

Partie II: Processus stochastiques Linéaires: Modèles AR et MA

N. Laïb

EISTI

9 décembre 2018

- 1 Introduction
- 2 Processus Stochastique Linéaires
- 3 Processus auto-régressifs $AR(p)$
- 4 processus moyenne-mobile : $MA(q)$
- 5 Prévision avec $MA(q)$ et $AR(p)$

Définition 1

(i) Un processus $(X_t)_t$ est dit linéaire si il peut s'écrire :

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

où $\epsilon_i \hookrightarrow BB(0, \sigma^2)$, $\psi_0 = 1$ et $\sum_i |\psi_i| < \infty$

(ii) $(X_t)_t$ est dit linéaire et causale si il est linéaire avec $\psi_i = 0$ pour $i < 0$:

$$X_t = \mu + \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}, \quad (2)$$

Remarques.

- Un tel processus est appelé parfois moyenne mobile infinie, noté $MA(\infty)$.
- L'étude des séries non causales conduit à des résultats non intuitifs difficilement utilisables.
- Le processus (2) (avec $\mu = 0$) est strictement stationnaire du fait que les ϵ_t le sont

↓

$$\mathbb{E}(X_t) = 0, \quad \text{Var}(X_t) = \sigma_\epsilon^2 \left(1 + \sum_{i=1}^{\infty} \psi_i^2\right), \quad \gamma_k = \text{Cov}(X_t, X_{t-k}) = \sigma_\epsilon^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} \quad (3)$$

- Un processus X_t , centré, stationnaire d'ordre 2, peut s'écrire selon l'expression (2).

Proposition 1

Si $(X_t, t \in \mathbb{Z})$ est un processus stationnaire, et si $(a_i, i \in \mathbb{Z})$ est une suite de réels $\mu + \sum_{i \in \mathbb{Z}} |a_i| < +\infty$, alors, le processus (Y_t) défini par

$$Y_t = \mu + \sum_{i \in \mathbb{Z}} a_i X_{t-i}, \quad \text{pour tout } t \in \mathbb{Z},$$

est un processus stationnaire.

Définition 2

Soit $(X_t, t \in \mathbb{Z})$ un processus du second ordre, on appelle **innovation** du processus à la date t , la variable $X_t - X_t^*$, où X_t^* est la régression affine de X_t sur $(X_s, s \leq t-1)$.

- X_t^* est une **prévision optimale** de X_t fondée sur l'information $I_{t-1} = \{X_{t-1}, X_{t-2}, \dots\}$.
On note

$$X_t^* := \widehat{X}_{t-1}(1) = \mathbb{E}(X_t | I_{t-1}).$$

- L'innovation $X_t - X_t^*$ est la partie de X_t non corrélée au passé, c'est l'**erreur de prévision** sur X_t , lorsque la **prévision est faite à la date $t-1$** .
- On montre que, si $(X_t, t \in \mathbb{Z})$ est stationnaire, alors le processus des innovations,

$$X_t - X_t^* = X_t - \widehat{X}_{t-1}(1), \quad t \in \mathbb{Z},$$

constitue un **bruit blanc**.

Définition 3

$$\text{Si l'on peut écrire : } \epsilon_t = X_t - \sum_{j=1}^{\infty} \pi_j X_{t-j} \quad \text{avec} \quad \sum_{j=1}^{\infty} |\pi_j| < \infty, \quad (4)$$

on dit alors que le modèle (2) est inversible.

$$\text{On a alors : } X_t = \epsilon_t + \sum_{j=1}^{\infty} \pi_j X_{t-j} = \epsilon_t + \hat{X}_t,$$

- \hat{X}_t est la meilleure prévision de X_t connaissant le passé
- ϵ_t est l'erreur de prévision commise.
- ϵ_t est le processus d'innovation, il est indépendant de X_{t-1}, \dots

▷ Afin d'obtenir des modèles comportant un nombre fini de paramètres, on considère les modèles (2) et (4) tronqués :

processus autorégressif d'ordre p , $AR(p)$:
$$X_t = \epsilon_t + \sum_{j=1}^p \pi_j X_{t-j}$$

processus moyenne mobile d'ordre q , $MA(q)$:
$$X_t = \epsilon_t + \sum_{j=1}^q \psi_j X_{t-j}$$

▷ Le $AR(p)$ est un modèle que l'on retrouve dans de nombreux domaines :

- en démographie pour la dynamique des populations,
- en économétrie pour la prédiction d'indices
- en automatique via le filtrage de Kalman.

Définition 4

Le processus autorégressif le plus simple est le processus de Markov AR(1) défini par :

$$X_t = \phi X_{t-1} + \epsilon_t \quad \text{où} \quad \epsilon_t \hookrightarrow BB(0, \sigma^2) \quad (5)$$

En écrivant cette relation pour les indices $t-1, \dots, t-k+1$, on obtient

$$X_t = \epsilon_t + \phi \epsilon_{t-1} + \dots + \phi^{k-1} \epsilon_{t-k+1} + \phi^k \epsilon_{t-k}$$

et en faisant tendre k vers l'infini on obtient la représentation $MA(\infty)$ de ce modèle :

$$X_t = \sum_{j=0}^{\infty} \phi^j \epsilon_{t-j} \quad t \in \mathbb{Z} \quad \text{à condition que } |\phi| < 1, \quad \psi_j = \phi^j, \quad j \in \mathbb{N}.$$

- ▷ Si $\phi = 1$, le processus (5) n'est pas stationnaire. C'est une marche aléatoire
- ▷ Si $\phi > 1$ le processus (5) est explosif.

▷ L'équation (5) s'écrit :

$$(1 - \phi B)X_t = \epsilon_t$$

▷ Pour $|\phi| < 1$, nous avons

$$X_t = \frac{1}{1 - \phi B} \epsilon_t.$$

▷ Comme $|\phi| < 1$ on peut effectuer un développement en série entière :

$$\frac{1}{1 - \phi B} = 1 + \phi B + \phi^2 B^2 + \dots \text{ et } X_t = \epsilon_t + \phi \epsilon_{t-1} + \phi^2 \epsilon_{t-2} + \dots$$

▷ Représentation $MA(\infty)$ du processus $AR(12)$: $X_t = -0.7X_{t-12} + \epsilon_t$

$$\begin{aligned} \frac{1}{1 + 0.7B^{12}} &= 1 - 0.7B^{12} + 0.49B^{24} + (-0.7)^3 B^{36} + \dots \\ &= \sum_{j=0}^{\infty} (-0.7)^j \epsilon_{t-12-j} \end{aligned}$$

▷ Ce calcul peut se faire avec R à l'aide de `ARMAtoMA()`

- La fonction d'autocovariance est définie par récurrence :

$$\begin{aligned}\gamma_k &= \mathbb{E}(X_t X_{t-k}) = \phi \mathbb{E}(X_{t-1} X_{t-k}) + \mathbb{E}(\epsilon_t X_{t-k}) \\ &= \phi^k \gamma_0 = \frac{\phi^k}{1 - \rho^2} \sigma^2\end{aligned}\tag{6}$$

- On voit que les variables X_t et X_{t-k} sont dépendantes, avec une corrélation qui décroît exponentiellement vers 0 quand k tend vers l'infini.
- La fonction d'autocovariance

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi^k, \quad k = 1, 2, 3, \dots$$

- La fonction d'autocorrélation partielle se déduit de la définition

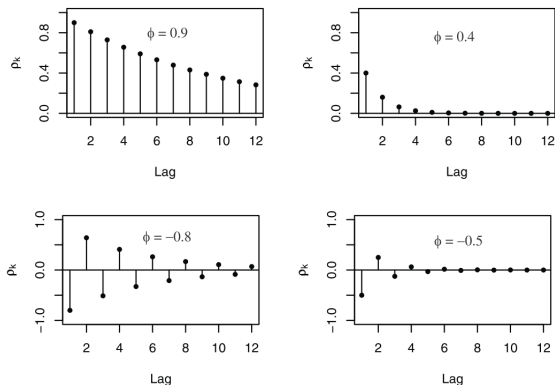
$$X_t = \phi X_{t-1} + \epsilon_t$$

puisque ϵ_t est indépendant de X_{t-1}, \dots, X_{t-k} :

$$\phi_{11} = \phi \quad \text{et} \quad \phi_{kk} = 0 \quad \text{pour} \quad k \geq 2.$$

- Puisque $|\phi| < 1$, la magnitude (l'ampleur) de la fonction d'autocorrélation décroît de façon exponentielle à mesure que le nombre de retards, k , augmente.
- Si $0 < \phi < 1$, toutes les corrélations sont positives
- Si $-1 < \phi < 0$, l'autocorrélation du décalage 1 est négative ($\rho_1 = \phi$) et les signes des autocorrélations successives alternent de positif à négatif, leurs amplitudes diminuant de manière exponentielle.

Figure: Fonctions d'autocorrélation pour plusieurs modèles AR (1)



- Si ϕ proche de ± 1 , la décroissance est assez lente, mais pour ϕ plus petit, la décroissance est assez rapide.
- Si ϕ près de ± 1 , la forte corrélation s'étendra sur de nombreux décalages et produira une série relativement lisse si ϕ est positive et une série très déchetée (série très en dents de scie) si ϕ est négatif.

Proposition 2

- (i) Si $\phi = \pm 1$, le processus (X_t) n'est pas stationnaire.
- (ii) Si $|\phi| \neq 1, \exists$ un unique processus stationnaire tel que $X_t - \phi X_{t-1} = \epsilon_t \quad \forall t \in \mathbb{Z}$
- (iii) Si $|\phi| < 1$ alors on peut inverser le polynôme, et

$$X_t = (1 - \phi B)^{-1} \epsilon_t = \sum_{i=1}^{\infty} \phi^i \epsilon_{t+i} \quad (\text{en fonction du future de } (\epsilon_t)). \quad (7)$$

- (iv) les fonctions d'auto-covariance et d'auto-corrélation sont respectivement

$$\gamma(h) = \sigma^2 \phi^h \frac{1}{1 - \phi^2} \quad \text{et} \quad \rho(h) = \phi^h \quad (\text{elles tendent vers zéro à vitesse exponentielle}),$$

- (v) la fonction d'auto -corrélation partielle est nulle pour tout retard ≥ 2 ,

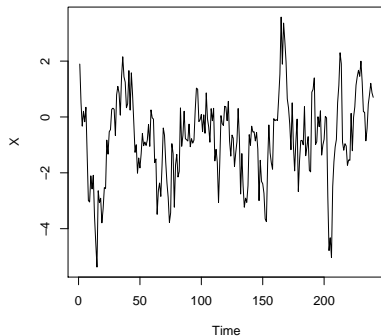
$$r(1) = \rho(1) \quad \text{et} \quad r(h) = 0, \quad \forall h \geq 2$$

Visualisation de la fonction d'autocorrélation d'un AR(1)

On peut visualiser cette autocorrélation sur les nuages de points $(X_{t-1}, X_t), (X_{t-2}, X_t)$, etc, pour un processus $AR(1)$ avec $\phi > 0$.

```
X=arima.sim(n = 240, list(ar = 0.8),sd = 1)  
plot(X)
```

Figure: Graphe de AR(1)



Un processus AR(1) : $X_t = \phi X_{t-1} + \epsilon_t$ sera :

- auto-corrélé positivement si $0 < \phi < 1$,
- auto-corrélé négativement si $-1 < \phi < 0$.
- Cette série oscille autour de 0, en s'en écartant suivant la valeur ϵ_t (si $-1 < \phi < +1$)
- Si $\phi = +1$, on obtient une marche aléatoire,
- si $\phi > +1$ ou $\phi < -1$ le processus n'est pas stationnaire, et on obtient un modèle qui explosera (à moyen terme).
- La valeur ϕ , dans le cas où le processus est **stationnaire**, est la corrélation entre deux dates consécutives $\phi = \text{corr}(X_t, X_{t-1})$

$$AR(1) : \left\{ \begin{array}{l} \text{Fonction d'autocorrélation} \left\{ \begin{array}{l} \phi > 0 \quad \text{décroissance exponentielle} \\ \phi < 0 \quad \text{sinusoïde amortie} \end{array} \right. \\ \text{Fonction d'autocorrélation partielle} \left\{ \begin{array}{l} \text{première non nulle (signe = signe de } \phi) \\ \text{toutes nulles après} \end{array} \right. \end{array} \right.$$

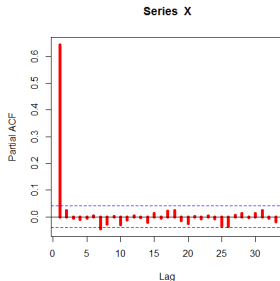
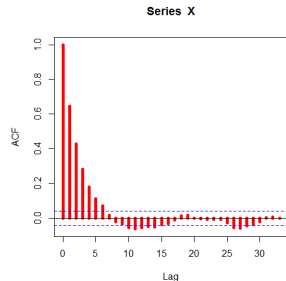
Exemple. Considérons un processus AR(1) stationnaire avec $\phi_1 = 0.6$.

```
X=arima.sim(n = 2400, list(ar = 0.6),sd = 1)
```

```
plot(X)
```

```
plot(acf(X),lwd=5,col="red")
```

```
plot(pacf(X),lwd=5,col="red")
```



Attention à la constante.

Par défaut, les modèles AR ne sont pas nécessairement centrés. La forme générale serait (pour un AR(1) par exemple)

$$X_t = c + \phi X_{t-1} + \epsilon_t$$

Si la série est stationnaire, de moyenne μ , alors μ devrait être solution de

$$\mu = c + \mu\phi \quad \Rightarrow \quad \mu = \frac{c}{1-\phi}.$$

Simulons un processus AR(1) de moyenne 2,

```
> X=arima.sim(list(order=c(1,0,0),ar=1/3),n=1000)+2
> mean(X)
[1] 1.931767
```

Le modèle $AR(1)$ est très simple puisque son évolution dépend seulement de sa valeur au temps précédent. Une généralisation consiste à introduire des temps supplémentaires.

Définition 5

$\{X_t\}$ est processus autoregressif d'ordre p , noté $AR(p)$, si il est stationnaire vérifiant une relation du type

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \epsilon_t \quad \forall t \in \mathbb{Z}, \quad (8)$$

où les ϕ_i sont des réels, avec $\phi_p \neq 0$, et (ϵ_t) est un bruit blanc de variance σ^2 .
L'équation (8) est équivalente à l'écriture :

$$\Phi(B)X_t = \epsilon_t \quad \text{où} \quad \Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$$

Remarque. En généralité, un processus $AR(p)$ peut être défini par la formule

$$X_t = c + \sum_{i=1}^p \phi_i X_{t-i} + \epsilon_t \quad \Rightarrow \quad \Phi(B)X_t = c + \epsilon_t \quad (9)$$

où c est un terme constant.

▷ Comme X_t est stationnaire,

$$\mu = E(X_t) = \sum_{i=1}^p \phi_i E(X_{t-i}) = c + \mu \sum_{i=1}^p \phi_i$$

Ce qui donne

$$\mu = \frac{c}{1 - \sum_{i=1}^p \phi_i} = \frac{c}{\Phi(1)}$$

On peut se ramener alors à (8) en posant :

$$Y_t = X_t - \mu \quad \text{où} \quad \mu = \frac{c}{\Phi(1)}.$$

▷ A quelle condition le processus $AR(p)$ admet une représentation $MA(\infty)$?

▷ Supposons $p=2$ et notons

$$P(z) = 1 - \phi_1 z - \phi_2 z^2 \quad \text{le polynôme caractéristique associé.}$$

Soient s_1 et s_2 ses racines réelles ou complexes \Rightarrow :

$$P(z) = (1 - z/s_1)(1 - z/s_2) \quad (10)$$

▷ Le développement en série entière de $P(z) = 1 - \phi_1 z - \phi_2 z^2$ est possible si chaque terme de (10) est inversible

⇔ les racines de $P(Z)$ sont strictement supérieures à 1.

D'après un résultat sur les polynômes (X_t) admet une représentation $MA(\infty)$, i.e.

$$X_t = \sum_{j=0}^{\infty} a_j \epsilon_{t-j} \quad \text{avec} \quad a_0 = 1, a_j \in \mathbb{R}, \quad \sum_{j=0}^{\infty} |a_j| < \infty$$

On sait que $\Phi(B)(X_t) = \epsilon_t \Rightarrow X_t = \Phi^{-1}(B)(\epsilon_t) \Rightarrow X_t$ stationnaire.

Proposition 3

Le processus autorégressif d'ordre p

$$X_t = \frac{1}{1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p} \epsilon_t \quad \epsilon_t \hookrightarrow BB(0, \sigma^2) \quad (11)$$

ou

$$X_t = \mu + \frac{1}{1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p} \epsilon_t \quad (\text{cas général}) \quad (12)$$

admet une représentation $MA(\infty)$ si les racines de l'équation :

$$1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = 0$$

sont strictement supérieures à 1 en module ; (11) est alors stationnaire ; c'est un $AR(p)$.

- La formule (12) sépare le niveau moyen μ , de l'erreur qui est une $AR(p)$ stationnaire.
- μ peut être une fonction du temps.
- Dans ce cas, la série est la somme d'une tendance déterministe et d'une erreur stationnaire.

▷ La fonction d'auto-corrélation vérifie l'équation de récurrence linéaire homogène d'ordre p suivante, dont les coefficients sont liés à ceux du polynôme autorégressif :

$$\rho(h) - \sum_{i=1}^p \phi_i \rho(h-i) = 0 \quad h > 0 \quad \text{où} \quad \rho(h) = \frac{\gamma(h)}{\gamma(0)}, h \in \mathbb{Z} \quad (13)$$

▷ En tenant compte de la parité de $\rho(h)$, on obtient les équations de Yule-Walker :

$$\begin{bmatrix} \rho(1) \\ \vdots \\ \rho(p) \end{bmatrix} = \begin{bmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(p-1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(p-2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho(p-1) & \rho(p-2) & \dots & \rho(1) & 1 \end{bmatrix} \begin{bmatrix} \phi(1) \\ \vdots \\ \phi(p) \end{bmatrix}$$

▷ Par inversion, il est possible d'exprimer les ϕ_i en fonction des $\rho(h)$.

$$\forall t, \quad X_t = c + \sum_{i=1}^p \phi_i X_{t-i} + \epsilon_t$$

où $p \geq 1$, $(\phi_i)_{1 \leq i \leq p}$ réels, et (ϵ_t) bruit blanc gaussien.

$$\Phi(B)X_t = c + \epsilon_t \quad \text{où} \quad \Phi(B)(X_t - \mu) = \epsilon_t \quad \text{et} \quad \Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$$

et

$$(1 - \phi_1 - \dots - \phi_p)\mu = c$$

Propriété. Un processus AR(p) est stationnaire si Φ n'a pas de racine unité.

Corrélogrammes d'un AR(p) stationnaire :

- ACF : $\rho(h)$ décroît de manière exponentielle vers 0
- PACF : $\forall h > p, r(h) = 0$

Intérêt : permet l'identification d'un AR(p).

Considérons le processus $AR(2)$

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \epsilon_t$$

Alors la fonction d'auto-corrélation vérifie l'équation de récurrence

$$\rho(h) = a_1 \rho(h-1) + a_2 \rho(h-2), \quad \forall h > 0.$$

Cette équation définit une suite récurrente linéaire d'ordre 2. Elle est donc entièrement déterminée par les racines du polynôme caractéristique ainsi que par les deux premiers termes de la suite. Or on sait que $\rho(0) = 1$, il reste donc à calculer $\rho(1)$.

$$\begin{aligned} \gamma(1) &= \text{cov}(X_t, X_{t-1}) = \mathbb{E}(X_t X_{t-1}) \\ &= a_1 \mathbb{E}(X_{t-1}^2) + a_2 \mathbb{E}(X_{t-1} X_t) + 0 \\ &= a_1 \gamma(0) + a_2 \gamma(2) \\ \Leftrightarrow \gamma(1) &= \frac{a_1 \sigma^2}{1 - a_2} \\ \Leftrightarrow \rho(1) &= \frac{\gamma(1)}{\sigma^2} = \frac{a_1}{1 - a_2} \end{aligned}$$

Appliquons l'exemple précédent au modèle AR(2) suivant,

$$X_t = X_{t-1} - \frac{1}{2}X_{t-2} + \varepsilon_t$$

Les racines du polynôme caractéristique sont complexes conjuguées telles que $z_1 = 1 - i = \sqrt{2}\exp(-i\pi/4)$, donc la fonction d'auto-corrélation est de la forme

$$\rho(h) = \sqrt{2}^h \left(K_1 \cos\left(h\frac{\pi}{4}\right) + K_2 \sin\left(h\frac{\pi}{4}\right) \right)$$

Or

$$\rho(0) = K_1 \quad \text{et} \quad \rho(1) = K_1 + K_2 = \frac{1}{1 - (-1/2)} \Rightarrow K_2 = -1/3.$$

Le comportement d'un processus

$$AR(2) : X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$$

- dépendra fortement des racines de son équation caractéristique $1 - \phi_1 z - \phi_2 z^2 = 0$.
- Si l'équation caractéristique a deux racines complexes conjuguées :

$r \exp(\pm i\theta)$ pour $r < 1$. Alors

- le processus est stationnaire (et oscille autour de 0, sans exploser)
- Le processus est aussi quasi-cyclique, de fréquence θ , avec un bruit aléatoire.

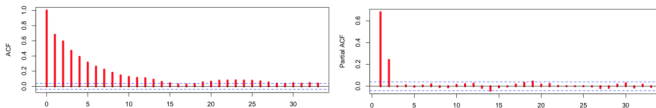
- AR(2) :

{	Fonction d'autocorrélation	{	décroissance exponentielle ou sinuséide amorti
	Fonction d'autocorrélation partielle		deux premières non nulles toutes nulles après

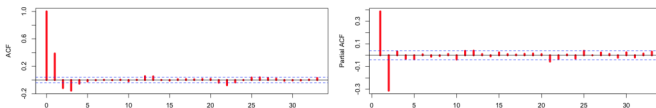
figure 28, Graphe des Fonctions ACF et PACF de AR(2).

Figure 3: Graphe des Fonctions ACF et PACF de AR(2)

```
> X=arima.sim(n = 2400, list(ar = c(0.6,0.4)),sd = 1)
> plot(acf(X),lwd=5,col="red")
> plot(pacf(X),lwd=5,col="red")
```



```
> X=arima.sim(n = 2400, list(ar = c(0.6,-0.4)),sd = 1)
> plot(acf(X),lwd=5,col="red")
> plot(pacf(X),lwd=5,col="red")
```



- ▷ L'algorithme de Durbin-Levinson, programmé dans la fonction `PacfDL()` de `FitAR`
- permet de calculer de façon itérative la PACF (théorique) à partir de l'ACF,
 - appliqué à l'**AFC empirique** permet de calculer les estimations $\hat{\phi}_{k,k}$, $k = 1, 2, \dots$
 - D'un point de vue pratique une **série suit un $AR(p)$** si les

$$\hat{\phi}_{k,k} \approx 0, \quad k > p, \quad \text{précisément :}$$

Proposition 4

Si (X_t) est un $AR(p)$, alors :

- (a) $\hat{\phi}_{p,p}$ converge vers $\phi_{p,p}$ quand $T \rightarrow \infty$,
- (b) $\hat{\phi}_{k,k}$, $\forall k > p$ converge vers 0 quand $T \rightarrow \infty$,
- (c) $Var(\hat{\phi}_{k,k}) \approx 1/T \quad \forall k > p$

- Si la PACF empirique d'une série n'est plus significativement différente de zéro à partir d'un certain ordre k , on peut lui ajuster un modèle $AR(k-1)$.

Exemple

Calculons à l'aide de l'algorithme de Durbin-Levinson la PACF d'un AR(2)

$$X_t = -0.7X_{t-1} + 0.2X_{t-2} + \varepsilon_t \quad (14)$$

puis les coefficients a_T jusqu'à $T = 4$ du système de Yule-Walker.

- Normalement on doit trouver des coefficients a_k nuls, pour $k > 2$.
- On calcule d'abord la fonction d'autocovariance théorique, g , à l'aide de `TacvfARMA()`
- La fonction `PacfDL()` a comme argument la fonction d'autocorrélation, $g/g[1]$

packages utilisés

require(caschrono)

require(FitARMA)

require(xtable)

require(polynom)

require(nlme)

require(dynlm)

$g = \text{TacvfARMA}(\text{phi}=\text{c}(-.7,.2),\text{lag.max}=4) :$ fonction d'autocovariance théorique
($a=\text{PacfDL}(g/g[1], \text{LinearPredictor} = \text{TRUE})$)

\$Pacf

```
[1] -8.750000e-01  2.000000e-01  -1.233581e-16  -3.563679e-16
```

\$ARCoefficients

```
[1] -7.000000e-01  2.000000e-01  -3.728156e-16  -3.563679e-16
```

\$ResidualVariance

```
[1] 0.225
```

```
g[1] - t(as.matrix(a$ARCoefficients))%*%as.matrix(g[-1])
```

```
  [,1]
```

```
[1,]  1
```

- **\$Pacf** : est le vecteur des $\phi_{k,k}$, effectivement nuls à partir de $k = 3$.
- **\$ARCoefficients** : est le vecteur (a_1, \dots, a_4) , avec $(a_1, a_2) = (\phi_1, \phi_2)$ et $a_k = 0$ à partir de $k = 3$.
- **\$ResidualVariance** : calcule la variance de l'erreur de prédiction à l'horizon 1. Elle vaut 1, comme on s'y attendait, car le bruit utilisé pour calculer la fonction d'autocovariance par `TacvfARMA()` est supposé de variance 1.

Exercice. Simuler une trajectoire de 200 valeurs d'un processus autorégressif obéissant à (14) et calculer la PCF empirique jusqu'au retard 4. Comparer avec l'exemple précédent.

Définition 6

$\{X_t\}$ est un processus moyenne mobile d'ordre q ($MA(q)$) si :

$$X_t = \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} \quad \epsilon_t \hookrightarrow BB(0, \sigma^2), \quad (15)$$

avec $\theta_q \neq 0$.

(15) peut s'écrire sous la forme

$$\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q \quad \text{ou} \quad X_t = \Theta(B) \epsilon_t$$

- ▷ Un $MA(q)$ est toujours stationnaire quelles que soient les valeurs des θ , il est de moyenne 0.
- ▷ Les processus $MA(\infty)$ sont stationnaires si et seulement si $\sum_{i=1}^{\infty} \theta_i^2 < \infty$.
- ▷ **Inversibilité** : on souhaite exprimer ce processus en fonction de son passé (observé) et pas seulement en fonction du bruit passé non observé.

▷ Cas de $MA(1)$:

$$X_t = \epsilon_t + \theta\epsilon_{t-1} = (1 + \theta B)\epsilon_t, \quad \epsilon_t \hookrightarrow BB(0, \sigma^2)$$

▷ Si $|\theta| < 1$, on peut développer $(1 + \theta B)^{-1}$ en série :

$$(1 + \theta B)^{-1} = 1 - \theta B + \theta^2 B^2 - \theta^3 B^3 + \dots$$

▷ Donc $MA(1)$ peut s'écrire comme une autorégression infinie :

$$X_t = \epsilon_t + \theta X_{t-1} - \theta^2 X_{t-2} + \dots$$

on dit qu'il est inversible.

- ▷ La condition d'inversibilité d'un $MA(1)$ est parallèle à la condition de représentation d'un $AR(1)$.
- ▷ Un $MA(q)$ est dit inversible si l'on peut le représenter comme une autoregression infinie.

Proposition 5

Un $MA(q)$ est inversible si les racines de

$$1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q = 0$$

sont, en module strictement supérieures à 1.

Proposition 6

$$\begin{aligned}
 \gamma(h) &= \mathbb{E}(X_t X_{t-h}) \\
 &= \mathbb{E}\left[\left[\epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q}\right] \left[\epsilon_{t-h} + \theta_1 \epsilon_{t-h-1} + \theta_2 \epsilon_{t-h-2} + \dots + \theta_q \epsilon_{t-h-q}\right]\right] \\
 &= \begin{cases} \left[\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_q \theta_{q-h}\right] \sigma^2 & \text{si } 1 \leq h \leq q \\ 0 & \text{si } h > q, \end{cases}
 \end{aligned}$$

avec, pour $h = 0$, la relation

$$\gamma(0) = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma^2$$

Cette dernière relation peut se réécrire

$$\gamma(k) = \sigma^2 \sum_{j=0}^q \theta_j \theta_{j+k} \quad \text{avec } \theta_0 = 1$$

D'où la fonction d'autocorrélation

$$\rho(h) = \frac{\theta_h + \theta_{h+1} \theta_1 + \dots + \theta_q \theta_{q-h}}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} \quad \text{si } 1 \leq h \leq q,$$

▷ On note en particulier que $\gamma(q) = \sigma^2\theta_q \neq 0$, alors que $\gamma(q+1) = 0$.

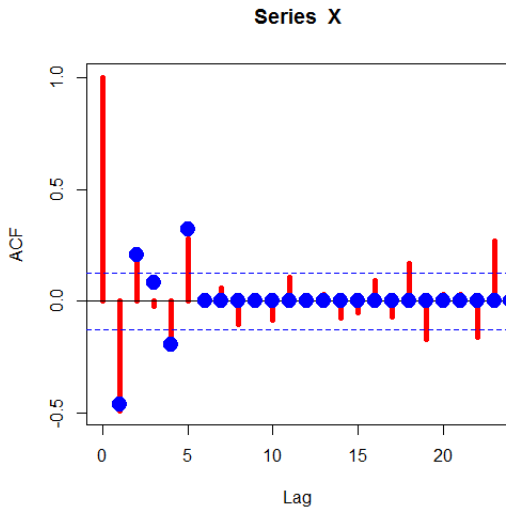
Cette permet d'estimer l'ordre de processus MA.

▷ On peut calculer les autocorrélations à la main

```
rho=function(h,theta){
theta=c(1,theta)
q=length(theta)
if(h>=q) ACF=0 else{sn=0;sd=0
for(i in 1:(q-h)) sn=sn+theta[i]*theta[i+h]
for(i in 1:(q)) sd=sd+theta[i]^2
ACF=sn/sd}
return(ACF)}
```

```
THETA=c(-.6,.7,-.4,0,.9)
X=arima.sim(list(ma = THETA),n=240)
acf(X,col="red",lwd=5)
```

```
R=function(h) rho(h,THETA)
points(1:24,Vectorize(R)(1:24),pch=19,cex=2,col="blue")
```



Un processus MA (1) est :

$$X_t = \epsilon_t + \theta \epsilon_{t-1}, \quad \text{pour tout } t \in \mathbb{Z}, \quad \epsilon_t \hookrightarrow BB(0, \sigma^2).$$

- Les autocorrélations sont données par

$$\rho(1) = \frac{\theta}{1 + \theta^2}, \quad \text{et } \rho(h) = 0, \quad \text{pour } h \geq 2.$$

- On montre que :

$$|\rho(1)| \leq 1/2$$

les modèles MA (1) ne peuvent avoir de fortes autocorrélations à l'ordre 1.

- L'autocorrélation partielle à l'ordre h est donnée par

$$r(h) = \frac{(-1)^h \theta^h (\theta^2 - 1)}{1 - \theta^{2(h+1)}},$$

▷ On peut visualiser l'autocorrélation sur les nuages de points $(X_{t-1}, X_t), (X_{t-2}, X_t)$, etc, pour un processus $MA(1)$ avec $\theta > 0$.

```
X=arima.sim(n = 240, list(ma = 0.8),sd = 1)
plot(X)
n=240;h=1
plot(X[1:(n-h)],X[(1+h):n])
library(ellipse)
lines(ellipse(.8/(1+.8^2)),type='l',col="red")
```

▷ **A retenir.**

$$MA(1): \begin{cases} \text{Fonction d'autocorrélation} \begin{cases} \text{première non nulle (signe = signe de } \theta) \\ \text{toutes nulles après} \end{cases} \\ \text{Fonction d'autocorrélation partielle} \begin{cases} \theta > 0 \text{ décroissance exponentielle} \\ \theta < 0 \text{ sinusoïde amortie} \end{cases} \end{cases}$$

Le processus MA (2) est : $X_t = \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2}$.

▷ La fonction d'autocorrélation est donnée par

$$\rho(h) = \begin{cases} \frac{\theta_1(1+\theta_2)}{1+\theta_1^2+\theta_2^2} & h=1 \\ \frac{\theta_2}{1+\theta_1^2+\theta_2^2} & h=2 \\ 0 & h \geq 3 \end{cases}$$

▷ Soit X un MA(q) de polynôme Θ et de bruit blanc d'innovation ϵ . La relation qui le lie à son bruit blanc d'innovation ϵ :

$$X_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}.$$

On alors :

$$\hat{X}_t(1) = -\theta_1 \epsilon_t - \dots - \theta_q \epsilon_{t-(q-1)}$$

$$\hat{X}_t(2) = -\theta_2 \epsilon_t - \dots - \theta_q \epsilon_{t-(q-2)}$$

⋮

$$\hat{X}_t(q) = -\theta_q \epsilon_t$$

$$\hat{X}_t(j) = 0, \quad j > q.$$

Il suffit ensuite d'exprimer explicitement les ϵ_t en fonction des X_t .

- ▷ La meilleure Prévision à l'horizon (BLB) $h = 1$ est

$$\mathbb{E}(X_{T+1}|X_T) = \mu + \phi X_T$$

- ▷ Cela peut être itérée pour obtenir la prévision à l'horizon k étapes ($k \geq 2$)

$$\mathbb{E}(X_{T+k}|X_T) = \mu \sum_{i=0}^{k-1} \phi^i + \phi_k X_T$$

- ▷ Lorsque $k \rightarrow \infty$, la prévision se rapproche de la valeur stationnaire $\mu/(1-\phi)$.
- ▷ L'erreur de prévision est donnée par

$$\mathbb{E}(Y_{T+k}|Y_T) - Y_{T+k} = \sum_{i=0}^{k-1} \phi^{k-1-i} \varepsilon_{T+i+1}$$

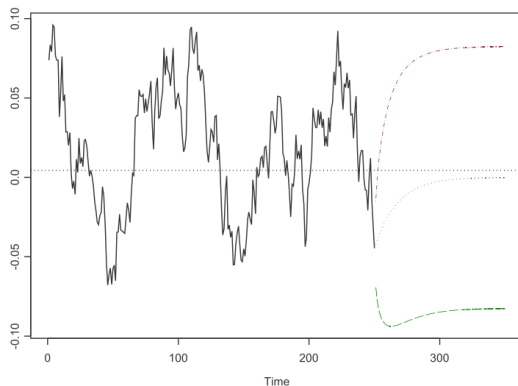
- ▷ La variance de l'erreur de prévision en k -pas est donnée par

$$\text{Var}[\mathbb{E}(Y_{T+k}|Y_T) - Y_{T+k}] = \sigma^2 \sum_{i=0}^{k-1} \phi^i = \sigma^2 \frac{1 - \alpha^{2k}}{1 - \phi^2}$$

Ainsi, lorsque $k \rightarrow \infty$ la variance augmente jusqu'au niveau stationnaire

Prévision avec AR(1)-modèle (suite)

Figure: Prévision de 100 jours avenir d'un AR(1) simulé avec $\mu = 0.00022$ et $\sigma = 0.013$. Les courbes en pointillés représentent des intervalles de confiance. La ligne horizontale représente le niveau stationnaire du processus.



▷ à mesure que nous avançons dans le futur, les prévisions approchent du niveau stationnaire, tandis que l'intervalle de confiance s'élargit.

$$AR(p): \quad X_t = \phi_0 + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \epsilon_t, \quad \epsilon_t \hookrightarrow BB(0, \sigma^2)$$

▷ A l'horizon $h=1$:

$$X_{t+1} = \phi_0 + \phi_1 X_t + \dots + \phi_p X_{t+1-p} + \epsilon_{t+1}, \quad \epsilon_t \hookrightarrow BB(0, \sigma^2)$$

▷ BLB de X_t à l'horizon $h=1$:

$$\hat{X}_t(1) := X_{t+1|t} = EL(X_{t+1} | X_t, X_{t-1}, \dots, X_1) = \phi_0 + \phi_1 X_t + \dots + \phi_p X_{t+1-p}$$

▷ l'erreur de prédiction associée : $e_1(t) = X_{t+1} - X_{t+1|t} = \epsilon_{t+1}$

▷ L'erreur quadratique moyenne de prévision est égale à :

σ_ϵ^2 = la variance de X_{t+1} conditionnellement au passé de la série.

▷ Le bruit blanc ϵ_{t+1} peut être interprété comme la

correction à la prédiction mécanique de X_{t+1} par les valeurs les plus récemment observées.

- ▷ **Prévision à l'horizon 2.** La prévision de

$$X_{t+2} = \phi_0 + \phi_1 X_t + \dots + \phi_p X_{t+2-p} + \epsilon_{t+2}, \quad \epsilon_t \hookrightarrow BB(0, \sigma^2)$$

connaissant le passé = espérance conditionnelle linéaire de X_{t+2} par rapport à son passé.

- ▷ que l'on prédit par

$$\hat{X}(2) = X_{t+2|t} = \phi_0 + \phi_1 \hat{X}_t(1) + \phi_2 X_t + \dots + \phi_p X_{t+2-p}.$$

- ▷ l'erreur de prédiction associée

$$e_2(t) = X_{t+2} - X_{t+2|t} = \epsilon_{t+2} + \phi_1 e_t(1) = \epsilon_{t+2} + \phi_1 \epsilon_{t+1}$$



$$\mathbb{E}(e_2(t)) = 0 \quad \text{et} \quad \text{Var}(e_2(t)) = \sigma_\epsilon^2 (1 + \phi_1^2)$$

- ▷ La variance de l'erreur de prévision augmente avec l'horizon de prévision.
- ▷ Dans la pratique, on remplace les ϕ par leurs estimations et on ne tient pas compte de la variabilité de ces dernières pour établir la prévision, qu'on note alors

$$\hat{X}_{t+h|t}$$

▷ Soit X un AR(p) de bruit blanc d'innovation ϵ et de polynôme canonique $\Phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p$.

▷ Les prévisions optimales $\hat{X}_t(h)$ pour $h > 0$ sont des combinaisons linéaires de $X_t, X_{t-1}, \dots, X_{t-p+1}$

$$\hat{X}_t(h) = \sum_{j=0}^{p-1} a_j(h) X_{t-j}$$

qui s'obtiennent par récurrence selon

$$\hat{X}_t(h) = \sum_{k=1}^p \phi_k \hat{X}_t(h-k)$$

avec les conditions initiales $\hat{X}_t(-j) = X_{t-j}$ pour $j \geq 0$.