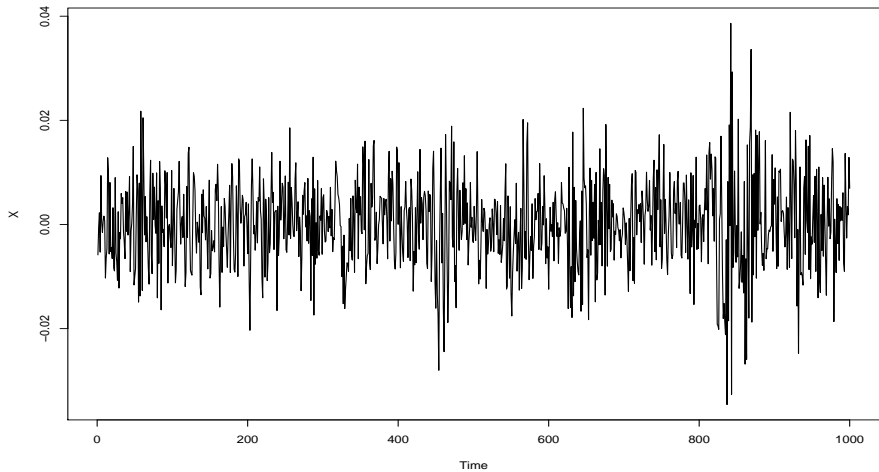
# Trajectoire Log-rendement SP500 d'octobre 1990 à octobre 2020



# Trajectoire Log-rendement SP500 09/2012 -> 09/2016



# Trajectoire d'un processus GARCH(1,1)



# Conditions de stationnarité d'un processus GARCH(1,1)

## Propriété

Un processus GARCH(1,1) est stationnaire si et seulement si  $\mathbb{E}[\log(a_1 \varepsilon_0^2 + b_1)] < 0$  et il vérifie alors p.s.

$$X_t = \sqrt{a_0} \varepsilon_t \left( 1 + \sum_{j=1}^{\infty} (a_1 \varepsilon_{t-1}^2 + b_1)(a_1 \varepsilon_{t-2}^2 + b_1) \times \cdots \times (a_1 \varepsilon_{t-j}^2 + b_1) \right)^{1/2}.$$

## Remarques :

- Si  $a_0 = 0$  alors  $X_t = 0$  p.s. pour tout  $t \in \mathbb{Z}$ .
- Si  $a_1 + b_1 < 1$ , en utilisant l'inégalité de Jensen,  
 $\mathbb{E}[\log(a_1 \varepsilon_0^2 + b_1)] \leq \log(\mathbb{E}[a_1 \varepsilon_0^2 + b_1]) \leq \log(a_1 + b_1) < 0$ .

# Preuve de la stationnarité d'un processus GARCH(1,1)

## Démonstration.

Soit  $\alpha(x) = a_1 x^2 + b_1$ . Alors  $\sigma_t^2 = a_0 + \alpha(\varepsilon_{t-1})\sigma_{t-1}^2$  pour tout  $t \in \mathbb{Z}$ . En itérant, pour  $m \geq 1$  :

$$\sigma_t^2 = a_0 \left( 1 + \sum_{j=1}^{m-1} \alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-j}) \right) + \alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-m})\sigma_{t-m}^2.$$

Supposons que  $\mathbf{E}[\log(a_1\varepsilon_0^2 + b_1)] = \mathbf{E}[\log(\alpha(\varepsilon_0))] < 0$ . D'après la loi forte des grands nombres :

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \log(\alpha(\varepsilon_{t-i})) &\xrightarrow[n \rightarrow \infty]{p.s.} \mathbf{E}[\log(\alpha(\varepsilon_0))] \\ \implies \left( \alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-n}) \right)^{1/n} &\xrightarrow[n \rightarrow \infty]{p.s.} \exp\left(\mathbf{E}[\log(\alpha(\varepsilon_0))]\right) < 1 \end{aligned}$$

Critère de Cauchy pour séries numériques : si  $|u_n|^{1/n} \xrightarrow[n \rightarrow +\infty]{} \ell < 1$  alors  $\sum |u_n|$  converge. D'où :

$$a_0 \left( 1 + \sum_{j=1}^n \alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-j}) \right) \xrightarrow[n \rightarrow \infty]{p.s.} s^2(t) = a_0 \left( 1 + \sum_{j=1}^{\infty} \alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-j}) \right)$$

La somme existe p.s. donc  $\alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-n}) \xrightarrow[n \rightarrow \infty]{p.s.} 0$ . Ainsi, pour obtenir une

solution stationnaire  $(X_t)$ , ce qui induit que  $(\sigma_t^2)$  est stationnaire, alors

$$\alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-n})\sigma_{t-n}^2 \xrightarrow[n \rightarrow \infty]{p.s.} 0. \quad \square$$

# Preuve de la stationnarité d'un processus GARCH(1,1)

## Démonstration.

On déduit :

$$X_t = \sqrt{a_0} \varepsilon_t \left( 1 + \sum_{j=1}^{\infty} \alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-j}) \right)^{1/2} \quad \text{presque sûrement.} \quad (2)$$

On peut montrer l'unicité de cette solution : si  $X_t = \varepsilon_t \sigma'(t)$  avec  $\sigma_t'^2 = a_0 + \alpha(\varepsilon_{t-1})\sigma_{t-1}'^2$  stationnaire, en utilisant l'itération précédente, on a

$$\sigma_t'^2 - \sigma_{t-n}'^2 = \alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-n}) (\sigma_{t-n}'^2 - \sigma_{t-n}'^2) \xrightarrow[n \rightarrow \infty]{p.s.} 0.$$

Si  $\mathbf{E}[\log(a_1 \varepsilon_0^2 + b_1)] = \mathbf{E}[\log(\alpha(\varepsilon_0))] > 0$ , où

$\sigma_t'^2 \geq a_0 \left( 1 + \sum_{j=1}^n \alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-j}) \right)$  pour tout  $n \in \mathbf{N}$ , en utilisant également la règle de Cauchy, on montre que  $\sum_{j=1}^n \alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-j}) \xrightarrow[n \rightarrow +\infty]{} +\infty$  donc il n'existe pas de solution stationnaire.

Si  $\mathbf{E}[\log(a_1 \varepsilon_0^2 + b_1)] = \mathbf{E}[\log(\alpha(\varepsilon_0))] = 0$ , on a

$\sigma_t'^2 \geq a_0 \left( 1 + \sum_{j=1}^n \alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-j}) \right)$  pour tout  $n \in \mathbf{N}$ . Pour qu'une solution stationnaire existe, il faut que la série converge p.s., donc que

$\alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-j}) \xrightarrow[n \rightarrow +\infty]{} 0$ , soit  $\sum_{i=1}^n \log(\alpha(\varepsilon_{t-i})) \xrightarrow[n \rightarrow \infty]{p.s.} -\infty$ . Mais

$(\log(\alpha(\varepsilon_{t-i})))_{i \geq 1}$  est une suite de v.a.i.i.d. centrées, donc  $\sum_{i=1}^n \log(\alpha(\varepsilon_{t-i}))$  est une marche aléatoire centrée qui ne peut converger p.s. vers  $-\infty$  ( $\mathbf{P}(\sum_{i=1}^n \log(\alpha(\varepsilon_{t-i})) \geq 0) \xrightarrow[n \rightarrow +\infty]{} 1/2$

par le TLC)  $\implies$  il n'existe pas de solution stationnaire dans ce cas. □

# Conditions de stationnarité d'ordre 2 d'un processus GARCH(1, 1)

## Propriété

Processus GARCH(1, 1) stationnaire d'ordre 2 si et seulement si  $a_1 + b_1 < 1$ .

## Démonstration.

$\Rightarrow$  Si  $X$  a un moment d'ordre 2, stationnaire d'ordre 2,  
 $\mathbf{E}[X_t^2] = \mathbf{E}[\sigma_t^2] = a_0 + a_1 \mathbf{E}[X_{t-1}^2] + b_1 \mathbf{E}[\sigma_{t-1}^2]$ , d'où l'équation  
 $\mathbf{E}[X_0^2] = a_0 + a_1 \mathbf{E}[X_0^2] + b_1 \mathbf{E}[X_0^2]$ , soit :

$$\mathbf{E}[X_0^2] = \frac{a_0}{1 - a_1 - b_1} \quad \Rightarrow \quad a_1 + b_1 < 1.$$

$\Leftarrow$  Si  $a_1 + b_1 < 1$  et  $\mathbf{E}(\varepsilon_0^2) = 1$  alors  $\mathbf{E}[\log(a_1\varepsilon_0^2 + b_1)] < 0$  d'où  $X$  stationnaire et s'écrit sous la forme (??). D'où  $X$  stationnaire d'ordre 2 si  $\mathbf{E}[X_0^2] < \infty$ . Avec (??) et l'indépendance des  $(\varepsilon_t)$

$$\begin{aligned} \mathbf{E}[X_t^2] &= a_0 \mathbf{E}[\varepsilon_t^2] \mathbf{E}\left[1 + \sum_{j=1}^{\infty} \alpha(\varepsilon_{t-1})\alpha(\varepsilon_{t-2}) \times \cdots \times \alpha(\varepsilon_{t-j})\right] \\ &= a_0 \left(1 + \sum_{j=1}^{\infty} \mathbf{E}[\alpha(\varepsilon_{t-1})] \mathbf{E}[\alpha(\varepsilon_{t-2})] \times \cdots \times \mathbf{E}[\alpha(\varepsilon_{t-j})]\right) \\ &= a_0 \left(1 + \sum_{j=1}^{\infty} (a_1 + b_1)^j\right) = \frac{a_0}{1 - a_1 - b_1}. \end{aligned}$$

# Conditions de stationnarité d'un processus GARCH( $p, q$ )

**Remarque** : CNS de stationnarité pour un GARCH( $p, q$ ) trop complexe !

Stationnarité d'ordre 2 : généralisation aux processus GARCH( $p, q$ ) :

## Propriété

Un processus GARCH( $p, q$ ) est stationnaire d'ordre 2 si et seulement si

$$\sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1.$$

Preuve condition suffisante très récente (2000). Condition nécessaire :

## Démonstration.

$\implies$  Si  $X$  a un moment d'ordre 2 et est stationnaire d'ordre 2 alors

$\mathbf{E}[X_t^2] = \mathbf{E}[\sigma_t^2] = a_0 + a_1 \mathbf{E}[X_{t-1}^2] + \dots + a_p \mathbf{E}[X_{t-p}^2] + b_1 \mathbf{E}[\sigma_{t-1}^2] + \dots + b_q \mathbf{E}[\sigma_{t-q}^2]$ , d'où l'équation  $\mathbf{E}[X_0^2] = a_0 + a_1 \mathbf{E}[X_0^2] + \dots + a_p \mathbf{E}[X_0^2] + b_1 \mathbf{E}[X_0^2] + \dots + b_q \mathbf{E}[X_0^2]$ , soit :

$$\mathbf{E}[X_0^2] = \frac{a_0}{1 - a_1 - \dots - a_p - b_1 - \dots - b_q} \implies \sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1.$$



# Propriétés d'un processus GARCH

De ceci, on peut déduire les propriétés suivantes :

## Propriété

Soit  $X = (X_k)_{k \in \mathbb{Z}}$  un processus GARCH( $p, q$ ) stationnaire d'ordre 2. Alors

- 1  $\mathbb{E}[X_t] = 0$ ,  $\text{var}(X_t) = \frac{a_0}{1 - \sum_{i=1}^p a_i - \sum_{j=1}^q b_j}$  et  $\text{cov}(X_0, X_k) = 0$  si  $k \neq 0$ .
- 2  $\mathbb{E}[X_t | X_{t-1}, X_{t-2}, \dots] = 0$  et  $\text{var}(X_t | X_{t-1}, X_{t-2}, \dots) = \sigma_t^2 \neq C^{te}$  :  
*Conditionnellement Hétéroscédastique*

## Démonstration.

- 1 Si  $X$  a un moment d'ordre 2 et est stationnaire d'ordre 2 alors  $\mathbb{E}[X_t] = \mathbb{E}[\varepsilon_t \sigma_t]$ . Mais  $\sigma_t$  fonction p.s. de  $(\varepsilon_{t-k})_{k \geq 1}$ , donc  $\varepsilon_t$  et  $\sigma_t$  indépendants, d'où  $\mathbb{E}[X_t] = \mathbb{E}[\varepsilon_t] \mathbb{E}[\sigma_t] = 0$ .  
Pour  $k \geq 1$ ,  $\text{cov}(X_0, X_k) = \mathbb{E}[\varepsilon_k (\varepsilon_0 \sigma_0 \sigma_k)] = \mathbb{E}[\varepsilon_k] \mathbb{E}[(\varepsilon_0 \sigma_0 \sigma_k)] = 0$  toujours grâce à la forme causale de  $\sigma_t$ .
- 2  $\mathbb{E}[X_t | X_{t-1}, X_{t-2}, \dots] = \mathbb{E}[\varepsilon_t \sigma_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots] = \sigma_t \mathbb{E}[\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots] = 0$ .  
 $\text{var}(X_t | X_{t-1}, X_{t-2}, \dots) = \mathbb{E}[X_t^2 | X_{t-1}, X_{t-2}, \dots] = \sigma_t^2 \mathbb{E}[\varepsilon_t^2] = \sigma_t^2$ .



# Propriétés d'un processus GARCH

## Propriété

Soit  $X = (X_k)_{k \in \mathbb{Z}}$  un processus GARCH( $p, q$ ) stationnaire d'ordre 2 tel que  $\mathbb{E}(\varepsilon^4) < \infty$ . Alors  $Y = (X_k^2)_{k \in \mathbb{Z}}$  est un processus ARMA( $\max(p, q), q$ ) faible non centré.

## Démonstration.

On considère  $v_t = X_t^2 - \sigma_t^2 = (\varepsilon_t^2 - 1)\sigma_t^2$ . La série  $(v_t)$  est un bruit blanc faible car  $(v_t)$  est stationnaire,  $\mathbb{E}(v_t) = 0$  et pour  $t > 0$ ,  $\text{cov}(v_0, v_t) = \text{cov}((\varepsilon_0^2 - 1)\sigma_0^2, (\varepsilon_t^2 - 1)\sigma_t^2) = 0$  car  $(\varepsilon_t^2 - 1)$  est indépendant des 3 autres termes et d'espérance nulle. De plus, on montre que

$$X_t^2 = v_t + (a_0 + \sum_{i=1}^p a_i X_{t-i}^2) + \sum_{j=1}^q b_j (X_{t-j}^2 - v_{t-j}) =$$
$$a_0 + \sum_{i=1}^p a_i X_{t-i}^2 + \sum_{j=1}^q b_j X_{t-j}^2 + v_t - \sum_{j=1}^q b_j v_{t-j}, \text{ d'où } (X_t^2) \text{ ARMA faible non centré. } \square$$

**Conséquences :** • Processus GARCH : bruit blanc faible mais pas fort

• Un processus GARCH ne peut jamais être un processus gaussien

# Plan du cours

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus ARMA( $p, q$ )-GARCH( $p', q'$ ) si

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

où  $\varepsilon_t = \xi_t \sigma_t$  et  $\sigma_t^2 = \omega_0 + c_1 \varepsilon_{t-1}^2 + \dots + c_{p'} \varepsilon_{t-p'}^2 + d_1 \sigma_{t-1}^2 + \dots + d_{q'} \sigma_{t-q'}^2$

avec  $(\xi_t)$  un bruit blanc de variance 1,  $\omega_0 > 0$ ,  $a_p, b_q, c_{p'}, d_{q'}$  non nuls.

**Remarque :** • On doit avoir  $P(z) \neq 0$  pour  $|z| \leq 1$  où

$$P(x) = 1 + a_1 x + \dots + a_p x^p \text{ et } \sum_{i=1}^{p'} c_i + \sum_{j=1}^{q'} d_j < 1$$

• ARMA-GARCH=ARMA avec bruit GARCH : même autocorrélation qu'un ARMA

• ARMA-GARCH : espérance et variance conditionnelles non constantes

# Processus ARMA-GARCH (preuves)

- ARMA-GARCH=ARMA avec bruit GARCH : même autocorrélation qu'un ARMA

## Démonstration.

Que  $(\varepsilon_t)$  soit un bruit blanc fort ou faible, ou même un processus stationnaire, on a  $X_t = \sum_{i=0}^{\infty} u_i \varepsilon_{t-i}$ . D'où  $r_x(k) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} u_i u_j \mathbf{E}[\varepsilon_{t-i} \varepsilon_{t-k-j}] = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} u_i u_j r_\varepsilon(k+i-j)$ .  
Donc  $r_x$  est la même autocovariance dès que  $r_\varepsilon(k) = 0$  pour  $k \geq 1$ . □

- ARMA-GARCH : espérance et variance conditionnelles non constantes

## Démonstration.

On a  $X_t = \varepsilon_t + \sum_{i=1}^{\infty} u_i \varepsilon_{t-i}$ . Donc  $\sigma(X_{t-1}, X_{t-2}, \dots) = \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots)$ . Par conséquent,  $\mathbf{E}[X_t | X_{t-1}, X_{t-2}, \dots] = \mathbf{E}[\varepsilon_t + \sum_{i=1}^{\infty} u_i \varepsilon_{t-i} | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots] = \sum_{i=1}^{\infty} u_i \varepsilon_{t-i} = X_t - \varepsilon_t = a_1 X_{t-1} + \dots + a_p X_{t-p} - b_1 \varepsilon_{t-1} - \dots - b_q \varepsilon_{t-q}$  : espérance conditionnelle non constante.  
 $\text{var}(X_t | X_{t-1}, X_{t-2}, \dots) = \mathbf{E}[(X_t - \mathbf{E}[X_t | X_{t-1}, X_{t-2}, \dots])]^2 | X_{t-1}, X_{t-2}, \dots] = \mathbf{E}[\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots] = \sigma_t^2$  : variance conditionnelle non constante. □

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus FARIMA( $p, d, q$ ) où  $-0.5 < d < 0.5$  et

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

$$\text{où } \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \left( \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \right) \xi_{t-j}$$

avec  $(\xi_t)$  bruit blanc de variance  $\sigma_\xi^2$ ,  $a_p, b_q$  non nuls,  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ .

**Commentaire :** • Avec  $P(x) = 1 + a_1 x + \dots + a_p x^p$ ,  $P(z) \neq 0$  pour  $|z| \leq 1$

• On a pour tout  $|x| < 1$ ,

$$(1-x)^{-d} = 1 + \sum_{k=1}^{\infty} \frac{d(d+1) \times \dots \times (d+k-1)}{k!} x^k = \sum_{k=0}^{\infty} \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)} x^k$$

# Processus FARIMA (2)

## Propriété

$(\varepsilon_t)$  processus FARIMA(0, d, 0)  $\implies (\varepsilon_t)$  processus linéaire stationnaire. Si  $0 < d < \frac{1}{2}$ ,  $r_\varepsilon(k) \underset{k \rightarrow \infty}{\sim} C|k|^{2d-1}$  ( $\varepsilon_t$ ) processus longue mémoire  $\sum_k |r_\varepsilon(k)| = \infty$

## Démonstration.

On a  $\varepsilon_t = \sum_{i=0}^{\infty} u_i \xi_{t-i}$  avec  $u_n = \frac{\Gamma(n+d)}{\Gamma(d)\Gamma(n+1)}$ . Par la formule de Stirling, pour  $z \rightarrow \infty$ ,

$\Gamma(z) \sim \sqrt{2\pi} z^{z-1/2} e^{-z}$ . D'où pour  $n \rightarrow \infty$ ,  $u_n \sim \frac{e^{1-d}}{\Gamma(d)} \frac{(n+d)^{n+d-1/2}}{(n+1)^{n+1/2}}$ . Après développements

limités, on trouve  $u_n \sim \frac{1}{\Gamma(d)} n^{d-1}$ . Ainsi  $u_n^2 \sim \frac{1}{\Gamma^2(d)} n^{2d-2}$  et comme  $\sum_n n^{2d-2} < \infty$  car  $-1/2 < d < 1/2$ , on en déduit que  $(\varepsilon_t)$  processus linéaire stationnaire.

Et  $r_\varepsilon(k) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} u_i u_j \mathbb{E}[\xi_{-i} \xi_{k-j}] = \sigma_\xi^2 \sum_{i=0}^{\infty} u_i u_{i+k}$ . Si  $i \rightarrow \infty$ ,  $u_i u_{i+k} \sim \frac{1}{\Gamma^2(d)} (i(i+k))^{d-1}$ , et

pour tout  $i$ ,  $u_i u_{i+k} \leq C k^{d-1}$ . Or  $\sum_{i=0}^{\infty} u_i u_{i+k} = \sum_{i=0}^{\lfloor k^{d/2} \rfloor} u_i u_{i+k} + \sum_{i=\lfloor k^{d/2} \rfloor + 1}^{\infty} u_i u_{i+k}$ . On a

$\sum_{i=0}^{\lfloor k^{d/2} \rfloor} u_i u_{i+k} \leq C \lfloor k^{d/2} \rfloor k^{d-1} \leq C k^{3d/2-1}$ . Par équivalence reste/intégrale, si  $k \rightarrow \infty$ ,  $\sum_{i=\lfloor k^{d/2} \rfloor}^{\infty} u_i u_{i+k} \sim \int_{k^{d/2}}^{\infty} (x(x+k))^{d-1} dx$ . Or  $(x(x+k))^{d-1} = (k/2)^{2d-2} ((1+2x/k)^2 - 1)^{d-1}$ .

Avec  $y = 1 + 2x/k$ , on obtient  $\int_{k^{d/2}}^{\infty} (x(x+k))^{d-1} dx = (k/2)^{2d-1} \int_{1+2k^{d/2-1}}^{\infty} (y^2 - 1)^{d-1} dy$ . Au

final,  $r_\varepsilon(k) \simeq \frac{\sigma_\xi^2}{\Gamma^2(d)} \left( \int_1^{\infty} (y^2 - 1)^{d-1} dy \right) (k/2)^{2d-1}$ . □

# Processus FARIMA (3)

## Propriété

$(X_t)$  processus FARIMA( $p, d, q$ )  $\implies (X_t)$  processus linéaire stationnaire. Si  $0 < d < \frac{1}{2}$ ,  $r_X(k) \underset{k \rightarrow \infty}{\sim} C |k|^{2d-1}$  ( $X_t$ ) processus longue mémoire  $\sum_k |r_X(k)| = \infty$

## Démonstration.

On a  $X_t = \sum_{i=0}^{\infty} v_i \varepsilon_{t-i}$  avec  $|v_n| \leq C \rho^n$ . Comme  $(\varepsilon_t)$  processus linéaire, par composition des filtres linéaires,  $(X_t)$  processus linéaire. Et comme  $\sum_i |v_i| < \infty$  et  $(\varepsilon_t)$  stationnaire et stationnaire d'ordre 2, alors  $(X_t)$  stationnaire et stationnaire d'ordre 2.

De plus,  $r_X(k) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} v_i v_j r_{\varepsilon}(k+i-j) = \sum_{|\ell| \leq \sqrt{k}} \sum_{j=-\ell}^{\infty} v_{j+\ell} v_j r_{\varepsilon}(k+\ell) + \sum_{|\ell| > \sqrt{k}} \sum_{j=-\ell}^{\infty} v_{j+\ell} v_j r_{\varepsilon}(k+\ell)$ . Soit  $w_{\ell} = \sum_{j=-\ell}^{\infty} v_{j+\ell} v_j$ . On a  $w_{\ell} \leq C \rho^{|\ell|}$ . On montre ainsi que

$\sum_{|\ell| \leq \sqrt{k}} w_{\ell} r_{\varepsilon}(k+\ell) \simeq C \sum_{|\ell| \leq \sqrt{k}} w_{\ell} (k+\ell)^{2d-1} \simeq C k^{2d-1} \sum_{\ell=-\infty}^{\infty} w_{\ell}$ . Par ailleurs,

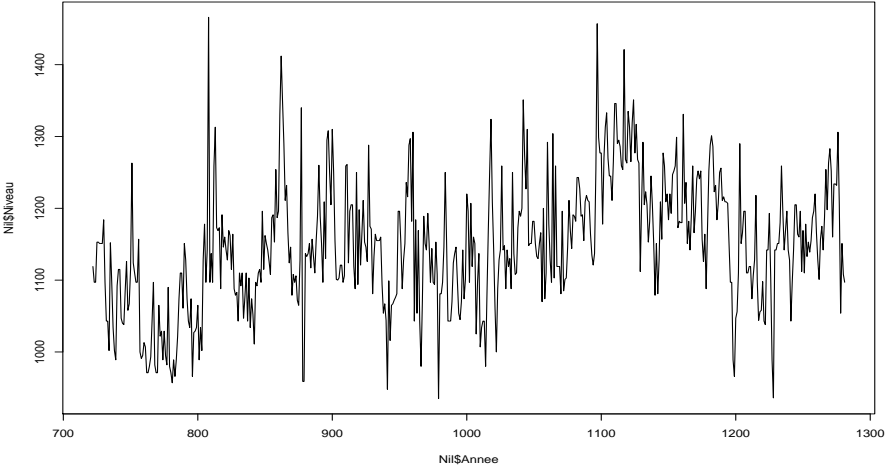
$\left| \sum_{|\ell| > \sqrt{k}} w_{\ell} r_{\varepsilon}(k+\ell) \right| \leq C \left| \sum_{|\ell| > \sqrt{k}} w_{\ell} \right| \leq C' \rho^{\sqrt{k}} 2 \sum_{\ell=0}^{\infty} \rho^{\ell} \leq C'' \rho^{\sqrt{k}}$ . D'où

$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} v_i v_j r_{\varepsilon}(k+i-j) \simeq C k^{2d-1} \sum_{\ell=-\infty}^{\infty} w_{\ell}$  pour  $k \rightarrow \infty$ . □

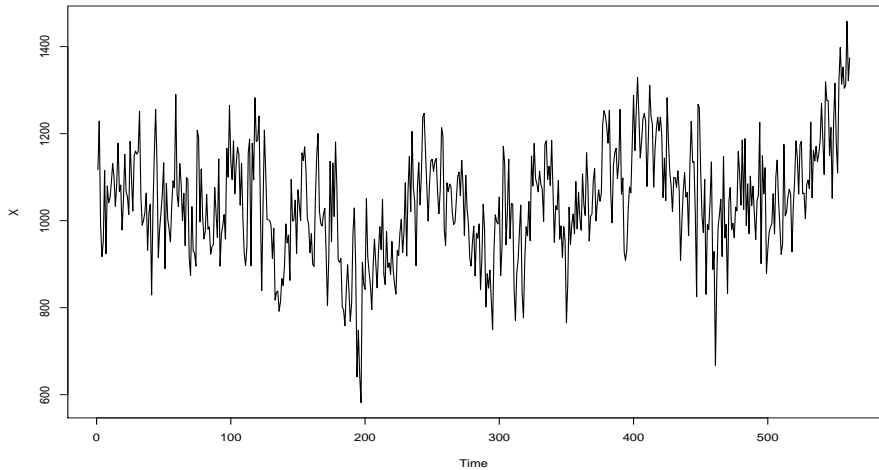
**Remarque :** On a  $\hat{\rho}(\cdot) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(\cdot)$  : ACF estime  $\rho$



# Niveau étiage du Nil



# Trajectoire d'un FARIMA(0, 0.4, 0)



## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus APARCH( $\delta, p, q$ ) si  $X_t = \varepsilon_t \sigma_t$  et

$$\sigma_t^\delta = \omega_0 + a_1(|X_{t-1}| - \gamma_1 X_{t-1})^\delta + \dots + a_p(|X_{t-p}| - \gamma_p X_{t-p})^\delta \\ + b_1 \sigma_{t-1}^\delta + \dots + b_q \sigma_{t-q}^\delta$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0, \delta > 0, a_i \geq 0, b_j \geq 0, \gamma_i \in [-1, 1], a_p \neq 0$  et  $b_q \neq 0$

**Remarque :** • Condition de stationnarité compliquée, voir après

- Bruit blanc faible :  $r_X(k) = 0$  pour  $|k| \geq 1$ .
- Variance conditionnelle non constante et dépendant du signe des valeurs

## Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + F_\theta(X_{t-1}, X_{t-2}, \dots)$$

$(\xi_t)$  bruit blanc de variance 1,  $\theta$  vecteur de paramètres,  $M_\theta$  et  $F_\theta$  fonctions.

**Remarque** : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que  $Q(z) \neq 0$  pour  $|z| \leq 1$   $Q(x) = 1 + b_1x + \dots + b_qx^q$

**Exemples** : • Si  $(X_t)$  AR( $p$ ),  $X_t = \xi_t - a_1 X_{t-1} - \dots - a_p X_{t-p}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi$ ,  $F_\theta(X_{t-1}, \dots) = -a_1 X_{t-1} - \dots - a_p X_{t-p}$

• Si  $(X_t)$  ARCH( $p$ ),  $X_t = \xi_t \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$ ,  $F_\theta(X_{t-1}, \dots) = 0$

## Nouvelle écriture des processus (2)

**Exemples :** • Si  $(X_t)$  MA(1),  $X_t = \xi_t + b_1 \xi_{t-1}$  avec  $|b_1| < 1$  alors

$$\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \implies X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad F_\theta(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si  $(X_t)$  ARMA( $p, q$ ),  $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$  quand  $Q(x) = 1 + b_1 x + \dots + b_q x^q$  a ses racines en dehors du disque unité

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad F_\theta(X_{t-1}, \dots) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si  $(X_t)$  GARCH( $p, q$ ), on montre que  $X_t = \xi_t \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}$

$$\implies M_\theta(X_{t-1}, \dots) = \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}, \quad F_\theta(X_{t-1}, \dots) = 0$$

## Processus affines causaux

$\mathbb{R}^\infty$  ensemble des suites n'ayant qu'un nombre fini de termes non nuls.

### Définition

Si  $(\varepsilon_t)_{t \in \mathbb{Z}}$  bruit blanc tel que  $\mathbb{E}[|\varepsilon_0|^r] < \infty$ ,  $r \geq 1$ ,  $(X_t)_{t \in \mathbb{R}}$  processus affine causal s'il existe  $F : \mathbb{R}^\infty \rightarrow \mathbb{R}$  et  $M : \mathbb{R}^\infty \rightarrow \mathbb{R}$  telles que

$$X_t = F((X_{t-k})_{k \geq 1}) + \varepsilon_t M((X_{t-k})_{k \geq 1}) \quad \text{pour tout } t \in \mathbb{Z}. \quad (3)$$

### Théorème (Doukhan et Wintenberger (2008))

Si  $(X_t)$  défini par (??) et  $F$  et  $G$  ont toutes leurs dérivées partielles existant presque partout sur  $\mathbb{R}^\infty$  avec

$$\alpha_i(F) = \sup_{x \in \mathbb{R}^\infty} \left| \frac{\partial}{\partial x_i} F(x) \right| \quad \text{et} \quad \alpha_i(M) = \sup_{x \in \mathbb{R}^\infty} \left| \frac{\partial}{\partial x_i} M(x) \right|.$$

Alors  $(X_t)$  stationnaire avec  $\mathbb{E}[|X_0|^r] < \infty$ , et  $\exists H : \mathbb{R}^\infty \rightarrow \mathbb{R}$  unique t.q.

$X_t = H((\varepsilon_{t-k})_{k \geq 0})$  dès que :

$$\sum_{i=1}^{\infty} \alpha_i(F) + (\mathbb{E}[|\varepsilon_0|^r])^{1/r} \sum_{i=1}^{\infty} \alpha_i(M) < 1. \quad (4)$$

## Processus $AR(\infty)$ et $ARCH(\infty)$

### Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus  $AR(\infty)$  stationnaire causal et  $\mathbb{E}[|X_0|^r] < \infty$  si

$$X_t = \varepsilon_t + \sum_{k=1}^{\infty} a_k X_{t-k} \quad \text{avec} \quad \sum_{i=1}^{\infty} |a_i| < 1$$

où  $(\varepsilon_t)$  bruit blanc tel que  $\mathbb{E}[|\varepsilon_0|^r] < \infty$ .

### Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus  $ARCH(\infty)$  stationnaire causal et  $\mathbb{E}[|X_0|^r] < \infty$  si

$$X_t = \varepsilon_t \left( a_0 + \sum_{k=1}^{\infty} a_k X_{t-k}^2 \right)^{1/2} \quad \text{avec} \quad (\mathbb{E}[|\varepsilon_0|^r])^{1/r} \sum_{i=1}^{\infty} \sqrt{a_i} < 1.$$

**Exercice :** Comparer les conditions de stationnarité pour  $AR(2)$ .

# Plan du cours



# Processus causaux et convergence

## Propriété

Soit  $(\xi_t)_{t \in \mathbb{Z}}$  un bruit blanc,  $H : \mathbb{R}^\infty \rightarrow \mathbb{R}$  et  $(X_t)_{t \in \mathbb{Z}}$  une série telle que  $X_t = H(\xi_t, \xi_{t-1}, \dots)$  pour  $t \in \mathbb{Z}$ .

① Si  $H(\xi_0, \xi_{-1}, \dots)$  existe p.s. alors  $(X_t)$  est une série stationnaire ;

② Si  $\mathbb{E}[|X_0|] = \mathbb{E}[|H(\xi_0, \xi_{-1}, \dots)|] < \infty$ ,

$$\bar{X}_n \xrightarrow[n \rightarrow \infty]{p.s.} \mathbb{E}[X_0] = \mathbb{E}[H(\xi_0, \xi_{-1}, \dots)]$$

③ Si  $\mathbb{E}[X_n | \xi_{n-1}, \xi_{n-2}, \dots] = 0$  et  $\mathbb{E}[X_0^2] < \infty$ ,  $(X_n)$  est une différence de martingale et

$$\sqrt{n} \bar{X}_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \mathbb{E}[X_0^2]).$$

**Exemple** : Un processus GARCH est une différence de martingale.

# Éléments de preuve des convergences

## Démonstration.

- ① Pour  $m \in \mathbf{N}^*$ , soit  $X_t^{(m)} = H(\xi_t, \dots, \xi_{t-m}, 0, \dots)$  pour  $t \in \mathbb{Z}$ . On montre alors que

$$(\xi_{t_1}, \dots, \xi_{t_1-m}, \dots, \xi_{t_k}, \dots, \xi_{t_k-m}) \stackrel{\mathcal{L}}{\sim} (\xi_{t_1+c}, \dots, \xi_{t_1+c-m}, \dots, \xi_{t_k+c}, \dots, \xi_{t_k+c-m})$$

Donc en appliquant  $H$ , on a bien  $(X_t^{(m)})$  stationnaire pour tout  $m$ . Et comme pour tout  $t$   $X_t^{(m)}$  tend en loi vers  $X_t$ , qui existe a.s., alors  $(X_t)$  stationnaire.

- ② Ceci peut être directement montré par le Théorème ergodique de Birkhoff, bien plus général. Pour le cas particulier traité, on peut comprendre pourquoi il y a convergence :  $(X_t^{(m)})_t$  est un processus stationnaire  $m$ -dépendant avec  $\mathbf{E}[|X_0^{(m)}|] < \infty$ , donc pour tout  $m \in \mathbf{N}^*$  et tout  $1 \leq k \leq (m+1)$ ,  $\frac{1}{n} \sum_{j=0}^{n-1} X_{k+j(m+1)}^{(m)} \xrightarrow[n \rightarrow \infty]{p.s.} \mathbf{E}[X_0^{(m)}]$ , d'où  $\bar{X}_{n(m+1)}^{(m)} \xrightarrow[n \rightarrow \infty]{p.s.} \mathbf{E}[X_0^{(m)}]$ . Comme  $\mathbf{E}[|X_0|] < \infty$ , on peut montrer que  $\mathbf{E}[X_0^{(m)}] \rightarrow \mathbf{E}[X_0]$  quand  $m \rightarrow \infty$ , tout comme  $X_t^{(m)} \rightarrow X_t$  p.s. quand  $m \rightarrow \infty$ .

- ③ Preuve délicate. Deux éléments :  $\mathbf{E}[X_k] = \mathbf{E}[\mathbf{E}[X_n | (\xi_{n-1}, \xi_{n-2}, \dots)]] = 0$ . Et pour  $k > 0$ ,

$$\begin{aligned} r_X(k) = \text{cov}(X_0, X_k) &= \mathbf{E}[X_0 X_k] = \mathbf{E}[\mathbf{E}[X_0 X_k | (\xi_{k-1}, \xi_{k-2}, \dots)]] \\ &= \mathbf{E}[X_0 \mathbf{E}[X_k | (\xi_{k-1}, \xi_{k-2}, \dots)]] = 0 \end{aligned}$$

donc  $(X_t)$  bruit blanc faible et on en déduit que  $\text{var}(\bar{X}_n) = \frac{1}{n} \mathbf{E}[X_0^2]$ .

# Convergence des moments

## Théorème (Doukhan et Wintenberger (2008))

Soit  $(X_t)_{t \in \mathbb{Z}}$  tel que  $X_t = M(X_{t-1}, X_{t-2}, \dots) \xi_t + F(X_{t-1}, X_{t-2}, \dots)$ .

On suppose que  $f$  et  $M$  de classe  $\mathcal{C}^1$  et il existe  $C > 0$ ,  $\rho \in [0, 1[$  tels que

$$\max \{ \alpha_i(F), \alpha_i(M) \} \leq C \rho^i \quad \text{pour tout } i \in \mathbb{N}.$$

- 1 Si  $\mathbf{E}[\xi_0^2] < \infty$  et  $\sum_{i=1}^{\infty} \alpha_i(F) + \alpha_i(M) (\mathbf{E}[\xi_0^2])^{1/2} < 1$ , alors  
$$\sqrt{n} (\bar{X}_n - \mathbf{E}[X_0]) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \gamma^2) \quad \text{avec } \gamma^2 = 2\pi f(0) = \sum_{\ell \in \mathbb{Z}} r_X(\ell).$$
- 2 Si  $\mathbf{E}[\xi_0^4] < \infty$  et  $\sum_{i=1}^{\infty} \alpha_i(F) + \alpha_i(M) (\mathbf{E}[\xi_0^4])^{1/4} < 1$ , alors  
$$\sqrt{n} (\hat{r}_n(i) - r_X(i))_{0 \leq i \leq m} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{m+1}(0, \Sigma(m)),$$
  
avec  $\Sigma(m) = (\sum_{k \in \mathbb{Z}} (\mathbf{E}[X_0 X_i X_k X_{k+j}] - r_X(i)r_X(j)))_{0 \leq i, j \leq m}$ .

## Démonstration.

Trop long



## Convergence des moments (2)

### Corollaire

Sous les mêmes hypothèses,

$$\sqrt{n} (\hat{\rho}_n(i) - \rho(i))_{1 \leq i \leq m} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_m(0, \Gamma(m))$$

$$\text{avec } \Gamma_{ij}(m) = \frac{1}{r_X^2(0)} \sum_{k \in \mathbb{Z}} (\mathbb{E}[X_0 X_i X_k X_{k+j}] - r_X(i)r_X(j)).$$

### Remarques :

- 1 Pour processus GARCH,  $\Gamma(m)$  diagonale, termes diagonaux =  $\frac{\mathbb{E}[X_0^2 X_i^2]}{r_X^2(0)}$   
donc pas tous égaux à 1 comme pour un bruit blanc fort
- 2 Non valide pour FARIMA( $p, d, q$ ) avec  $0 \leq d \leq \frac{1}{2}$  : vitesse en  $n^{1-2d}$  !

# Plan du cours

# Estimation par maximum de vraisemblance

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + F_\theta(X_{t-1}, X_{t-2}, \dots)$$

On veut estimer  $\theta \in \mathbb{R}^d$  à partir de  $(X_1, \dots, X_n)$  trajectoire observée.

**Première idée** : Utiliser le maximum de vraisemblance

- Requiert la connaissance de la loi de  $(X_1, \dots, X_n)$
- En dehors du cas gaussien, quasiment impossible car dépendance !
- Et les GARCH( $p, q$ ) ne sont jamais gaussiens !

## Estimation par maximum de vraisemblance (2)

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + F_\theta(X_{t-1}, X_{t-2}, \dots)$$

**Seconde idée :** Utiliser le maximum de vraisemblance conditionnel

⇒ On calcule la densité de  $(X_1, \dots, X_n)$  sachant  $(X_0, X_{-1}, \dots)$ .

On obtient alors la vraisemblance conditionnelle :

$$\begin{aligned} V_\theta(X_1, \dots, X_n) &= f(X_n | X_{n-1}, X_{n-2}, \dots) \\ &\quad \times f(X_{n-1} | X_{n-2}, X_{n-3}, \dots) \times \dots \times f(X_1 | X_0, X_{-1}, \dots) \end{aligned}$$

Si  $(\xi_t)$  gaussien  $\mathcal{N}(0, 1)$ ,  $X_k$  est gaussien sachant  $X_{k-1}, X_{k-2}, \dots$  d'où :

$$f(X_k | X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta(X_{k-1}, X_{k-2}, \dots)} \exp\left(-\frac{1}{2} \frac{(X_k - F_\theta(X_{k-1}, X_{k-2}, \dots))^2}{M_\theta^2(X_{k-1}, X_{k-2}, \dots)}\right)$$

## Estimation par maximum de vraisemblance (3)

On obtient alors la log-vraisemblance conditionnelle :

$$L_{\theta}(X_1, \dots, X_n) = -\frac{1}{2} \left( n \log(2\pi) + \sum_{t=1}^n \log(M_{\theta}^2(X_{t-1}, X_{t-2}, \dots)) + \frac{(X_t - F_{\theta}(X_{t-1}, X_{t-2}, \dots))^2}{M_{\theta}^2(X_{t-1}, X_{t-2}, \dots)} \right)$$

**Mais**  $X_0, X_{-1}, \dots$  non observées alors que par exemple pour un MA(1) on a

$$F_{\theta}(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\Rightarrow \begin{cases} \hat{M}_{\theta}^t = M_{\theta}(X_{t-1}, \dots, X_1, 0, 0, \dots) \\ \hat{F}_{\theta}^t = F_{\theta}(X_{t-1}, \dots, X_1, 0, 0, \dots) \end{cases} \text{ remplace } \begin{cases} M_{\theta}(X_{t-1}, X_{t-2}, \dots) \\ F_{\theta}(X_{t-1}, X_{t-2}, \dots) \end{cases}$$



## Estimation par quasi-maximum de vraisemblance (4)

On obtient alors la quasi-log-vraisemblance conditionnelle :

$$\widehat{L}_\theta(X_1, \dots, X_n) = -\frac{1}{2} \left( \sum_{t=1}^n \log((\widehat{M}_\theta^t)^2) + \frac{(X_t - \widehat{F}_\theta^t)^2}{(\widehat{M}_\theta^t)^2} \right)$$

On définit l'estimateur par quasi-maximum de vraisemblance QMLE :

$$\widehat{\theta}_n = \underset{\theta \in \Theta}{\text{Argmax}} \widehat{L}_\theta(X_1, \dots, X_n)$$

**Exemples :**

- Pour un processus AR(p),

$$(\widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(a_1, \dots, a_p) \in \mathbb{A} \subset \mathbb{R}^p}{\text{Argmin}} \sum_{k=p+1}^n (X_k - (a_1 X_{k-1} + \dots + a_p X_{k-p}))^2$$

- Pour un processus ARCH(p),

$$(\widehat{\omega}_0, \widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(\omega_0, a_1, \dots, a_p) \in \mathbb{A} \subset \mathbb{R}^{p+1}}{\text{Argmin}} \sum_{k=p+1}^n \log(\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2)$$

$$+ \frac{X_k^2}{\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2}$$

# Convergence du QMLE

Des conditions doivent être rajoutées pour obtenir la convergence :

- **(A0) Stationnarité d'ordre 4** : on note  $\Theta$  un compact de  $\Theta(r) = \{\theta \in \mathbb{R}^d, \sum_{i=1}^{\infty} \alpha_i(F_\theta) + \alpha_i(M_\theta) (\mathbf{E}[|\xi_0|^4])^{1/4} < 1\}$
- **(A1) Existence** : On suppose que  $\inf_{\theta \in \Theta, (x_n)_n \in \mathbb{R}^\infty} M_\theta((x_n)_n) = \underline{M} > 0$
- **(A2) Identification** :  $\theta \neq \theta' \iff F_\theta \neq F_{\theta'}$  ou  $M_\theta \neq M_{\theta'}$
- **(A3) Décorrélation** :  $\sup_{\theta \in \Theta} \left\{ |\alpha_i(M_\theta)| + |\alpha_i(F_\theta)| \right\} \leq C i^{-\kappa}, \kappa > 1$

## Convergence de l'estimateur QML (2)

On peut montrer que :

### Théorème

Soit  $(X_t)$  vérifiant :

$$X_t = M_{\theta^*}(X_{t-1}, X_{t-2}, \dots) \xi_t + F_{\theta^*}(X_{t-1}, X_{t-2}, \dots)$$

avec  $\mathbb{E}[\xi_0^4] < \infty$  et les hypothèses **(A0)**, **(A1)**, **(A2)**, et **(A3)**. Alors, même si  $(\xi_t)$  **non gaussien** on a  $\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{p.s.} \theta^*$ .

**Exemples** : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

**Remarque** : Le théorème est aussi valable pour  $\mathbb{E}[\xi_0^2] < \infty$  avec  $\kappa > 3/2$

# Preuve de la convergence presque sûre

## Démonstration.

- 1 On montre qu'avec  $q_t(\theta) = -\frac{1}{2} \left( \log((M_\theta^t)^2) + \frac{(X_t - F_\theta^t)^2}{(M_\theta^t)^2} \right)$  on a :

$$\sup_{\theta \in \Theta} \left\{ \left| \frac{1}{n} L_\theta(X_1, \dots, X_n) - \mathbb{E}[q_0(\theta)] \right| \right\} = \sup_{\theta \in \Theta} \left\{ \left| \frac{1}{n} \sum_{t=1}^n q_t(\theta) - \mathbb{E}[q_0(\theta)] \right| \right\} \xrightarrow[n \rightarrow \infty]{p.s.} 0.$$

Cela est obtenu car  $(q_t(\theta))_{t \in \mathbb{Z}}$  est une suite stationnaire causal et  $\mathbb{E}[\sup_{\theta \in \Theta} |q_0(\theta)|] < \infty$ .  
Pour ce dernier point,

$\sup_{\theta \in \Theta} |q_0(\theta)| \leq \sup_{\theta \in \Theta} \left\{ \frac{1}{2M^2} 2(X_0^2 + (F_\theta^0)^2) + \frac{1}{2} (|\log(M^2)| + (M_\theta^0)^2) \right\}$ . Mais

$$(M_\theta^0)^2 \leq 2((M_\theta^0 - M_\theta(0, \dots))^2 + M_\theta^2(0, \dots)) \leq 2 \left( \left( \sum_{i=1}^{\infty} \sup_{\theta \in \Theta} \alpha_i(M_\theta) X_{-i} \right)^2 + M_\theta^2(0, \dots) \right).$$

Or  $(\sum_{i=1}^{\infty} \sup_{\theta \in \Theta} \alpha_i(M_\theta) X_{-i})^2 \leq (\sum_{i=1}^{\infty} \sup_{\theta \in \Theta} \alpha_i(M_\theta)) (\sum_{i=1}^{\infty} \sup_{\theta \in \Theta} \alpha_i(M_\theta) X_{-i}^2)$  par l'inégalité de Cauchy-Schwarz. D'où  $\mathbb{E}[\sup_{\theta \in \Theta} (M_\theta^0)^2] < \infty$  car  $\sum_{i=1}^{\infty} \sup_{\theta \in \Theta} \alpha_i(M_\theta) < \infty$ .  
Même chose avec  $\mathbb{E}[\sup_{\theta \in \Theta} (F_\theta^0)^2]$ . D'où  $\mathbb{E}[\sup_{\theta \in \Theta} |q_0(\theta)|] < \infty$ .

- 2 On a également  $\sup_{\theta \in \Theta} \left\{ \left| \frac{1}{n} L_\theta(X_1, \dots, X_n) - \frac{1}{n} \widehat{L}_\theta(X_1, \dots, X_n) \right| \right\} \xrightarrow[n \rightarrow \infty]{p.s.} 0$ . Ceci se fait en

majorant  $\mathbb{E}[\sup_{\theta \in \Theta} \{|q_t(\theta) - \widehat{q}_t(\theta)|\}] \leq C \mathbb{E} \left[ \sup_{\theta \in \Theta} \left\{ (|M_\theta^t - \widehat{M}_\theta^t|^3 + |F_\theta^t - \widehat{F}_\theta^t|^3)^{1/3} \right\} \right]$

et  $|M_\theta^t - \widehat{M}_\theta^t| \leq \sum_{k=t}^{\infty} \sup_{\theta \in \Theta} \alpha_k(M_\theta) |X_{t-k}|$  soit

$\mathbb{E}[\sup_{\theta \in \Theta} \{|q_t(\theta) - \widehat{q}_t(\theta)|\}] \leq C \sum_{k=t}^{\infty} \sup_{\theta \in \Theta} \{\alpha_k(M_\theta) + \alpha_k(F_\theta)\}$ . Puis on utilise le

fait que pour toute série  $(Z_t)$  si  $\sum_t \mathbb{E}[|Z_t|]/t < \infty$  alors  $\frac{1}{n} \sum_{t=1}^n Z_t \xrightarrow[n \rightarrow \infty]{p.s.} 0$ .

## Preuve de la convergence presque sûre (2)

### Démonstration.

3 On montre enfin que  $\max_{\theta \in \Theta} \mathbf{E}[q_0(\theta)] = \mathbf{E}[q_0(\theta^*)]$ . En effet, pour tout  $\theta \in \Theta$ ,

$$\begin{aligned}\mathbf{E}[q_0(\theta^*) - q_0(\theta)] &= \frac{1}{2} \mathbf{E} \left[ \left( \log \left( \frac{(M_\theta^0)^2}{(M_{\theta^*}^0)^2} \right) + \frac{(X_0 - F_\theta^0)^2}{(M_\theta^0)^2} - \frac{(X_0 - F_{\theta^*}^0)^2}{(M_{\theta^*}^0)^2} \right) \right] \\ &= \frac{1}{2} \mathbf{E} \left[ \left( \log \left( \frac{(M_\theta^0)^2}{(M_{\theta^*}^0)^2} \right) + \frac{(M_{\theta^*}^0 \xi_0 + F_{\theta^*}^0 - F_\theta^0)^2}{(M_\theta^0)^2} - \xi_0^2 \right) \right] \\ &= \frac{1}{2} \mathbf{E} \left[ \mathbf{E} \left[ \left( \log \left( \frac{(M_\theta^0)^2}{(M_{\theta^*}^0)^2} \right) + \frac{(M_{\theta^*}^0 \xi_0 + F_{\theta^*}^0 - F_\theta^0)^2}{(M_\theta^0)^2} - \xi_0^2 \right) \mid (X_{-1}, X_{-2}, \dots) \right] \right] \\ &= \frac{1}{2} \mathbf{E} \left[ \left( \log \left( \frac{(M_\theta^0)^2}{(M_{\theta^*}^0)^2} \right) + \frac{(M_{\theta^*}^0)^2}{(M_\theta^0)^2} - 1 + \frac{(F_{\theta^*}^0 - F_\theta^0)^2}{(M_\theta^0)^2} \right) \right] \\ &= \frac{1}{2} \mathbf{E} \left[ h \left( \frac{(M_{\theta^*}^0)^2}{(M_\theta^0)^2} \right) + \frac{(F_{\theta^*}^0 - F_\theta^0)^2}{(M_\theta^0)^2} \right]\end{aligned}$$

et  $h(x) = -\log(x) + x - 1 > 0$  pour  $x \neq 1$ . D'où, avec **(A2)**,  $\max_{\theta \in \Theta} \mathbf{E}[q_0(\theta)] = \mathbf{E}[q_0(\theta^*)]$ .

On conclut par  $\sup_{\theta \in \Theta} \left\{ \left| \frac{1}{n} \widehat{L}_\theta(X_1, \dots, X_n) - \mathbf{E}[q_0(\theta)] \right| \right\} \xrightarrow[n \rightarrow \infty]{p.s.} 0$  et  $\widehat{\theta}_n$  maximisant  $\widehat{L}_\theta(X_1, \dots, X_n)$   $\square$

# Normalité asymptotique du QMLE

Conditions à rajouter pour obtenir la normalité asymptotique :

- **(A4) Régularité** : On suppose que pour tout  $\theta \in \Theta$ ,  $\frac{\partial^3 F_\theta(x)}{\partial x \partial \theta^2}$  et  $\frac{\partial^3 M_\theta(x)}{\partial x \partial \theta^2}$  existent sur  $\mathbb{R}^\infty$  et il existe  $\ell' > 3/2$  tel que

$$\sup_{\theta \in \Theta} \sup_{x \in \mathbb{R}^\infty} \left( \left\| \frac{\partial^2}{\partial \theta \partial x_j} F_\theta(x) \right\| + \left\| \frac{\partial^2}{\partial \theta \partial x_j} M_\theta(x) \right\| + \left\| \frac{\partial^3}{\partial \theta^2 \partial x_j} F_\theta(x) \right\| + \left\| \frac{\partial^3}{\partial \theta^2 \partial x_j} M_\theta(x) \right\| \right) = O(j^{-\ell'}).$$

- **(A5) Covariance asymptotique** : On suppose que la matrice  $\left( \mathbb{E} \left[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} q_0(\theta^*) \right] \right)_{ij}$  est inversible.

## Normalité asymptotique du QMLE (2)

On peut montrer que :

### Théorème

Sous les conditions du théorème précédent et avec **(A4)** et **(A5)** on a

$$\sqrt{n}(\hat{\theta}_n - \theta^*) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_d(0, F(\theta^*)^{-1} G(\theta^*) F(\theta^*)^{-1}),$$

avec  $F(\theta^*) = \left( \mathbf{E} \left[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} q_0(\theta^*) \right] \right)_{ij}$  et  $G(\theta^*) = \left( \mathbf{E} \left[ \frac{\partial}{\partial \theta_i} q_0(\theta^*) \frac{\partial}{\partial \theta_j} q_0(\theta^*) \right] \right)_{ij}$

**Exemples** : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

## Normalité asymptotique du QMLE (3)

**Remarque** : En utilisant la convergence p.s. de  $\hat{\theta}_n$  vers  $\theta^*$ , on a

$$\begin{cases} \left( \frac{\partial^2}{\partial \theta_i \partial \theta_j} \frac{1}{n} \log (\widehat{L}_{\hat{\theta}_n}(X_1, \dots, X_n)) \right)_{ij} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} F(\theta^*) \\ \left( \frac{\partial}{\partial \theta_i} \frac{1}{n} \log (\widehat{L}_{\hat{\theta}_n}(X_1, \dots, X_n)) \frac{\partial}{\partial \theta_j} \frac{1}{n} \log (\widehat{L}_{\hat{\theta}_n}(X_1, \dots, X_n)) \right)_{ij} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} G(\theta^*) \end{cases}$$

Avec le Lemme de Slutsky, on en déduit :

$$\sqrt{n} (\widehat{F}_n \widehat{G}_n^{-1} \widehat{F}_n)^{1/2} (\hat{\theta}_n - \theta^*) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_d(0, I_d).$$

On peut alors obtenir des intervalles de confiances, des tests de Wald,...



# Preuve de la normalité asymptotique

## Démonstration.

La preuve se fait en quatre temps :

- ① Par la formule de Taylor-Lagrange, avec  $\bar{\theta} = c\hat{\theta}_n + (1-c)\theta^*$  et  $0 < c < 1$ , on a

$$n^{-1/2} \frac{\partial \widehat{L}_{\theta^*}(X_1, \dots, X_n)}{\partial \theta} = n^{-1/2} \frac{\partial \widehat{L}_{\hat{\theta}_n}(X_1, \dots, X_n)}{\partial \theta} + n^{1/2}(\theta^* - \hat{\theta}_n) \frac{1}{n} \frac{\partial^2 \widehat{L}_{\bar{\theta}}(X_1, \dots, X_n)}{\partial \theta^2}.$$

Mais  $\frac{\partial \widehat{L}_{\hat{\theta}_n}(X_1, \dots, X_n)}{\partial \theta} = 0$  car  $\hat{\theta}_n$  extremum. D'où

$$n^{-1/2} \frac{\partial \widehat{L}_{\theta^*}(X_1, \dots, X_n)}{\partial \theta} = n^{1/2}(\theta^* - \hat{\theta}_n) \frac{1}{n} \frac{\partial^2 \widehat{L}_{\bar{\theta}}(X_1, \dots, X_n)}{\partial \theta^2} \quad (5)$$

- ②  $n^{-1/2} \frac{\partial L_{\theta^*}(X_1, \dots, X_n)}{\partial \theta} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_d(0, G(\theta^*))$  par CLT pour différence de martingale car

$$\mathbb{E}\left[\frac{\partial q_t(\theta^*)}{\partial \theta} \mid X_{t-1}, X_{t-2}, \dots\right] = 0.$$

- ③ On montre que  $n^{-1/2} \mathbb{E}\left[\sup_{\theta \in \Theta} \left\{ \left\| \frac{\partial L_{\theta}(X_1, \dots, X_n)}{\partial \theta} - \frac{\partial \widehat{L}_{\theta}(X_1, \dots, X_n)}{\partial \theta} \right\| \right\}\right] \xrightarrow[n \rightarrow +\infty]{} 0$  grâce à **(A4)**, d'où

$$n^{-1/2} \left\| \frac{\partial \widehat{L}_{\theta^*}(X_1, \dots, X_n)}{\partial \theta} - \frac{\partial L_{\theta^*}(X_1, \dots, X_n)}{\partial \theta} \right\| \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0 \quad (6)$$



## Preuve de la normalité asymptotique (2)

### Démonstration.

④ On a  $\mathbf{E} \left[ \sup_{\theta \in \Theta} \left\{ \left\| \frac{1}{n} \frac{\partial^2 L_{\theta}(X_1, \dots, X_n)}{\partial \theta^2} - \frac{1}{n} \frac{\partial^2 \widehat{L}_{\theta}(X_1, \dots, X_n)}{\partial \theta^2} \right\| \right\} \right] \xrightarrow[n \rightarrow +\infty]{} 0$  grâce à **(A4)** et  $\frac{1}{n} \frac{\partial^2 L_{\theta^*}(X_1, \dots, X_n)}{\partial \theta^2} \xrightarrow[n \rightarrow \infty]{p.s.} F(\theta^*)$  par LFGN ergodique. D'où

$$\frac{1}{n} \frac{\partial^2 \widehat{L}_{\theta}(X_1, \dots, X_n)}{\partial \theta^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} F(\theta^*) \quad (7)$$

De (??) et avec **(A5)**, on obtient que pour  $n$  suffisamment grand  $\frac{1}{n} \frac{\partial^2 \widehat{L}_{\theta}(X_1, \dots, X_n)}{\partial \theta^2}$  est inversible et converge en probabilité vers  $F(\theta^*)^{-1}$  qui existe. Par suite, en considérant (??), et avec (??), on en déduit que

$$n^{1/2}(\theta^* - \widehat{\theta}_n) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} F(\theta^*)^{-1} Z_n \quad \text{où} \quad Z_n \stackrel{\mathcal{L}}{\sim} \mathcal{N}_d(0, G(\theta^*)).$$

D'où le résultat final. □

**Exercice :** Pour un processus AR(1), avec  $\mu_4^* = \mathbf{E}[\xi_0^4]$ , démontrer que

$$\sqrt{n} \begin{pmatrix} \widehat{\alpha}_n - \alpha^* \\ \widehat{\sigma}_n^2 - \sigma^{*2} \end{pmatrix} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 - \alpha^{*2} & 0 \\ 0 & \sigma^{*4}(\mu_4^* - 1) \end{pmatrix} \right)$$

# Correction de l'exercice

## Démonstration.

On a  $\theta = {}^t(\alpha, \sigma^2)$  et  $M_\theta^0 = \sigma$ ,  $F_\theta^0 = \alpha X_{-1}$ . Les conditions **(A1)**-**(A4)** sont vérifiées. D'où  $q_0(\theta) = -\frac{1}{2} \left( \log(\sigma^2) + \frac{(X_0 - \alpha X_{-1})^2}{\sigma^2} \right)$ . On en déduit :

$$\frac{\partial q_0(\theta)}{\partial \theta} = \left( \begin{array}{c} \frac{X_{-1}(X_0 - \alpha X_{-1})}{\sigma^2} \\ \frac{(X_0 - \alpha X_{-1})^2 - \sigma^2}{2\sigma^4} \end{array} \right) \quad \text{et} \quad \frac{\partial^2 q_0(\theta)}{\partial \theta^2} = - \left( \begin{array}{cc} \frac{X_{-1}^2}{\sigma^2} & \frac{X_{-1}(X_0 - \alpha X_{-1})}{\sigma^4} \\ \frac{X_{-1}(X_0 - \alpha X_{-1})}{\sigma^4} & \frac{(X_0 - \alpha X_{-1})^2}{\sigma^6} - \frac{1}{2\sigma^4} \end{array} \right)$$

En prenant l'espérance pour  $\theta = \theta^*$ , d'où  $X_0 - \alpha^* X_{-1} = \sigma^* \xi_0$  indépendant de  $X_{-1}$ , et avec  $\mathbb{E}[X_0^2] = \frac{\sigma^{*2}}{1 - \alpha^{*2}}$ , on en déduit que :

$$F(\theta^*) = \left( \begin{array}{cc} -\frac{1}{1 - \alpha^{*2}} & 0 \\ 0 & -\frac{1}{2\sigma^{*4}} \end{array} \right) \quad \text{et} \quad G(\theta^*) = \left( \begin{array}{cc} \frac{1}{1 - \alpha^{*2}} & 0 \\ 0 & \frac{\mu_4^* - 1}{4\sigma^{*4}} \end{array} \right)$$

et **(A5)** est vérifiée. D'où le résultat. □

# Plan du cours

## Sélection de modèles

**Problématique** : On suppose  $(X_1, \dots, X_n)$  trajectoire observée. Quel modèle choisir pour modéliser cette trajectoire ?

**Exemple** : Soit la trajectoire des log-rendements d'un indice financier. Est-il préférable de la modéliser par un GARCH(1, 1) ou par un ARCH(2, 0) ? Ou bien un ARMA(3, 2) ?

### Définition

On considère  $\mathcal{M}$  une famille de modèles affines causaux et on notera  $m \in \mathcal{M}$  un modèle, vérifiant, avec  $\theta(m) \in \mathbb{R}^d$

$$X_t = M_{\theta(m)}(X_{t-1}, X_{t-2}, \dots) \xi_t + f_{\theta(m)}(X_{t-1}, X_{t-2}, \dots)$$

**Exemple** :  $\mathcal{M}$  famille des  $AR(p)$  pour  $0 \leq p \leq 10$  :  $\theta(m) = (a_1, \dots, a_{10}, \sigma_\xi^2)$

## Sélection de modèles (2)

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

- le critère AIC pour ce modèle :  $\widehat{AIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + 2 |m|$  ;
- le critère BIC pour ce modèle :  $\widehat{BIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + \log(n) |m|$

où  $|m|$  est le nombre composantes non nulles de  $\theta(m)$ .

La famille de modèle peut-être hiérarchique (exemple : AR( $p$ )) ou non (exemple : ARMA( $p, q$ ) ou ARMA( $p, q$ )+GARCH( $p', q'$ )).

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

$$\widehat{m}_{AIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{AIC}(m) \quad \text{et} \quad \widehat{m}_{BIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{BIC}(m).$$

## Sélection de modèles (3)

### Théorème

Soit  $\mathcal{M}$  famille de taille finie de modèles affines causaux. S'il existe un "vrai" modèle  $m^* \in \mathcal{M}$  tel que  $(X_1, \dots, X_n)$  trajectoire observée issue de  $m^*$ , sous **(A1)-(A5)** et s'il existe  $\ell'' > 5/2$  tel que

$$\sup_{\theta(m) \in \Theta} \sup_{x \in \mathbb{R}^\infty} \left( \left\| \frac{\partial}{\partial x_j} F_{\theta(m)}(x) \right\| + \left\| \frac{\partial}{\partial x_j} M_{\theta(m)}(x) \right\| + \left\| \frac{\partial^2}{\partial \theta \partial x_j} F_{\theta(m)}(x) \right\| \right. \\ \left. + \left\| \frac{\partial^2}{\partial \theta \partial x_j} M_{\theta(m)}(x) \right\| + \left\| \frac{\partial^3}{\partial \theta^2 \partial x_j} F_{\theta(m)}(x) \right\| + \left\| \frac{\partial^3}{\partial \theta^2 \partial x_j} M_{\theta(m)}(x) \right\| \right) = O(j^{-\ell''}).$$

Alors,  $\mathbb{P}(\hat{m}_{BIC} = m^*) \xrightarrow[n \rightarrow +\infty]{} 1$  et  $\mathbb{P}(m^* \subset \hat{m}_{AIC}) \xrightarrow[n \rightarrow +\infty]{} 1$ ,

$$\sqrt{n} (\hat{\theta}_n(\hat{m}_{BIC}) - \theta^*(m^*)) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{|m^*|} (0, F(\theta^*(m^*))^{-1} G(\theta^*(m^*)) F(\theta^*(m^*)))$$

**Exemples** : Le théorème est valable pour les processus ARMA, GARCH, APARCH, ARMA-GARCH,...

## Sélection de modèles (4)

**Remarque :** AIC reste intéressant car minimise les erreurs de prédictions

**Remarque :** Pour les modèles affines causaux, il est toujours possible de trouver  $M_\theta$  et  $F_\theta$ , où  $\theta \in \Theta \subset \mathbb{R}^d$  tel que pour tout  $m \in \mathcal{M}$ ,  $\theta(m) \in \Theta$ .

### Démonstration.

Si  $m_0 \in \mathcal{M}$  est tel que  $\hat{m} = m_0$  et  $m^* \subset m_0$ , ("overfitted" model). On obtient

$$\begin{aligned}\mathbb{P}(\widehat{BIC}(m_0) \leq \widehat{BIC}(m^*)) &= \mathbb{P}\left(-2\widehat{L}_n(\widehat{\theta}(m_0)) + |m_0| \log(n) \leq -2\widehat{L}_n(\widehat{\theta}(m^*)) + |m^*| \log(n)\right) \\ &= \mathbb{P}\left(\frac{1}{\log(n)}(\widehat{L}_n(\widehat{\theta}(m^*)) - \widehat{L}_n(\widehat{\theta}(m_0))) \leq \frac{(|m^*| - |m_0|)}{2}\right) \xrightarrow{n \rightarrow +\infty} 0,\end{aligned}$$

après avoir montré que  $\mathbb{E}\left[\frac{1}{\log(n)}|\widehat{L}_n(\widehat{\theta}(m^*)) - \widehat{L}_n(\widehat{\theta}(m_0))|\right] \xrightarrow{n \rightarrow +\infty} 0$ .

Si  $m_0 \in \mathcal{M}$  est tel que  $\hat{m} = m_0$  et  $m^* \not\subset m_0$ , ("misspecified" model)

$$\frac{\widehat{BIC}(m_0) - \widehat{BIC}(m^*)}{n} = D_{KL}(m^*, m_0) + \frac{\log(n)}{n}(|m_0| - |m^*|) + o_{p.s.}(1)$$

où  $D_{KL}(m^*, m_0) = \mathbb{E}[q_0(\theta(m^*)) - q_0(\theta^*(m_0))] > 0$  et  $\theta^*(m_0) = \text{Argmin}_{\theta(m_0)} \mathbb{E}[q_0(\theta(m_0))]$  □



## Test d'adéquation

**Problématique** : Une fois un modèle  $\hat{m}_{BIC}$  obtenu pour modéliser  $(X_1, \dots, X_n)$  comment s'assurer que ce modèle convient ?

$\implies$  Test d'adéquation de type portemanteau

$$\begin{cases} H_0 : \exists m^* \in \mathcal{M}, \text{ tel que } (X_1, \dots, X_n) \text{ trajectoire de } X \text{ avec } \theta^*(m^*) \\ H_1 : \nexists m \in \mathcal{M}, \text{ tel que } (X_1, \dots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m) \end{cases}$$

**Remarque** : Sous  $H_0$ , on a  $X_t = M_{\theta^*(m^*)} \xi_t + f_{\theta^*(m^*)}$  pour  $t \in \mathbb{Z}$

## Test d'adéquation (2)

On définit les résidus pour le modèle  $m$  :

$$\widehat{\xi}_t(m) := (\widehat{M}_{\widehat{\theta}(m)}^t)^{-1} (X_t - \widehat{f}_{\widehat{\theta}(m)}^t),$$

Pour  $K \in \mathbb{N}^*$  fixé, vecteur d'autocorrélations des carrés des résidus :

$$\widehat{\rho}(m) := (\widehat{\rho}_1(m), \dots, \widehat{\rho}_K(m))',$$

$$\text{où } \widehat{\rho}_k(m) := \frac{\widehat{\gamma}_k(m)}{\widehat{\gamma}_0(m)} \text{ et } \widehat{\gamma}_k(m) := \frac{1}{n} \sum_{t=k+1}^n (\widehat{\xi}_t^2(m) - 1)(\widehat{\xi}_{t-k}^2(m) - 1)$$

**Remarque** : On a  $\mathbb{E}[\xi_0^2] = 1$  et on montre que  $\mathbb{E}[\widehat{\xi}_0^2] \xrightarrow{n \rightarrow +\infty} 1$  sous  $H_0$ .

Ceci explique pourquoi 1 remplace  $\widehat{\xi}_{nn}^2$ .

## Test d'adéquation (2)

### Théorème

Sous les hypothèses du théorème "BIC", si  $\mathbb{E}[\xi_0^3] = 0$  et  $\mu_4 = \mathbb{E}[\xi_0^4]$ , avec

$$V_K(\theta^*(m^*)) = I_K + \frac{J_K(m^*) F(\theta^*(m^*))^{-1}}{(\mu_4 - 1)^2} \left( G(\theta^*(m^*)) F(\theta^*(m^*))^{-1} + (\mu_4 - 1) I_K \right)^t J_K(m^*),$$

où  $J_K(m^*) = -2 \left( \mathbb{E} \left[ (\xi_0^2 - 1) \frac{\partial}{\partial \theta_j} \log (M_{\theta^*(m^*)}^i) \right] \right)_{1 \leq i \leq K, j \in m^*}$ , on a

$$\sqrt{n} \hat{\rho}(m^*) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_K(0, V_K(\theta^*(m^*))).$$

**Remarque :** Concrètement, on peut remplacer  $V_K(\theta^*(m^*))$  par  $\hat{V}_K(m^*)$ , où

$$\begin{cases} \hat{V}_K(m^*) = I_K + \frac{1}{(\hat{\mu}_4 - 1)^2} \hat{J}_K F(\hat{\theta}(m^*))^{-1} \left( G(\hat{\theta}(m^*)) F(\hat{\theta}(m^*))^{-1} + (\hat{\mu}_4 - 1) I_K \right)^t \hat{J}_K \\ \hat{J}_K = \frac{1}{n} \sum_{t=1}^{n-i} (\hat{\xi}_t^2(m^*) - 1) \left( \frac{\partial \log (\hat{M}_{\hat{\theta}(m^*)}^{t+i})}{\partial \theta_j} \right)_{1 \leq i \leq K, j}, \quad \hat{\mu}_4 = \frac{1}{n} \sum_{t=1}^n \hat{\xi}_t^4(m^*) \end{cases}$$

# Preuve de la normalité asymptotique de $\widehat{\rho}(m^*)$

## Démonstration.

Cette preuve peut se décomposer en 3 parties :

1. Tout d'abord on prouve que  $\sqrt{n} \sup_{\theta \in \Theta(m^*)} \{|\widehat{\gamma}_k(\theta(m^*)) - \gamma_k(\theta(m^*))|\} \xrightarrow[n \rightarrow \infty]{p.s.} 0$  où

$$\gamma = \frac{1}{n} \sum_{t=1}^{n-k} (\xi_t^2(\theta(m)) - 1) (\xi_{t+k}^2(\theta(m)) - 1) \text{ où } \xi_t(\theta(m)) := (M_{\theta(m)}^t)^{-1} (X_t - f_{\theta(m)}^t).$$

2. Par Taylor-Lagrange  $\sqrt{n} \gamma_k(\widehat{\theta}(m^*)) = \sqrt{n} \gamma_k(\theta^*(m^*)) + \partial_{\theta} \gamma_k(\bar{\theta}^{(k)}) \sqrt{n} ((\widehat{\theta}(m^*))_i - \theta_i^*)_{i \in m^*}$  et  $(\partial_{\theta} \gamma_k(\bar{\theta}^{(k)}))_{1 \leq k \leq K} \xrightarrow[n \rightarrow \infty]{p.s.} J_K(m^*)$ . Ainsi

$$(\partial_{\theta} \gamma_k(\bar{\theta}^{(k)}))_{1 \leq k \leq K} \sqrt{n} ((\widehat{\theta}(m^*))_i - \theta_i^*)_{i \in m^*} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_K \left( 0, J_K(m^*) F(\theta^*(m^*))^{-1} G(\theta^*(m^*)) F(\theta^*(m^*))^{-1} J_K(m^*) \right).$$

3. On a aussi  $\sqrt{n} (\gamma_k(\theta^*))_{1 \leq k \leq K} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_K(0, (\mu_4 - 1)^2 I_K)$ .

4. Enfin, après calcul :

$$\text{cov}(\sqrt{n} \gamma(\theta^*), (\partial_{\theta} \gamma_k(\bar{\theta}^{(k)}))_k \sqrt{n} ((\widehat{\theta}(m^*))_i - \theta_i^*)_{i \in m^*}) \xrightarrow[n \rightarrow +\infty]{} \frac{1}{2} (\mu_4 - 1) J_K(m^*) F(\theta^*(m^*))^{-1} J_K(m^*).$$

## Test d'adéquation (3)

- Si  $m^*$  est connu : on définit la statistique de test :

$$\hat{Q}_K(m^*) := n^t \hat{\rho}(m^*) (\hat{V}_K(m^*))^{-1} \hat{\rho}(m^*)$$

### Proposition

Sous les hypothèses du théorème "BIC", et si  $m^*$  est le vrai modèle

$$\hat{Q}_K(m^*) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2(K).$$

### Démonstration.

On a facilement  $(V_K(\theta^*(m^*)))^{-1/2} \sqrt{n} \hat{\rho}(m^*) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_K(0, I_K)$ . On en déduit que

$n^t \hat{\rho}(m^*) (V_K(\theta^*(m^*)))^{-1} \hat{\rho}(m^*) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2(K)$ . On a, grâce aux hypothèses **A0** et **A4**,

$\hat{V}_K(m^*) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} V_K(\theta^*(m^*))$  et le Lemme de Slutsky permet de conclure. □

**Conséquence** : Pour tester :  $\begin{cases} H_0^* : \text{le vrai modèle est } m^* \\ H_1^* : \text{le vrai modèle n'est pas } m^* \end{cases}$

## Test d'adéquation (4)

- Si  $m^*$  est inconnu mais  $m^* \in \mathcal{M}$ , on définit la statistique de test :

$$\widehat{Q}_K(\widehat{m}_{BIC}) := n^t \widehat{\rho}(\widehat{m}_{BIC}) \widehat{V}_K^{-1} \widehat{\rho}(\widehat{m}_{BIC})$$

$$\text{où } \begin{cases} \widehat{V}_K = I_K + \frac{1}{(\widehat{\mu}_4 - 1)^2} \widehat{J}_K F(\theta(\widehat{m}_{BIC}))^{-1} \left( G(\theta(\widehat{m}_{BIC})) F(\theta(\widehat{m}_{BIC}))^{-1} + (\widehat{\mu}_4 - 1) I_K \right)^t \widehat{J}_K \\ \widehat{J}_K = \frac{1}{n} \sum_{t=1}^{n-i} (\widehat{e}_t^2(\widehat{m}_{BIC}) - 1) \left( \frac{\partial \log(\widehat{M}_{\widehat{\theta}(\widehat{m}_{BIC})}^{t+i})}{\partial \theta_j} \right)_{1 \leq i \leq K, j}, \quad \widehat{\mu}_4 = \frac{1}{n} \sum_{t=1}^n \widehat{\varepsilon}_t^4(\widehat{m}_{BIC}) \end{cases}$$

### Théorème

Sous les hypothèses du théorème "BIC" et sous  $H_0$ ,  $\widehat{Q}_K(\widehat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2(K)$

**Remarque :** • Le test est utilisable pour les processus ARMA, GARCH, APARCH, ARMA-GARCH,...

- Pour un ARMA,  $\widehat{Q}_K(\widehat{m}_{BIC})$  peut être remplacé par  $n \|\widehat{\rho}(\widehat{m}_{BIC})\|^2$
- Pour un GARCH,  $\widehat{V} \simeq I_K - \frac{1}{4} \widehat{J}_K (\widehat{G}(\widehat{\theta}(\widehat{m}_{BIC})))^{-1} \widehat{J}_K$

# Plan du cours

## Prédiction à partir d'un modèle

On suppose observée  $(X_1, \dots, X_n)$ .

**Objectif** : On veut prédire  $X_{n+h}$ , soit  $\hat{X}_{n+h}$  ( $h$  horizon)

**Méthode 1** : On suppose que  $X_t = M_{\theta^*}^t \xi_t + f_{\theta^*}^t$  pour  $t \in \mathbb{Z}$

**Procédure** : On veut écrire  $\hat{X}_{n+h} = g(X_1, \dots, X_n)$  et on cherche

$$\hat{g} = \text{Arg min}_{g \in \mathbf{L}^2} \mathbf{E}[(X_{n+h} - g(X_1, \dots, X_n))^2]$$

$$\implies \hat{X}_{n+h} = \mathbf{E}[X_{n+h} \mid (X_1, \dots, X_n)]$$



## Prédiction à partir d'un modèle (2)

**Exemples** : On veut calculer  $\mathbb{E}[X_{n+h} \mid (X_1, \dots, X_n)]$

① Si  $(X_t)$  AR(1) causal  $X_t = \hat{\alpha} X_{t-1} + \xi_t$  d'où :

$$\blacktriangleright \hat{X}_{n+1} = \mathbb{E}[\hat{\alpha} X_n + \xi_{n+1} \mid (X_1, \dots, X_n)] = \hat{\alpha} X_n$$

$$\blacktriangleright \hat{X}_{n+2} = \mathbb{E}[\hat{\alpha}^2 X_n + \hat{\alpha} \xi_{n+1} + \xi_{n+2} \mid (X_1, \dots, X_n)] = \hat{\alpha}^2 X_n$$

$$\blacktriangleright \hat{X}_{n+h} = \mathbb{E}[\hat{\alpha}^h X_n + \sum_{i=1}^h \hat{\alpha}^{h-i} \xi_{n+i} \mid (X_1, \dots, X_n)] = \hat{\alpha}^h X_n$$

② Si  $(X_t)$  MA(1) inversible

$$X_t = \xi_t + \hat{\beta} \xi_{t-1} \implies X_t = \xi_t - \sum_{k=1}^{\infty} (-\hat{\beta})^k X_{t-k}$$

$$\blacktriangleright \hat{X}_{n+1} = \mathbb{E}[\xi_{n+1} - \sum_{k=0}^{\infty} (-\hat{\beta})^{k+1} X_{n-k} \mid (X_1, \dots, X_n)] = - \sum_{k=0}^{n-1} (-\hat{\beta})^{k+1} X_{n-k}$$

$$\blacktriangleright \hat{X}_{n+h} = \mathbb{E}[\xi_{n+h} + \hat{\beta} \xi_{n+h-1} \mid (X_1, \dots, X_n)] = 0 \text{ pour } h \geq 2$$

③ Si  $(X_t)$  GARCH( $p, q$ ) alors  $X_t = \sigma_t \xi_t$  d'où  $\hat{X}_{n+h} = 0$  si  $h \geq 1$

## Exercice

Exercice : Si  $(X_t)$  est un GARCH(1, 1), déterminer  $\hat{\sigma}_{n+1}^2$  ?  $\hat{\sigma}_{n+2}^2$  ?

### Démonstration.

- $\hat{\sigma}_{n+1}^2 = \mathbb{E}[\sigma_{n+1}^2 | (X_1, \dots, X_n)]$ . Mais  $\sigma_{n+1}^2 = \omega_0 + a_1 X_n^2 + b_1 \sigma_n^2$ , soit encore  $\sigma_{n+1}^2 = \omega_0 + a_1 X_n^2 + b_1(\omega_0 + a_1 X_{n-1}^2 + b_1 \sigma_{n-1}^2) = \omega_0(1 + b_1) + a_1 X_n^2 + a_1 b_1 X_{n-1}^2 + b_1^2 \sigma_{n-1}^2$ . Par itération,

$$\sigma_{n+1}^2 = \omega_0(1 + b_1 + \dots + b_1^{n-1}) + \sum_{i=0}^n b_1^i X_{n-i}^2 + b_1^n \sigma_0^2$$

En conséquence  $\hat{\sigma}_{n+1}^2 = \hat{\omega}_0(1 + \hat{b}_1 + \dots + \hat{b}_1^{n-1}) + \hat{a}_1 \sum_{i=0}^n \hat{b}_1^i X_{n-i}^2$

- $\hat{\sigma}_{n+2}^2 = \mathbb{E}[\sigma_{n+2}^2 | (X_1, \dots, X_n)]$ . Mais  $\sigma_{n+2}^2 = \omega_0 + a_1 X_{n+1}^2 + b_1 \sigma_{n+1}^2$ , soit encore  $\sigma_{n+2}^2 = \omega_0 + a_1(\omega_0 + a_1 X_n^2 + b_1 \sigma_n^2) + b_1 \sigma_{n+1}^2 = \omega_0(1 + a_1) + a_1^2 X_n^2 + a_1 b_1 \sigma_n^2 + b_1 \sigma_{n+1}^2$ . D'où

$$\begin{aligned} \sigma_{n+2}^2 &= \omega_0(1 + a_1) + a_1^2 X_n^2 + a_1 b_1(\omega_0(1 + b_1 + \dots + b_1^{n-2}) + a_1 \sum_{i=0}^{n-1} b_1^i X_{n-1-i}^2 + b_1^n \sigma_0^2) \\ &\quad + b_1(\omega_0(1 + b_1 + \dots + b_1^{n-1}) + a_1 \sum_{i=0}^n b_1^i X_{n-i}^2 + b_1^n \sigma_0^2) \end{aligned}$$

# Plan du cours

# Prédiction sans modèle

**Méthode 2** : On détermine  $\hat{X}_{n+h}$  sans modéliser  $(X_1, \dots, X_n)$

⇒ Influence décroissante (exponentielle) aux données lointaines

⇒ Filtrage exponentiel, moindres carrés pondérés

① Filtrage simple de Holt-Winters simple : si  $\beta_h \in [0, 1]$ , on cherche :

$$\hat{X}_{n+h} = \hat{a}_h = \underset{a_h \in \mathbb{R}}{\text{Argmin}} \sum_{j=0}^{n-1} \beta_h^{n-j} (X_j - a_h)^2$$

$$\Rightarrow \hat{a}_h = ({}^t J \Omega(\beta_h) J)^{-1} {}^t J \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\text{avec } \Omega(\beta) = \begin{pmatrix} \beta^{n-1} & 0 & 0 & \dots & 0 \\ 0 & \beta^{n-1} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad \text{et } J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

## Prédiction sans modèle (2)

- ② Filtrage de Holt-Winters de degré 1 : pour  $\beta_h \in [0, 1]$ , on cherche :

$$\hat{X}_{n+h} = \hat{a}_h(n+h) + \hat{b}_h \quad \text{où} \quad (\hat{a}_h, \hat{b}_h)^t = \text{Arg min}_{\hat{a}_h, \hat{b}_h \in \mathbb{R}} \sum_{j=0}^{n-1} \beta_h^{n-j} (X_j - (a_h j + b_h))^2$$
$$\implies \hat{a}_h = ({}^t J \Omega(\beta_h) J)^{-1} {}^t J \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \text{ avec } J = \begin{pmatrix} 1 & 1 \\ \vdots & \\ 1 & n \end{pmatrix}$$

- ③ Filtre de Holt-Winter double

$\implies$  Filtrage exponentiel sur les  $(X_t)_t$  de paramètre  $\beta_h$  et sur les accroissement  $(X_{t+1} - X_t)_t$  de paramètre  $\delta_h$ .

Si  $\delta_h = \beta_h$  revient au filtre de degré 1

## Prédiction sans modèle (3)

**Problème** : Comment choisir  $\beta_h$  ?

Utilisation d'une **validation croisée** particulière :

- Pour  $m$  grand, on choisit une grille  $B_m = \{0, 1/m, 2/m, \dots, 1\}$  pour  $\beta_h$
- Pour  $M$  grand, et  $n - M \leq t \leq n - h$ , pour chaque  $\beta_h$  choisi,

$\implies \hat{X}_{t+h}(\beta_h)$  calculé à partir de  $(X_1, \dots, X_t)$

- On choisit  $\hat{\beta}_h = \operatorname{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} (X_{t+h} - \hat{X}_{t+h}(\beta_h))^2$

**Remarque** :  $\beta_h \rightarrow 1$  même poids pour tous,  $\beta_h \rightarrow 0$  dernière valeur

# Plan du cours

# Value-at-Risk

**Problème** : Déterminer la perte maximale d'un actif (portefeuille), la value-at-risk, à un horizon de  $h$  jours et une confiance de  $1 - \alpha$  (risque  $\alpha$ ) ?

## Définition

Pour  $(I_t)$  l'évolution d'un actif ou d'un portefeuille,

$$\mathbb{P}(I_{t+h} - I_t \leq \text{VAR}_{t,h}(\alpha) \mid I_t, I_{t-1}, \dots) = \alpha$$

**Exemple** : Si  $I_{t+1} - I_t$  ARCH(1) et  $(\xi_t)_t$  bruit blanc  $\mathcal{N}(0, 1)$ , alors  $\mathbb{P}(\xi_t \sqrt{a_0 + a_1 (I_t - I_{t-1})^2} \leq \text{VAR}_{t,1}(\alpha) \mid I_t, I_{t-1}, \dots) = \alpha$

$$\implies \text{VAR}_{t,1}(\alpha) = \sqrt{a_0 + a_1 (I_t - I_{t-1})^2} q_\alpha \text{ où } q_\alpha \text{ quantile}$$

$$\implies \widehat{\text{VAR}}_{t,1}(\alpha) = \sqrt{\widehat{a}_0 + \widehat{a}_1 (I_t - I_{t-1})^2} q_\alpha$$

**Exercice** : Et pour  $\text{VAR}_{t,h}(\alpha)$  ?



## Value-at-Risk (2)

### Démonstration.

Avec  $X_t = I_t - I_{t-1}$ , on veut résoudre  $\mathbf{P}(X_{t+h} + \dots + X_{t+1} \leq \text{VAR}_{t,h}(\alpha) \mid X_t, X_{t-1}, \dots) = \alpha$ .  
Mais :

$$\begin{aligned}\mathbf{E}[X_{n+k} \mid X_t, X_{t-1}, \dots] &= \mathbf{E}[\sqrt{(a_0 + a_1 X_{n+k-1}^2) \mid X_t, X_{t-1}, \dots}] \mathbf{E}[\xi_{n+k} \mid X_t, X_{t-1}, \dots] = 0; \\ \text{var}[X_{n+k} \mid X_t, X_{t-1}, \dots] &= \mathbf{E}[\xi_{n+k}^2 \mid X_t, X_{t-1}, \dots] \mathbf{E}[(a_0 + a_1 X_{n+k-1}^2) \mid X_t, X_{t-1}, \dots] \\ &= a_0(1 + a_1 + \dots + a_1^{k-1}) + a_1^k X_n^2.\end{aligned}$$

$$\text{D'où } \begin{cases} \mathbf{E}[X_{t+h} + \dots + X_{t+1} \mid X_t, X_{t-1}, \dots] &= 0 \\ \text{var}[X_{t+h} + \dots + X_{t+1} \mid X_t, X_{t-1}, \dots] &= a_0 \sum_{k=0}^{h-1} (h-k) a_1^k + \frac{a_1(1-a_1^h)}{1-a_1} X_n^2 \end{cases} \text{ car} \\ \text{cov}(X_{t+i}, X_{t+j} \mid X_t, X_{t-1}, \dots) = 0.$$

Avec l'inégalité de Bienaymé-Tchebychev, pour tout  $\text{VAR}_{t,h}(\alpha) < 0$  si le bruit est symétrique,

$$\mathbf{P}(X_{t+h} + \dots + X_{t+1} \leq \text{VAR}_{t,h}(\alpha) \mid X_t, X_{t-1}, \dots) \leq \frac{a_0 \sum_{k=0}^{h-1} (h-k) a_1^k + \frac{a_1(1-a_1^h)}{1-a_1} X_n^2}{2(\text{VAR}_{t,h}(\alpha))^2}$$

$$\text{On posera donc } \widehat{\text{VAR}}_{t,h}(\alpha) = \left( \frac{2\alpha}{\widehat{a}_0 \sum_{k=0}^{h-1} (h-k) \widehat{a}_1^k + \frac{\widehat{a}_1(1-\widehat{a}_1^h)}{1-\widehat{a}_1} X_n^2} \right)^{1/2}.$$

