

Les processus AR et MA

MAP-STA2 : Séries chronologiques

Yannig Goude yannig.goude@edf.fr

2020-2021

Contents

Théorème de représentation de Wold	1
Représentation spectrale	2
Opérateurs retard et avance	5
Définition	5
Inversion de polynômes en L	6
Processus auto-régressif (AR)	7
Processus moyennes mobiles (MA)	10
Processus autorégressifs moyennes mobiles (ARMA)	13
Identification et estimation de modèles ARMA	15

Théorème de représentation de Wold

Ce théorème motive les résultats présentés dans ce chapitre. Nous n'en ferons pas la démonstration pour des raisons de temps.

théorème soit $(X_t)_{t \in \mathbf{Z}}$ un processus centré et stationnaire du second ordre, alors on a la décomposition suivante:

$$X_t = \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t-j} + C_t, \quad t \in \mathbf{Z}$$

où

- $\psi_0 = 1$ et $\sum_{j=0}^{+\infty} |\psi_j| < +\infty$
- ε_t est le processus d'innovation de X_t
- C_t est un processus déterministe

les moyennes mobiles s'avèrent donc intéressantes pour modéliser des processus stationnaires. La question reste alors l'estimation des coefficients ψ . Nous verrons par la suite que certaines familles de modèles permettent une représentation parcimonieuse de cette moyenne mobile.

Représentation spectrale

Nous avons jusqu'à présent étudié les processus stationnaires du second ordre dans leur représentation temporelle. On peut également s'intéresser à leur représentation dans le domaine des fréquences, approche largement considérée en traitement du signal, en décomposant le processus en composantes périodiques aléatoires.

Considérons, pour simplifier, le cas des moyennes mobiles infinies.

$$X_t = \sum_{i \in \mathbf{Z}} \theta_i \varepsilon_{t-i}, \quad \sum_{i \in \mathbf{Z}} |\theta_i| < +\infty$$

et ε_t est un bruit blanc de variance σ^2 .

Ce processus a pour fonction de covariance:

$$\gamma(h) = \sigma^2 \sum_{i \in \mathbf{Z}} \theta_i \theta_{i-h}$$

la série $\gamma(h)$ est absolument sommable (à vérifier).

la densité spectrale du processus est la fonction définie par:

$$f(\omega) = \frac{1}{2\pi} \sum_{h \in \mathbf{Z}} \gamma(h) e^{i\omega h}, \quad \forall \omega \in \mathbf{R}$$

cette fonction existe car la série $(\gamma(h) e^{i\omega h})_{h \in \mathbf{Z}}$ est absolument sommable.

D'autre part c'est une fonction réelle car γ est paire:

$$\begin{aligned} f(\omega) &= \frac{1}{2\pi} \left[\gamma(0) + \sum_{h=1}^{+\infty} \gamma(h) (e^{i\omega h} + e^{-i\omega h}) \right] \\ f(\omega) &= \frac{\gamma(0)}{2\pi} + \frac{1}{\pi} \sum_{h=1}^{+\infty} \gamma(h) \cos(\omega h) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} \gamma(h) \cos(\omega h) \end{aligned}$$

propriété connaître la fonction d'auto-covariance d'un processus est équivalent à connaître sa densité spectrale et on a:

$$\gamma(h) = \int_{-\pi}^{\pi} f(\omega) \cos(\omega h) d\omega = \int_{-\pi}^{\pi} f(\omega) e^{i\omega h} d\omega$$

preuve à faire

propriété soient $(X_t)_{t \in \mathbf{Z}}$ un processus moyenne mobile infinie $X_t = \sum_{i \in \mathbf{Z}} \theta_i \varepsilon_{t-i}$ et $(Y_t)_{t \in \mathbf{Z}}$ le processus défini par $Y_t = \sum_{j \in \mathbf{Z}} \theta_j X_{t-j}$ avec $\sum_{i \in \mathbf{Z}} |\theta_i| < +\infty$ alors on a la relation suivante entre les deux densités spectrales de X et Y :

$$f_y(\omega) = f_x(\omega) \left| \sum_{j \in \mathbf{Z}} \theta_j e^{i\omega j} \right|^2$$

preuve à faire

exemples

- bruit blanc: la densité spectrale d'un bruit blanc de variance σ^2 est constante et vaut:

$$f(\omega) = \frac{1}{2\pi} \sum_{h \in \mathbf{Z}} \gamma(h) e^{i\omega h} = \frac{\gamma(0)}{2\pi} = \frac{\sigma^2}{2\pi}$$

inversement, si X est un processus stationnaire de densité spectrale constante $f(\omega) = c$ on a:

$$\gamma(h) = \int_{-\pi}^{\pi} c \cos(\omega h) d\omega = 0, \text{ si } h \neq 0$$

ce qui est la définition d'un bruit blanc faible.

- moyenne mobile d'ordre 1: MA(1). Soit $X_t = \varepsilon_t + \theta \varepsilon_{t-1}$:

$$f(\omega) = \frac{\sigma^2}{2\pi} (1 + \theta^2 + 2\theta \cos(\omega))$$

- moyenne mobile infinie: MA(∞).

soit $X_t = \sum_{j \in \mathbf{Z}} \theta^j \varepsilon_{t-j}$, on a alors:

$$\gamma(h) = \sigma^2 \sum_{j \in \mathbf{Z}} \theta^{2j+|h|}$$

et

$$f(\omega) = \frac{\sigma^2}{2\pi} |\Theta(e^{i\omega})|^2$$

avec $\Theta(z) = \sum_{j=0}^{+\infty} \theta^j z^j$, $z \in \mathbf{C}$.

d'ou:

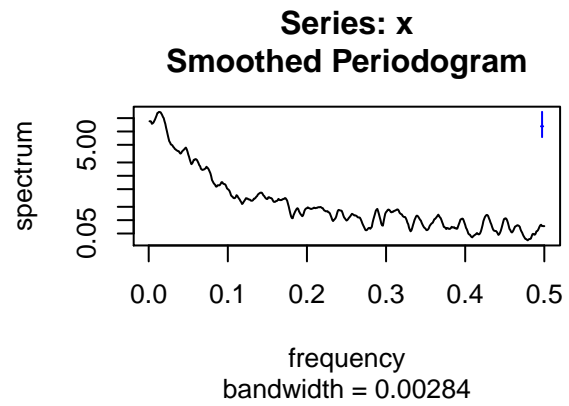
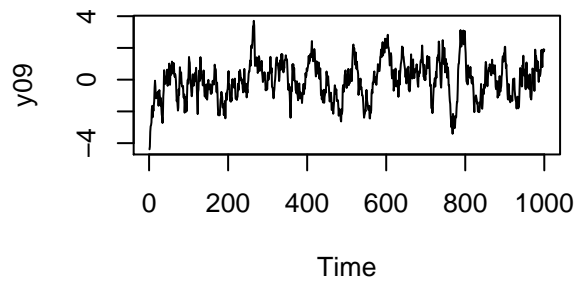
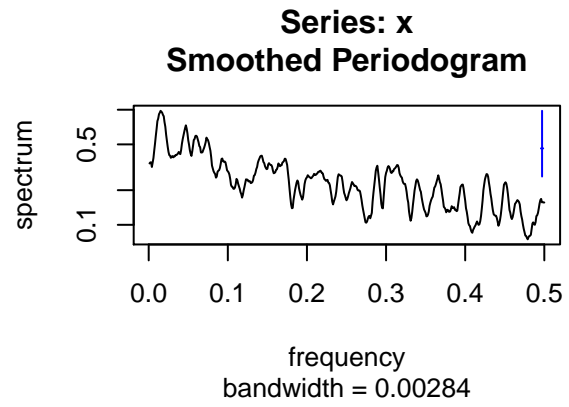
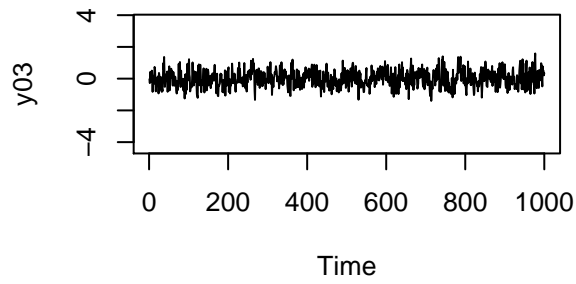
$$f(\omega) = \frac{\sigma^2}{2\pi} \frac{1}{|1 - \theta e^{i\omega}|} = \frac{\sigma^2}{2\pi} \frac{1}{1 + \theta^2 - 2\theta \cos(\omega)}$$

ci-dessous un exemple de densités spectrales obtenues pour deux processus AR(1), de coefficients $\theta = 0.3$ et $\theta = 0.9$. Le cas $\theta = 0.9$ fait clairement apparaître une prédominance des basses fréquences.

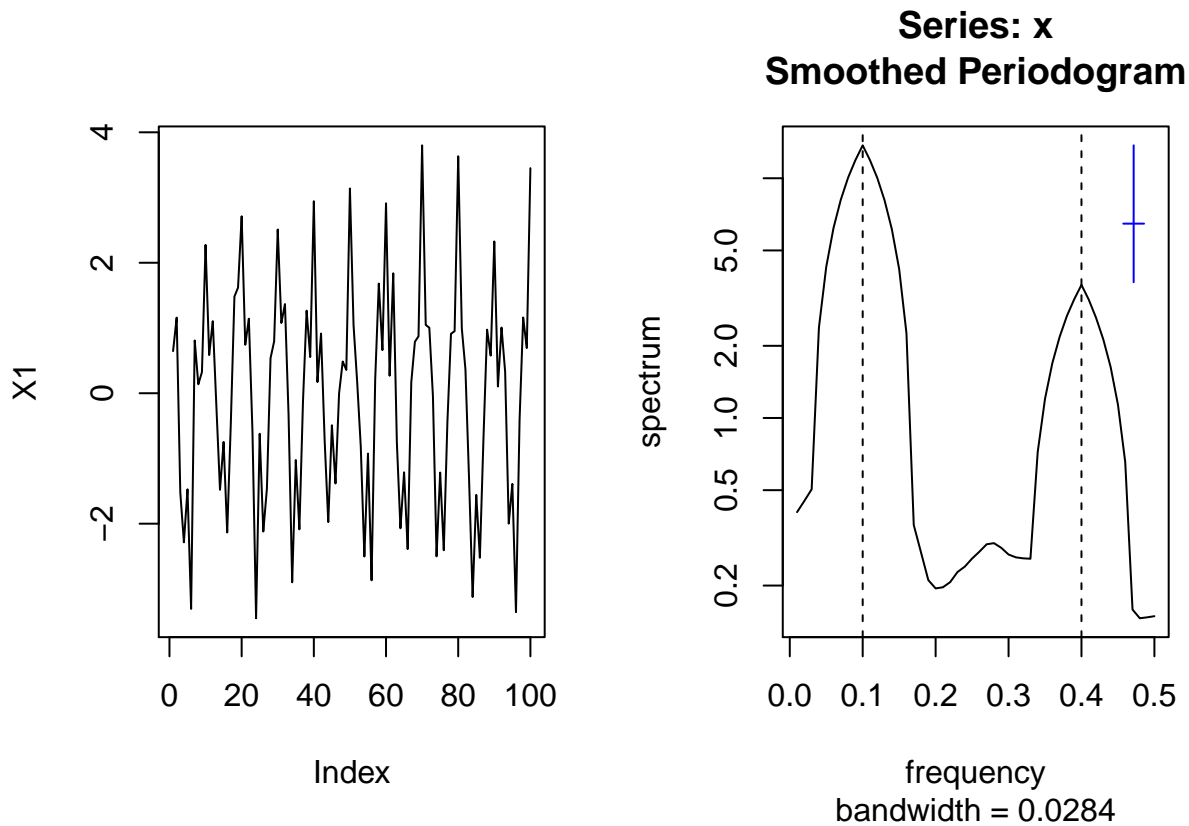
```
set.seed(131)
n<-1000
sigma<-1/2
eps<-rnorm(n,0,sd=sigma)

y03<-arima.sim(n = n, list(ar = c(0.3)),innov=eps)
y09<-arima.sim(n = n, list(ar = c(0.9)),innov=eps)

par(mfrow=c(2,2))
plot(y03,type='l',ylim=range(y03,y09))
spectrum(y03,kernel("daniell", c(3, 3)))
plot(y09,type='l')
spectrum(y09,kernel("daniell", c(3, 3)))
```



Un autre exemple avec le processus: $X_t = 2 \cos(2\pi t/10) + \cos(8\pi t/10) + \varepsilon_t$:



Opérateurs retard et avance

Définition

définition on appelle opérateur retard L l'opérateur qui associe à un processus $(X_t)_{t \in \mathbf{Z}}$ le processus $(Y_t)_{t \in \mathbf{Z}}$ tel que $Y_t = LX_t = X_{t-1}$. Cet opérateur est:

- linéaire
- inversible: son inverse est $L^{-1} = F$ tel que $FY_t = X_{t+1}$ et est appelé l'opérateur avance

De plus on a:

$$L^n X_t = X_{t-n}$$

et

$$\left(\sum_{i=1}^p a_i L^i \right) X_t = \sum_{i=1}^p a_i X_{t-i}$$

séries en L on peut définir à l'aide de cet opérateur des séries en L (ou F). Soit un processus stationnaire $(X_t)_{t \in \mathbf{Z}}$ et une suite $(a_i)_{i \in \mathbf{Z}}$ de nombres réels absolument sommable $\sum_{i \in \mathbf{Z}} |a_i| < \infty$, alors le processus défini par:

$$Y_t = \sum_{i \in \mathbf{Z}} a_i X_{t-i} = \left(\sum_{i \in \mathbf{Z}} a_i L^i \right) X_t$$

est stationnaire (voir le cours précédent).

Inversion de polynômes en L

Avec ces notations le processus AR(1) $Y_t = aY_{t-1} + \varepsilon_t$, avec $|a| < 1$ s'écrit ainsi:

$$(1 - aL)Y_t = \varepsilon_t$$

ainsi $Y_t = (1 - aL)^{-1}\varepsilon_t$ et l'écriture moyenne mobile de ce processus peut-être vue comme un problème d'inversion du polynôme $x \rightarrow 1 - ax$. En se rappelant qu'au voisinage de 0:

$$\frac{1}{1 - ax} = 1 + ax + a^2x^2 + \dots + a^kx^k + \dots$$

on a:

$$Y_t = \left(\sum_{i=0}^{\infty} a^i L^i \right) \varepsilon_t$$

Plus rigoureusement, $1 - aL$ est une application de l'ensemble des processus stationnaires dans lui même. La série $\sum_{i=0}^{\infty} a^i$ étant absolument sommable, la série $\sum_{i=0}^{\infty} a^i L^i$ est définie ($|a| < 1$) et nous avons:

$$\left(\sum_{i=0}^{\infty} a^i L^i \right) (1 - aL) = \sum_{i=0}^{\infty} a^i L^i - a^{i+1} L^{i+1} = L^0 = 1$$

$1 - aL$ est donc inversible et son inverse vaut $(1 - aL)^{-1} = \sum_{i=0}^{\infty} a^i L^i$. Remarquons que le processus $X_t = \sum_{i=0}^{\infty} a^i Y_{t-i}$ est l'unique processus stationnaire satisfaisant $(1 - aL)X_t = Y_t$ mais pas le seul processus. En effet, si Z est une v.a. quelconque, Za^t est solution (non-stationnaire!) de $(1 - aL)X_t = 0$.

Si $|a| > 1$, $1 - aL = -aL(1 - \frac{1}{a}F)$, $-aL$ est inversible et son inverse est $-\frac{1}{a}F$. D'autre part, la série $\sum_{i=0}^{\infty} \frac{1}{a^i} F^i$ existe et est l'inverse de $1 - F/a$.

On a donc:

$$\frac{1}{1 - aL} = \left(-\frac{1}{a}F \right) \left(\sum_{i=0}^{\infty} \frac{1}{a^i} F^i \right) = \sum_{i=1}^{\infty} \frac{1}{a^i} F^i$$

Dans le cas $|a| = 1$, $1 - aL$ n'est pas inversible. Si par exemple $a = 1$, tout processus constant est annulé par l'opérateur $1 - L$, l'application $1 - L$ n'est donc pas injective.

Plus généralement, si l'on considère un polynôme:

$$\Phi(z) = 1 + \phi_1 z + \phi_2 z^2 + \dots + \phi_p z^p$$

de racines $z_j = 1/a_j$ de modules supérieurs à 1. Alors on sait qu'il existe une série entière $\Psi(z) = \sum_{i \in \mathbf{Z}} \psi_i z^i$ telle que

$$\sum_{i=0}^{\infty} |\psi_i| < \infty \text{ et } \Phi(z)\Psi(z) = 1$$

En appliquant ce résultat aux séries en L , $\Phi(L) = 1 + \phi_1 L + \phi_2 L^2 + \dots + \phi_p L^p$ est inversible et son inverse est $\Psi(L)$.

Pour déterminer les coefficients ϕ_i on utilisera une des 3 méthodes suivantes:

- expliciter la relation $\Phi(z)\Psi(z) = 1$ en identifiant les coefficients de même degré et donc résoudre le système récursif suivant:

$$\begin{cases} \psi_0 & = 1 \\ \psi_1 + \phi_1 & = 0 \\ \psi_2 + \phi_1\psi_1 + \phi_2 & = 0 \\ \dots & \dots \\ \psi_p + \phi_1\psi_{p-1} + \dots + \phi_{p-1}\psi_1 + \phi_p & = 0 \\ \dots & \dots \\ \psi_n + \phi_1\psi_{n-1} + \dots + \phi_{p-1}\psi_{n-p+1} + \phi_p\psi_{n-p} & = 0 \quad \forall n > p \end{cases}$$

- effectuer la division de 1 par $\Phi(x)$ selon les puissances croissantes.
- décomposer en éléments simples la fraction $1/\Phi(x)$, puis on écrit le développement en série de chaun des termes de la décomposition. Ainsi, si toutes les racines λ_i de Φ sont réelles distinctes on a:

$$\frac{1}{\Phi(x)} = \frac{1}{\prod_{i=1}^p (1 - \lambda_i x)} = \sum_{i=1}^p \frac{a_i}{1 - \lambda_i x} = \sum_{i=1}^p a_i \sum_{j=0}^{\infty} \lambda_i^j z^j = \sum_{j=0}^{\infty} \left[\sum_{i=1}^p a_i \lambda_i^j \right] z^j$$

L'inverse de Φ existe dès que ses racines sont de module différent de 1. En effet, supposons que les r premières racines soient de module supérieur à 1 et que les $p - r$ dernières de module inférieur à 1, on a de même que pour $1 - aL$:

$$\Phi(L) = \prod_{i=1}^p (1 - \lambda_i L) = \prod_{i=1}^r (1 - \lambda_i L) \prod_{i=r+1}^p \left(1 - \frac{A}{\lambda_i} F\right) \prod_{i=r+1}^p (-\lambda_i L)$$

Soit:

$$\Phi(L) = \Phi_1(L)\Phi_2(F)\lambda L^{p-r}$$

Φ_1 et Φ_2 sont inversibles (racines de module >1) donc $\Phi(L)$ est inversible et son inverse vaut:

$$\frac{1}{\Phi(L)} = \frac{1}{\Phi_1(L)} \frac{1}{\Phi_2(F)} \frac{1}{\lambda} F^{p-r}$$

Processus auto-régressif (AR)

définition on appelle processus autorégressif d'ordre p , usuellement noté AR(p), un processus stationnaire $(X_t)_{t \in \mathbf{Z}}$ vérifiant une relation du type:

$$X_t + \sum_{i=1}^p \phi_i X_{t-i} = \varepsilon_t, \quad \forall t \in \mathbf{Z}$$

avec $\phi_i \in \mathbf{R}$ et ε_t un bruit blanc de variance σ^2 .

Cette relation s'écrit également:

$$\Phi(L)X_t = \varepsilon_t$$

où $\Phi(L) = 1 + \phi_1 L + \dots + \phi_p L^p$

Écriture moyenne mobile infinie la définition du processus AR(p) proposée n'implique pas forcément que cette équation a une solution unique stationnaire. Pour cela, il faut relier ce problème à la question de l'inversion du polynôme en L : $\Phi(L)$ vu précédemment.

- si Φ a toutes ses racines de module différent de 1 on peut inverser $\Phi(L)$ et l'équation a une unique solution avec une écriture moyenne mobile infinie:

$$X_t = \Phi(L)^{-1} \varepsilon_t = \sum_{j=-\infty}^{+\infty} h_j \varepsilon_{t-j}, \quad \sum_{j=-\infty}^{+\infty} |h_j| < +\infty$$

on voit que dans ce cas la valeur de X_t peut dépendre des valeurs passées, présentes et futures de ε_t . Dans le cas où les racines de Φ sont de module strictement supérieur à 1, $\Phi^{-1}(L)$ ayant un développement ne contenant que des puissances positives de L , la valeur de X_t ne dépend que des valeurs passées de ε_t :

$$X_t = \sum_{j=0}^{+\infty} h_j \varepsilon_{t-j}, \quad \sum_{j=0}^{+\infty} |h_j| < +\infty, \quad h_0 = 1$$

on voit donc que X_{t-1} ne dépend (linéairement) que de $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$. Comme ε est un bruit blanc, X_{t-1} est donc non corrélée à ε_t . La projection linéaire de X_t sur son passé s'écrit donc:

$$P_{M_{t-1}^{-\infty}}(X_t) = - \sum_{i=1}^p \phi_i X_{t-i}$$

et

$$X_t - P_{M_{t-1}^{-\infty}}(X_t) = \varepsilon_t$$

Le bruit blanc ε_t est appelé le processus des innovations associé à X . Lorsque ε_t est un bruit blanc fort, l'indépendance des innovations implique que $P_{M_{t-1}^{-\infty}}(X_t)$ est non-seulement la régression linéaire de X_t sur son passé mais aussi son espérance conditionnelle.

Modification des racines lorsque les racines de Φ sont de module différent de 1 on peut montrer que, quitte à changer de bruit blanc générateur, on peut supposer que ces racines sont de module supérieur à 1.

En utilisant la densité spectrale on a:

$$\frac{\sigma^2}{2\pi} = f_X(\omega) \left| 1 + \sum_{j=1}^p \phi_j e^{i\omega j} \right|^2$$

et

$$f_X(\omega) = \frac{\sigma^2}{2\pi \left| 1 + \sum_{j=1}^p \phi_j e^{i\omega j} \right|^2} = \frac{\sigma^2}{2\pi |\Phi(e^{i\omega})|^2}$$

en notant z_j les racines de Φ , on a:

$$f_X(\omega) = \frac{\sigma^2}{2\pi \left| \prod_{j=1}^p \left(1 - \frac{e^{i\omega}}{z_j} \right) \right|^2}$$

supposons que les r premières racines de Φ soient de module supérieur à 1 et les $p - r$ de module inférieur à 1. On note:

$$\Phi^* = \prod_{j=1}^r \left(1 - \frac{L}{z_j}\right) \prod_{j=r+1}^p (1 - Lz_j)$$

le polynôme construit en remplaçant les racines de module inférieur à 1 par leur inverse dans Φ .

Le processus:

$$\eta_t = \Phi^* X_t$$

est un bruit blanc:

$$\begin{aligned} f_\eta(\omega) &= |\Phi^*(e^{i\omega})|^2 f_X(\omega) \\ &= \frac{\sigma^2 |\Phi^*(e^{i\omega})|^2}{2\pi |\Phi(e^{i\omega})|^2} \\ &= \frac{\sigma^2}{2\pi} \prod_{j=r+1}^p \frac{|1 - e^{i\omega} z_j|^2}{|1 - \frac{e^{i\omega}}{z_j}|^2} = \frac{\sigma^2}{2\pi} \prod_{j=r+1}^p |z_j|^2 \end{aligned}$$

ce qui est bien indépendant de ω . Donc η_t est bien un bruit blanc de variance $\sigma^2 \prod_{j=r+1}^p |z_j|^2$, variance inférieure à σ^2 variance de ε . Cette nouvelle représentation de X_t fait donc intervenir un bruit blanc s'interprétant comme l'innovation (on peut développer X_t en moyenne mobile infinie ne dépendant que du passé de η_t) et dont la variance est plus faible que celle du bruit blanc initial.

On voit donc qu'en choisissant une racine ou son inverse on peut obtenir de nombreuses représentations de X_t . On appelle la représentation dans laquelle toutes les racines sont de module supérieur à 1 la représentation canonique du processus X_t , celle dans laquelle le bruit blanc générateur est l'innovation.

Fonction d'autocorrélation d'un AR(p) soit un processus stationnaire $AR(p)$ défini par

$$\Phi(L)X_t = X_t + \sum_{i=1}^p \phi_i X_{t-i} = \varepsilon_t$$

dont les racines sont de modules > 1 .

Alors la fonction d'autocorrélation:

$$\rho(h) + \sum_{j=1}^p \phi_j \rho(h-j) = 0, \quad h = 1, 2, \dots$$

est donc définie comme la solution d'une suite récurrente d'ordre p , donc racines non-nulles de l'équation:

$$\lambda^p + \sum_{j=1}^p \phi_j \lambda^{p-j} = 0$$

soit:

$$1 + \sum_{j=1}^p \phi_j \lambda^{-j} = 0$$

c'est à dire que $\frac{1}{\lambda}$ est racine de Φ donc $|\lambda| < 1$. On sait donc que $\rho(h)$ est une combinaison de différentes composantes: exponentielles décroissantes pour les racines réelles, sinusoïdes amorties pour les racines complexes, exponentielles décroissantes et termes polynomiaux pour les racines multiples. Par conséquent, $\rho(h)$ tend exponentiellement vers 0 avec h .

Remarquons également qu'on a la relation matricielle (équation de Yule-Walker):

$$\begin{pmatrix} -\phi_1 \\ \vdots \\ -\phi_p \end{pmatrix} = [\rho(i-j)]^{-1} \begin{pmatrix} \rho(1) \\ \vdots \\ \rho(p) \end{pmatrix}$$

Ce qui donne un algorithme d'estimation des coefficients d'un processus AR(p) en remplaçant $\rho(h)$ par leur autocorrélation empirique. Remarquons également que la variance de ce processus vaut:

$$\gamma(0) = - \sum_{j=1}^p \phi_j \gamma(j) + \sigma^2 = \frac{\sigma^2}{1 + \sum_{j=1}^p \phi_j \rho(j)}$$

Fonction d'autocorrélation partielle d'un AR(p) soit un processus stationnaire $AR(p)$ défini par

$$\Phi(L)X_t = X_t + \sum_{i=1}^p \phi_i X_{t-i} = \varepsilon_t$$

dont les racines sont à l'extérieur du disque unité.

On a que l'autocorrélation partielle $r(h) = 0$ si $h \geq p + 1$ car la projection de X_t sur $M(X_{t-1}, \dots, X_{t-h})$ est $-\sum_{j=1}^p \phi_j X_{t-j}$ et le coefficient associé à X_{t-h} est nul.

Cette propriété est très utile en pratique lorsque l'on cherche à identifier l'ordre d'un processus AR. On peut ainsi calculer les autocorrélation partielles empiriques et regarder quand celles-ci sont négligeables (non-significativement différentes de 0). Notons également que $r(p) = -\phi_p$.

Processus moyennes mobiles (MA)

définition on appelle processus moyenne mobile d'ordre q , usuellement noté MA(q) (moving average), un processus $(X_t)_{t \in \mathbf{Z}}$ défini par:

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad \forall t \in \mathbf{Z}$$

où les θ_i sont des réels et ε_t est un bruit blanc de variance σ^2 .

De même que pour les processus autorégressifs cette relation s'écrit:

$$X_t = (1 + \theta_1 L + \dots + \theta_q L^q) \varepsilon_t = \Theta(L) \varepsilon_t$$

Contrairement à l'AR(p) la définition du MA(q) est explicite et le processus X_t est stationnaire.

Ecriture autorégressive infinie par analogie avec ce qu'on a vu pour les AR(p), si $\Theta(z)$ a des racines de modules différents de 1, on peut écrire:

$$\sum_{i=-\infty}^{+\infty} \pi_i X_{t-i} = \varepsilon_t, \text{ avec } \sum_{i=-\infty}^{+\infty} |\pi_i| < +\infty$$

si les racines sont toutes de module supérieur à 1, $\pi_i = 0$ pour $i < 0$ et ε_t est l'innovation du processus X_t (démonstration similaire à celle effectuée pour les AR, à faire en exercice). On dit dans ce cas que le processus est inversible.

Modification des racines si $\Theta(z)$ a certaines de ses racines à l'intérieur du disque unité, on peut, quitte à changer de bruit blanc, se ramener à une moyenne mobile inversible.

Par exemple, si

$$X_t = \Theta(L)\varepsilon_t = (1 - z_1 L)(1 - z_2 L)\dots(1 - z_q L)$$

où $z_1^{-1}, \dots, z_q^{-1}$ sont les racines de Θ .

Supposons que les r premières racines soient de module > 1 .

La densité spectrale de X est donnée par:

$$\begin{aligned} f_X(\omega) &= \frac{\sigma^2}{2\pi} \prod_{j=1}^q |1 - z_j e^{i\omega}|^2 \\ &= \frac{\sigma^2}{2\pi} \prod_{j=1}^r |z_j|^2 \prod_{j=1}^r |1 - z_j e^{i\omega}|^2 \prod_{j=r+1}^q |1 - z_j^{-1} e^{-i\omega}|^2 \\ &= \frac{\sigma^2}{2\pi} \prod_{j=1}^r |z_j|^2 |\Theta^*(e^{i\omega})|^2 \end{aligned}$$

et le polynôme Θ^* a toutes ses racines à l'extérieur du disque unité.

En posant $\eta_t = \Theta^*(L)^{-1} X_t$ on obtient que:

$$f_\eta(\omega) = \frac{f_X(\omega)}{|\Theta^*(e^{i\omega})|^2} = \frac{\sigma^2}{2\pi} \prod_{j=1}^r |z_j|^2$$

donc η est un bruit blanc de variance $\sigma^2 \prod_{j=1}^r |z_j|^2$ et $X_t = \Theta^*(L)\eta_t$ est un MA(q) inversible.

Rappelons que la variance d'un MA(q) vaut:

$$\text{var}(X_t) = \sigma^2(1 + \theta_1^2 + \dots + \theta_q^2)$$

Fonction d'autocorrélation d'un MA(q) la fonction d'autocovariance est:

$$\begin{aligned} \gamma(h) &= E(X_t X_{t-h}) \\ &= E[(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q})(\varepsilon_{t-h} + \theta_1 \varepsilon_{t-h-1} + \dots + \theta_q \varepsilon_{t-h-q})] \end{aligned}$$

et donc

$$\begin{cases} \gamma(h) = (\theta_h + \theta_{h+1}\theta_1 + \dots + \theta_q\theta_{q-h})\sigma^2, & \text{si } 1 \leq h \leq q \\ = 0 & \text{sinon} \end{cases}$$

et la fonction d'autocorrélation s'écrit:

$$\begin{cases} \rho(h) = \frac{\theta_h + \theta_{h+1}\theta_1 + \dots + \theta_q\theta_{q-h}}{1 + \theta_1^2 + \dots + \theta_q^2}, & \text{si } 1 \leq h \leq q \\ = 0 & \text{sinon} \end{cases}$$

ce résultat est important en pratique car $\rho(h)$ s'annule pour $h \geq q$ et $\rho(q) = \theta_q / (1 + \theta_1^2 + \dots + \theta_q^2)$ est non nul si θ_q est non nul ie que X_t est un MA(q). Pour déterminer l'ordre d'un MA(q) en pratique, on calcule donc ses autocorrélations empiriques et on regarde quand celles-ci sont négligeables (non-significativement différentes de 0). L'autocorrélation pour les MA(q) est donc l'analogue de l'autocorrélation partielle pour les AR(p).

Fonction d'autocorrélation partielle d'un MA(q) la fonction d'autocorrélation partielle peut se calculer par l'algorithme de durbin-watson mais son expression est compliquée. D'autre part, contrairement aux processus AR(p), $r(h)$ ne s'annule pas à partir d'une certaine valeur de h .

Processus autorégressifs moyennes mobiles (ARMA)

Les processus ARMA(p,q) généralise les modèles autorégressifs et moyennes mobiles. Ces modèles sont très utiles en pratique pour modéliser des séries réelles en nécessitant moins de paramètres que les modèles AR ou MA simples.

définition un processus stationnaire X_t admet une représentation ARMA(p,q) minimale s'il satisfait

$$X_t + \sum_{i=1}^p \phi_i X_{t-i} = \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}, \quad \forall t \in \mathbf{Z}$$

soit, avec les notations précédentes

$$\Phi(L)X_t = \Theta(L)\varepsilon_t, \quad \forall t \in \mathbf{Z}$$

avec les conditions suivantes:

1. $\phi_p \neq 0$ et $\theta_q \neq 0$
2. Φ et Θ n'ont pas de racines communes et leurs racines sont de modules > 1
3. ε_t est un BB de variance σ^2

propriété si X_t est un processus stationnaire de représentation ARMA(p,q) minimale:

$$\Phi(L)X_t = \Theta(L)\varepsilon_t, \quad \forall t \in \mathbf{Z}$$

alors:

1. X admet la représentation MA(∞):

$$X_t = \frac{\Theta(L)}{\Phi(L)}\varepsilon_t = \sum_{j=0}^{+\infty} h_j \varepsilon_{t-j}, \quad h_0 = 1$$

2. X admet la représentation AR(∞):

$$\frac{\Phi(L)}{\Theta(L)}X_t = \sum_{j=0}^{+\infty} \pi_j X_{t-j} = \varepsilon_t, \quad \pi_0 = 1$$

3. l'innovation de X est ε

autocovariance d'un ARMA(p,q) si X_t est un processus stationnaire de représentation ARMA(p,q) minimale:

$$\Phi(L)X_t = \Theta(L)\varepsilon_t, \quad \forall t \in \mathbf{Z}$$

sa fonction d'autocovariance vérifie:

$$\gamma(h) + \sum_{j=1}^p \phi_j \gamma(h-j) = 0, \quad \forall h \geq q+1$$

preuve on a

$$X_t + \sum_{i=1}^p \phi_i X_{t-i} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad \forall t \in \mathbf{Z}$$

en multipliant par X_{t-h} , $h \geq q+1$ puis en prenant l'espérance on obtient bien:

$$\gamma(h) + \sum_{j=1}^p \phi_j \gamma(h-j) = \text{cov}(X_{t-h}, \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}) = 0$$

la covariance satisfait donc une équation de récurrence d'ordre p à partir du rang $q+1$, dont le polynôme caractéristique est $z^p \Phi(1/z)$. Ainsi, l'autocovariance (et l'autocorrélation) d'un processus ARMA(p,q) tendent exponentiellement vers 0 avec h à partir de l'ordre $q+1$.

Si Φ a toutes ses racines distinctes égales à z_j , $j = 1, \dots, p$, $|z_j| > 1$ les covariances sont de la forme:

$$\gamma(h) = \sum_{j=1}^p A_j \frac{1}{z_j^h}, \quad h \geq \max(q+1-p, 0)$$

pour obtenir les valeurs initiales des covariances on exploite le développement MA(∞):

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

et la définition de l'ARMA(p,q):

$$X_t + \sum_{j=1}^p \phi_j X_{t-j} = \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}$$

En multipliant par X_{t-h} et en prenant l'espérance on a:

$$\gamma(h) + \sum_{j=1}^p \phi_j \gamma(h-j) = \text{cov}(\varepsilon_t + \sum_{j=1}^q \varepsilon_{t-j}, \sum_{j=0}^{+\infty} h_j \varepsilon_{t-j})$$

et donc:

$$\gamma(h) + \sum_{j=1}^p \phi_j \gamma(h-j) = \sigma^2(\theta_h + h_1 \theta_{h+1} + \dots + h_{q-h} \theta_q), \quad 0 \leq h \leq q$$

développement autorégressif infini d'un ARMA(p,q)

on cherche la série en L π vérifiant $\pi(L)\Theta(L) = \Phi(L)$ soit:

$$\sum_{j=0}^{\infty} \pi_j L^j \sum_{j=0}^q \theta_j L^j = \sum_{j=0}^p \phi_j L^j$$

cela implique:

$$\sum_{j=0}^q \theta_j \pi_{h-j} = 0, \quad \forall h \geq \max(p+1, q)$$

les coefficients du développement $AR(\infty)$ satisfont une équation de récurrence d'ordre q à partir du rang $\max(p+1, q)$. Le polynôme caractéristique est $z^q\Theta(1/z)$.

développement moyenne mobile infinie d'un ARMA(p,q)

de même que pour le développement AR infini, les coefficients h_j de la représentation moyenne infinie vérifient:

$$\sum_{j=0}^p \phi_j h_{l-j} = 0, \quad \forall l \geq \max(q+1, p)$$

les coefficients du développement $MA(\infty)$ satisfont une équation de récurrence d'ordre p à partir du rang $\max(q+1, p)$. Le polynôme caractéristique est $z^p\Phi(1/z)$.

densité spectrale d'un ARMA(p,q)

$$f_X(\omega) = \frac{\sigma^2 |\Theta(e^{i\omega})|^2}{2\pi |\Phi(e^{i\omega})|^2}$$

détermination des ordres

l'intérêt des modèles ARMA tient dans leur parcimonie. En effet, pour un ajustement aux données de la même qualité, un ARMA nécessite moins de paramètres qu'un AR ou un MA pur (on le comprend en regardant la décomposition AR ou MA infinie). En revanche, le choix des ordres est plus complexe que pour les modèles AR ou MA purs.

Identification et estimation de modèles ARMA

En pratique, lorsque l'on doit ajuster un modèle AR, MA ou ARMA à des données réelles la première question qui se pose est celle du choix des ordres p et q du modèle ARMA (on considère que les AR et MA sont un cas particulier d'ARMA avec respectivement $q=0$ et $p=0$). Pour choisir ces ordres, nous pouvons exploiter les résultats suivants:

Type de processus	Définition	autocorrélations	autocorrélations partielles
AR(p)	$\Phi(L)X_t = \varepsilon_t$	$\rho(h) \searrow 0$	$r(h) = 0$ pour $h \geq p+1$
MA(q)	$X_t = \Theta(L)\varepsilon_t$	$\rho(h) = 0$ pour $h \geq q+1$	$r(h)$ rien de particulier
ARMA(p,q)	$\Phi(L)X_t = \Theta(L)\varepsilon_t$	$\rho(h) \searrow 0, h \geq q+1$	$r(h) \searrow 0, h \geq \max(q+1, p+1)$

Ses résultats sont facilement exploitables dans le cas d'un AR ou MA pur. Dans le cas d'un ARMA il existe un moyen de déterminer les ordres p et q à l'aide de la méthode du coin basée sur certains déterminant de matrices de corrélations que nous ne développons pas ici. Une autre approche est de considérer un ensemble crédible de modèles ARMA(p,q) puis de sélectionner un candidat par une méthode de sélection de modèle basée sur des critères d'information de type BIC ou AIC.

Une fois l'ordre et le type de modèle choisis, pour estimer les coefficients des modèles plusieurs approches sont possibles.

1. Equations de Yule-Walker. Nous avons obtenus précédemment que dans le cas d'un processus AR, les coefficients vérifient:

$$\begin{pmatrix} -\phi_1 \\ \vdots \\ -\phi_p \end{pmatrix} = [\rho(i-j)]^{-1} \begin{pmatrix} \rho(1) \\ \vdots \\ \rho(p) \end{pmatrix}$$

en exploitant le fait que les autocorrélations empiriques sont un estimateur convergent des autocorrélations, on peut obtenir un estimateur des coefficients ϕ_j .

Pour obtenir ces équations, nous avons exploités les relations d'orthogonalités. Rappel:

$$X_t + \sum_{j=1}^p X_{t-j} = \varepsilon_t$$

en multipliant par X_{t-h} , en prenant l'espérance et en remarquant que pour $h > 0$, $E(\varepsilon_t X_{t-h}) = 0$ on obtient la relation à l'origine des équations de Yule-Walker:

$$\gamma(h) + \sum_{j=1}^p \gamma(h-j) = 0$$

Cela revient à calculer la projection linéaire de X_t sur son passé donc à minimiser en ϕ_j le risque quadratique:

$$E\left(X_t - \sum_{j=1}^p \phi_j X_{t-j}\right)^2$$

Une autre approche consiste à minimiser directement le risque empirique:

$$\sum_{t=p+1}^n \left(X_t - \sum_{j=1}^p \phi_j X_{t-j}\right)^2$$

on peut montrer que le comportement asymptotique de cet estimateur est le même que celui obtenu par Yule-Walker.

- Moindre carrés conditionnels. si ε_t est gaussien, le vecteur des observations (X_1, \dots, X_n) est gaussien et on peut alors calculer sa vraisemblance. Une approche simplifiée est de calculer la densité de (X_{p+1}, \dots, X_n) conditionnellement à (X_1, \dots, X_p) qui est gaussienne. Sachant que $X_t/X_{t-1}, \dots, X_{t-p}$ est une gaussienne d'espérance $\sum_{j=1}^p \phi_j X_{t-j}$ et de variance σ^2 on a, en utilisant la formule des conditionnements successifs:

$$\begin{aligned} f(X_{p+1}, \dots, X_n / X_1, \dots, X_p) &= \prod_{t=p+1}^n f(X_t / X_{t-1}, \dots, X_{t-p}) \\ &= \frac{1}{(2\pi\sigma^2)^{\frac{n-p}{2}}} \exp\left(-\frac{\sum_{t=p+1}^n (X_t - \sum_{j=1}^p \phi_j X_{t-j})^2}{2\sigma^2}\right) \end{aligned}$$

maximiser cette vraisemblance conditionnelle en ϕ revient à, en passant au log, minimiser:

$$(\phi_1, \dots, \phi_p, \sigma^2) \rightarrow \frac{\sum_{t=p+1}^n (X_t - \sum_{j=1}^p \phi_j X_{t-j})^2}{\sigma^2} + (n-p) \ln(\sigma^2)$$

à σ^2 fixé cela correspond au critère des moindres carrés.

La démarche est donc d'obtenir $\hat{\phi}_{css}$ (pour *conditional sum of square*) en minimisant:

$$\hat{\phi}_{css} = \operatorname{argmin}_{\phi} \sum_{t=p+1}^n (X_t - \sum_{j=1}^p \phi_j X_{t-j})^2$$

puis

$$\hat{\sigma}_{css}^2 = \frac{1}{n-p} \sum_{t=p+1}^n (X_t - \sum_{j=1}^p \hat{\phi}_{css} X_{t-j})^2$$

cette approche se généralise aux ARMA, de même que le calcul de la vraisemblance exacte que nous n'aborderons pas ici.

Revenons sur le choix de l'ordre d'un ARMA(p,q). Nous avons vu précédemment que l'étude des autocorrélations et autocorrélations partielles peuvent permettre de préselectionner un certain nombre de modèles plausibles. On peut ensuite, une fois les paramètres de ces modèles estimés sélectionner celui qui minimise les critères suivants:

- **AIC** *Akaike Information Criterion*, adapté au problème de la prévision, défini par:

$$\text{AIC}(\phi, \theta, \sigma^2) = -2 \log(L(\theta, \phi, \sigma^2)) + 2k$$

où L est la vraisemblance, k est le nombre de paramètres dans le modèle donc $p+q$ ou $p+q+1$ si la constante est dans le modèle.

- **BIC** *Bayesian Information Criterion*, adapté au problème de la prévision, défini par:

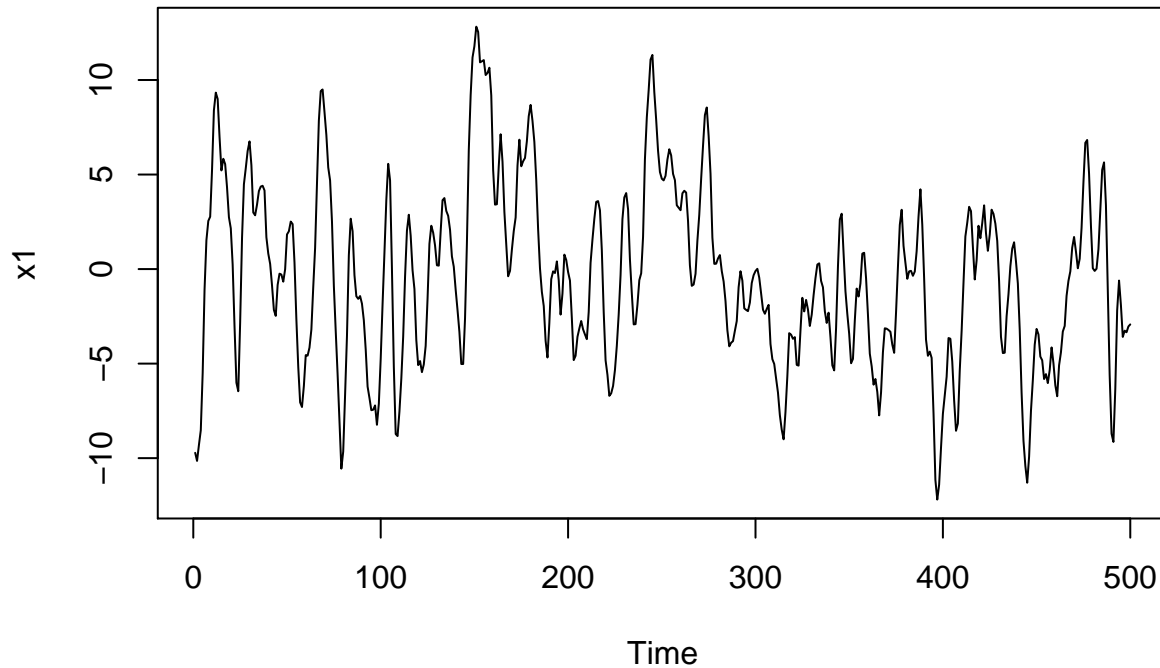
$$\text{BIC}(\phi, \theta, \sigma^2) = -2 \log(L(\theta, \phi, \sigma^2)) + \log(n)k$$

le principe est donc de sélectionner un modèle qui colle bien aux données (grande vraisemblance) tout en pénalisant les modèles comprenant trop de paramètres.

Voilà ci-dessous un exemple obtenu pour un processus AR(p) simulé ainsi:

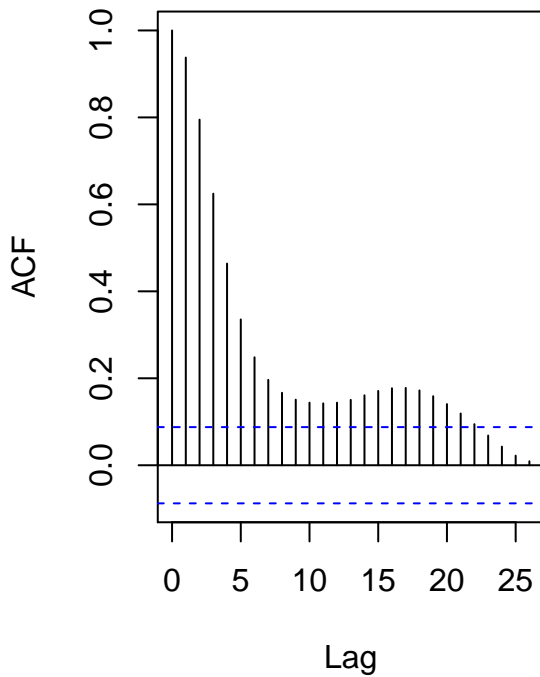
```
library(polynom)
n<-500
sigma<-1
eps<-rnorm(n,0,sd=sigma)
pol<-poly.calc(x=c(3,2,4,1.5,5))
pol<-pol/pol[1]
pol

## -1 + 1.95*x - 1.447222*x^2 + 0.5111111*x^3 - 0.08611111*x^4 + 0.005555556*x^5
x1<-arima.sim(n = n, list(ar = pol[2:length(pol)],innov=eps))
plot(x1)
```

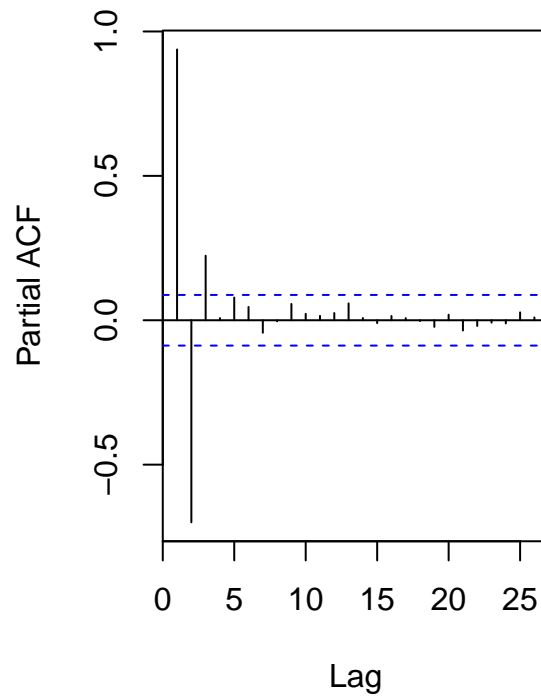


```
par(mfrow=c(1,2))
acf(x1)
pacf(x1)
```

Series x1



Series x1



L'estimation d'un modèle AR(p) par maximum de vraisemblance sur cette série se fait via la commande:

```
x1.model<-arima(x1, order = c(p,0,0), method = c("ML"), SSinit = c("Rossignol2011"),
  optim.method = "BFGS", include.mean = F)
```

On effectue ce calcul pour $p = 1, \dots, 20$ et on obtient:

