Université de Paris 5 – René Descartes
UFR de Mathématiques et Informatique
45, rue des Saints-Pères 75270 Paris cedex 06

Statistique des Processus - Maîtrise MASS

Corrigé de l'examen du 10 juin 2003

9h - 12h

Ex 1.

1. Vrai. Si X et Y sont deux processus faiblement stationnaires indépendants de fonctions d'autocovariance γ_X et γ_Y, alors $X+Y$ est aussi un processus faiblement stationnaire de fonction d'autocovariance $\gamma_{X+Y} = \gamma_X + \gamma_Y$. Si X et Y sont de plus des processus MA, alors γ_X et γ_Y s'annulent à partir d'un certain rang ; il en est donc de même pour γ_{X+Y}. Or cela est caractéristique d'un processus MA.
2. Vrai. Si X et Y sont deux processus faiblement stationnaires indépendants tels que $\phi(B)(X) = \varepsilon$ et $Y = \theta(B)(\varepsilon')$, alors $X+Y$ est aussi un processus faiblement stationnaire tel que $\phi(B)(X+Y) = \varepsilon + \phi(B) \circ \theta(B)(\varepsilon')$. Or le processus $\varepsilon + \phi(B) \circ \theta(B)(\varepsilon')$ est un processus MA comme somme de deux processus MA indépendants (cf question précédente). Ce qui montre que $X+Y$ est un ARMA.
3. Faux (sauf cas particulier). Exemple : $X = \varepsilon$ et Y faiblement stationnaire tel que $(I - 0,5B)(Y) = \varepsilon'$ avec ε et ε' indépendants. Alors $X+Y$ vérifie $(I-0,5B)(X+Y) = (I-0,5B)(\varepsilon) + \varepsilon'$. La fonction d'autocovariance de $(I-0,5B)(\varepsilon) + \varepsilon'$ ne s'annule pas en 1, mais s'annule à partir du rang 2 : il s'agit donc d'une MA(1). Autrement dit, $X+Y$ est un ARMA$(1,1)$.
4. Vrai. Un processus faiblement stationnaire dont la fonction d'autocorrélation s'annule à partir du rang 1 inclus est un bruit blanc faible ; sa fonction d'autocorrélation partielle s'annule donc aussi à partir du rang 1 inclus.
5. Faux. Il existe une représentation de Wold pour tout processus faiblement stationnaire.
6. Vrai si le processus ARMA possède une représentation inversible. En effet, si X est un processus ARMA tel que $\phi(B)(X) = \theta(B)(\varepsilon)$, alors la densité spectrale vaut
$$f_X(\theta) = \frac{1}{2\pi}\left|\frac{\theta(e^{i\theta})}{\phi(e^{i\theta})}\right|^2$$
Elle s'annule si et seulement s'il existe θ_0 tel que $\theta(e^{i\theta_0}) = 0$, autrement dit si la transformée en z de $\theta(B)$ s'annule sur le cercle unité.
7. Faux, si la période est connue. Par exemple, un modèle du type $Y_t = \alpha\cos(t\omega) + \beta\sin(\omega t) + X_t$ avec α, β paramètres et X bruit corrélé modélise une tendance périodique.
8. Vrai. Un bruit corrélé n'est autre que l'image par une transformation linéaire inversible d'un bruit blanc ordinaire qu'on peut dès lors tester.

9. Faux. Un processus ARIMA(p, d, q) n'est pas faiblement stationnaire dès lors que $d \geq 1$.
10. Faux. L'histoire ne se répète pas... Un processus d'innovation est une famille de variables décorrélées, et ne peut être périodique que s'il est nul.

Ex 2.

1. X est un processus AR(2), Y un ARMA(1, 1) et Z un MA(2).
   - On a $\phi(B)(X) = \varepsilon$ avec $\phi(z) = 1 + 1,6z + 0,8z^2$. La représentation est inversible de nature pour un processus AR. Elle est causale si et seulement si $\phi(z)$ ne s'annule pas sur le disque unité. Or

$$\phi(z) = \left(1 + \frac{z}{1 + i/2}\right)\left(1 + \frac{z}{1 - i/2}\right)$$

Les racines sont donc de module strictement supérieur à 1, la représentation est donc aussi causale. On en déduit que ε est le bruit blanc d'innovation de X. Notons

$$\beta = \frac{1}{|1 + i/2|} = \frac{2}{\sqrt{5}} \text{ et } \alpha = \arctan(1/2)$$

Alors

$$\begin{aligned}
\frac{1}{\phi(z)} &= \frac{1}{(1 + z\beta e^{-i\alpha})(1 + z\beta e^{i\alpha})} \\
&= \frac{1}{1 - e^{2i\alpha}} \frac{1}{1 + z\beta e^{-i\alpha}} + \frac{1}{1 - e^{-2i\alpha}} \frac{1}{1 + z\beta e^{i\alpha}} \\
&= \sum_{u \geq 0} (-1)^u z^u \beta^u \left(\frac{e^{-iu\alpha}}{1 - e^{2i\alpha}} + \frac{e^{iu\alpha}}{1 - e^{-2i\alpha}}\right) \\
&= \frac{1}{e^{i\alpha} - e^{-i\alpha}} \sum_{u \geq 0} (-1)^u z^u \beta^u \left(e^{i(u+1)\alpha} - e^{-i(u+1)\alpha}\right) \\
&= \frac{1}{\sin(\alpha)} \sum_{u \geq 0} (-1)^u z^u \beta^u \sin((u+1)\alpha)
\end{aligned}$$

On en déduit

$$\forall t \in \mathbb{Z}, \quad X_t = \phi(B)^{-1}(\varepsilon)_t = \frac{1}{\sin(\alpha)} \sum_{u \geq 0} (-1)^u \varepsilon_{t-u} \beta^u \sin((u+1)\alpha)$$

et, pour tout $t \geq 0$,

$$\begin{aligned}
\gamma_X(t) &= \operatorname{cov}(X_t, X_0) \\
&= \frac{1}{\sin(\alpha)^2} \sum_{u \geq 0} \sum_{v \geq 0} (-1)^u \beta^u \sin((u+1)\alpha) (-1)^v \beta^v \sin((v+1)\alpha) \operatorname{cov}(\varepsilon_{t-u}, \varepsilon_{-v}) \\
&= \frac{\sigma^2}{\sin(\alpha)^2} \sum_{v \geq 0} (-1)^{t+2v} \beta^{t+2v} \sin((t+v+1)\alpha) \sin((v+1)\alpha)
\end{aligned}$$

car $\text{cov}(\varepsilon_{t-u}, \varepsilon_{-v}) = \sigma^2 \delta_{u=t+v}$. D'où

$$\begin{aligned}
\gamma_X(t) &= \frac{\sigma^2}{2\sin(\alpha)^2}(-1)^t\beta^t \sum_{v\geq 0} \beta^{2v} \left(\cos(t\alpha) - \cos\left((t+2v+2)\,\alpha\right)\right) \\
&= \frac{\sigma^2}{2\sin(\alpha)^2}(-1)^t\beta^t \left(\frac{\cos(t\alpha)}{1-\beta^2} - \Re\left(\frac{e^{i(t+2)\alpha}}{1-\beta^2 e^{2i\alpha}}\right)\right) \\
&= \frac{\sigma^2}{2\sin(\alpha)^2}(-1)^t\beta^t \left(\frac{\cos(t\alpha)}{1-\beta^2} - \frac{\cos\left((t+2)\,\alpha\right) - \beta^2\cos(t\alpha)}{1-2\beta^2\cos(2\alpha)+\beta^4}\right)
\end{aligned}$$

On peut encore simplifier l'expression ci-dessus, en remarquant que $\sin(\alpha)^2 = 1 - 1/(1+\tan(\alpha)^2) = 1/5$:

$$\gamma_X(t) = \sigma^2 \frac{25}{2}\beta^t \left(\cos(\alpha t) - \frac{1}{17}\left(5\cos\left((t+2)\,\alpha\right) - 4\cos(t\alpha)\right)\right)$$

On en déduit en particulier

$$\begin{aligned}
\gamma_X(0) &= \frac{225}{17}\sigma^2 \\
\gamma_X(1) &= -\frac{200}{17}\sigma^2 \\
\rho_X(1) &= -\frac{8}{9} \simeq -0,89
\end{aligned}$$

en utilisant

$$\begin{aligned}
\cos(\alpha) &= \Re\,(1+i/2)\,\beta = \frac{2}{\sqrt{5}} \\
\cos(2\alpha) &= 2\cos(\alpha)^2 - 1 = \frac{3}{5} \\
\cos(3\alpha) &= 4\cos(\alpha)^3 - 3\cos(\alpha) = \frac{2}{5\sqrt{5}}
\end{aligned}$$

- On a $\phi'(B)(Y) = \theta'(B)(\varepsilon$, avec $\phi'(z) = 1 - 0,7z$ et $\theta'(z) = 1 + 0,4z$. Les racines de ces deux polynômes étant de module strictement supérieur à 1, on en déduit que la représentation est causale et inversible, et que ε est le bruit blanc d'innovation de X. De plus,

  $$\frac{\theta'(z)}{\phi'(z)} = (1+0,4z)\sum_{u\geq 0}(0,7)^u z^u = 1 + \sum_{u\geq 1} 1,1(0,7)^{u-1} z^u$$

  Il résulte de ce calcul la décomposition de Wold de Y :

  $$X_t = \phi'^{-1}(B) \circ \theta'(B)(\varepsilon)_t = \varepsilon_t + \sum_{u\geq 1} 1,1(0,7)^{u-1}\varepsilon_{t-u}$$

  Grâce à quoi on peut déterminer la fonction d'autocovariance de Y : pour tout $t \geq 0$,

  $$\gamma_Y(t) = \text{cov}(Y_t, Y_0) = \text{cov}\left(\varepsilon_t + \sum_{u\geq 1} 1,1(0,7)^{u-1}\varepsilon_{t-u}, \varepsilon_0 + \sum_{v\geq 1} 1,1(0,7)^{v-1}\varepsilon_{-v}\right)$$

D'où

$$
\begin{aligned}
\gamma_X(0) &= \sigma^2\left(1+\sum_{u\geq 1}1,1^2(0,7)^{2v-2}\right) = \sigma^2\left(1+\frac{1,1^2}{1-0,49}\right) = \sigma^2\frac{172}{51} \\
\gamma_X(t) &= \sigma^2\left(1,1(0,7)^{t-1}+\sum_{v\geq 1}1,1^2(0,7)^{t+2v-2}\right) \\
&= \sigma^2(0,7)^{t-1}\left(1,1+1,1^2\frac{0,7}{1-0,49}\right) = \sigma^2(0,7)^{t-1}\frac{140,8}{51}
\end{aligned}
$$

En particulier, $\rho_Y(1) = 140,5/172 \simeq 0,82$.

- On a $Z = \theta''(B)(\varepsilon)$ avec $\theta''(z) = 1+0,4z-0,5z^2$. La représentation est causale de nature pour un processus MA. Comme les racines de ce polynôme sont $2\left(-0,2\pm\sqrt{0,54}\right)$, de module strictement supérieur à 1, la représentation est aussi inversible, et le bruit blanc ε est donc le bruit blanc d'innovation de Z. La fonction d'autocovariance vaut

$$
\begin{aligned}
\gamma_Z(0) &= \operatorname{var}(\varepsilon_t+0,4\varepsilon_{t-1}-0,5\varepsilon_{t-2}) = \sigma^2\left(1+0,4^2+0,5^2\right) = 1,41\sigma^2 \\
\gamma_Z(1) &= \operatorname{cov}(\varepsilon_t+0,4\varepsilon_{t-1}-0,5\varepsilon_{t-2},\varepsilon_{t+1}+0,4\varepsilon_t-0,5\varepsilon_{t-1}) \\
&= \sigma^2(0,4-0,5\times 0,4) = 0,2\sigma^2 \\
\gamma_Z(2) &= \operatorname{cov}(\varepsilon_t+0,4\varepsilon_{t-1}-0,5\varepsilon_{t-2},\varepsilon_{t+2}+0,4\varepsilon_{t+1}-0,5\varepsilon_t) \\
&= -0,5\sigma^2
\end{aligned}
$$

les autres valeurs de la fonction d'autocovariance étant nulles, car Z est un MA(2).

2. Calculons les densités spectrales :
   - Cas de X :

$$
\begin{aligned}
f_X(\lambda) &= \frac{\sigma^2}{2\pi}\frac{1}{\left|\phi(e^i\lambda)\right|^2} \\
&= \frac{\sigma^2}{2\pi}\frac{1}{\left(1+1,6e^{i\lambda}+0,8e^{2i\lambda}\right)\left(1+1,6e^{-i\lambda}+0,8e^{-2i\lambda}\right)} \\
&= \frac{\sigma^2}{2\pi}\frac{1}{4,2+5,76\cos(\lambda)+1,6\cos(2\lambda)} \\
&= \frac{\sigma^2}{2\pi}\frac{1}{2,6+5,76\cos(\lambda)+3,2\cos(\lambda)^2}
\end{aligned}
$$

La dérivée du polynôme $2,6+5,766x+3,2x^2$ s'annule en $x_0 = -5,76/6,4$, ce qui correspond à $\theta_0 = \arccos(x_0) = 2,69$, soit à une fréquence $\varphi_0 = 2,69/2\pi = 0,43$. Compte-tenu des sens de variation, on en déduit que le densité spectrale est croissante entre 0 et $0,43\times 2\pi$, et décroissante entre $0,43\times 2\pi$ et π. En conséquence, on identifie le graphe B comme étant la densité spectrale de X.

– Cas de Y :

$$\begin{aligned} f_Y(\lambda) &= \frac{\sigma^2}{2\pi}\left|\frac{\theta'(e^{i\lambda})}{\phi'(e^{i\lambda})}\right|^2 \\ &= \frac{\sigma^2}{2\pi}\frac{\left(1+0,4e^{i\lambda}\right)\left(1+0,4e^{-i\lambda}\right)}{\left(1-0,7e^{i\lambda}\right)\left(1-0,7e^{-i\lambda}\right)} \\ &= \frac{\sigma^2}{2\pi}\frac{1,16+0,8\cos(\lambda)}{1,49-1,4\cos(\lambda)} \end{aligned}$$

La fraction rationnelle

$$\frac{1,16+0,8x}{1,49-1,4x} = cte + \frac{cte}{1,49-1,4x}$$

est monotone sur $]-\infty, 1,49/1,4[$, en pratique croissante. On en déduit que f_Y est décroissante sur $[0,\pi]$. On identifie donc le graphe C comme celui de f_Y.

– Par élimination, le graphe A est celui de la densité spectrale de Z... Vérifions-le :

$$\begin{aligned} f_Z(\lambda) &= \frac{\sigma^2}{2\pi}\left|\theta''(e^{i\lambda})\right|^2 \\ &= \frac{\sigma^2}{2\pi}\left(1+0,4e^{i\lambda}-0,5e^{2i\lambda}\right)\left(1+0,4e^{-i\lambda}-0,5e^{-2i\lambda}\right) \\ &= \frac{\sigma^2}{2\pi}\left(1,41+0,4\cos(\lambda)-\cos(2\lambda)\right) \\ &= \frac{\sigma^2}{2\pi}\left(2,41+0,4\cos(\lambda)-2\cos(\lambda)^2\right) \end{aligned}$$

La dérivée du polynôme $2,41+0,4x-2x^2$ s'annule en $x_1 = 0,1$, ce qui correspond à $\theta_1 = \arccos(0,1) = 1,47$ soit à une fréquence $\varphi_1 = 1,47/2\pi = 0,23$. Compte-tenu des sens de variation, on en déduit que f_Z croît entre 0 et $0,23 \times 2\pi$, puis décroît entre $0,23 \times 2\pi$ et π. Ce qui correspond bien au graphe A.

3. – La série temporelle A a une allure presque non stationnaire, quasiment de marche aléatoire, la régularité ne pouvant apparaître que sur de longues périodes : cela correspond à une densité spectrale prenant ses plus fortes valeurs à l'origine, pour des fréquences basses. La série temporelle correspond donc à la densité spectrale C, et donc au processus Y.
   – Les séries temporelles B et C ont toutes deux un comportement périodique marqué, la période étant plus courte pour la série C, ce qui correspond à une fréquence plus élévée. Comme le pic de la densité spectrale B a lieu en une fréquence plus grande que le pic de la densité spectrale A, on en déduit que la série temporelle C correspond à la densité spectrale B, et donc au processus X. La série temporelle B correspond à la densité spectrale A et au processus Z.
4. – L'autocorrélogramme A est remarquable surtout par les deux très fortes valeurs prises par l'autocorrélation partielle (estimée) en 1 et 2, les valeurs suivantes étant peu significativement différentes de 0. L'autocorrélation (estimée) est proche de 0 à partir du rang 3 également, mais

la valeur en 2, nettement positive interdit qu'il s'agisse là de l'autocorrélogramme du processus MA(2) Z : en effet, $\rho_Z(2) = \gamma_Z(2)/\gamma_Z(0) = -0,36$. D'autre part, $\rho_Y(1) \simeq 0,82$ est très nettement positive, ce qui ne concorde pas avec l'autocorrélogramme A. Il paraît donc vraisemblable que cet autocorrélogramme corresponde au processus AR(2) X, dont la valeur théorique $\rho_X(1) \simeq -0,89$ paraît effectivement proche de la valeur ici estimée.

- L'autocorrélogramme B peut être l'autocorrélogramme d'un processus MA(2), car l'autocorrélation (estimée) est proche de 0 à partir du rang 3. De plus, la valeur en 2 de l'autocorrélation (estimée) est de l'ordre de $-0,4$, ce qui est effectivement proche de la valeur théorique $\rho_Z(2) = -0,36$. On en déduit que l'autocorrélogramme B correspond au processus Z.
- Par élimination, l'autocorrélogramme C correspond au processus Y, à la série temporelle A. Notons, pour achever de s'en convaincre, que l'autocorrélation (estimée) en 1 est de l'ordre de $0,8$, effectivement proche de la valeur théorique $\rho_Y(1) = 0,82$. D'autre part, une décroissance régulière de la fonction d'autocorrélation (estimée) peut être le signe d'une allure proche d'une marche aléatoire.

Ex 3.

1. (a) Soit $\hat{\rho}_X$ l'estimateur empirique de la fonction d'autocorrélation de X :

$$\forall t = 0, \ldots, N-1, \quad \hat{\rho}_X(t) = \frac{\frac{1}{N-t}\sum_{u=1}^{N-t}(X_u - \bar{X}_N)(X_{u+t} - \bar{X})}{\frac{1}{N}\sum_{u=1}^{N}(X_u - \bar{X}_N)^2}$$

qu'on peut remplacer ici par

$$\forall t = 0, \ldots, N-1, \quad \hat{\rho}_X(t) = \frac{\frac{1}{N-t}\sum_{u=1}^{N-t} X_u X_{u+t}}{\frac{1}{N}\sum_{u=1}^{N} X_u^2}$$

dans la mesure où l'on sait que le processus X est centré. Si X est un processus AR(p), tel que $X_t + \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} = \varepsilon_t$, alors l'estimateur préliminaire de $(\varphi_1, \ldots, \varphi_p)$ vérifie les équations de Yule-Walker

$$\begin{pmatrix} 1 & \hat{\rho}_X(1) & \cdots & \hat{\rho}_X(p-1) \\ \hat{\rho}_X(1) & 1 & \ddots & \hat{\rho}_X(p-2) \\ \vdots & \ddots & \ddots & \vdots \\ \hat{\rho}_X(p-1) & \hat{\rho}_X(p-2) & \cdots & 1 \end{pmatrix} \begin{pmatrix} \hat{\varphi}_1 \\ \hat{\varphi}_2 \\ \vdots \\ \hat{\varphi}_p \end{pmatrix} = \begin{pmatrix} \hat{\rho}_X(1) \\ \hat{\rho}_X(2) \\ \vdots \\ \hat{\rho}_X(p) \end{pmatrix}$$

Ici $p = 1$, et on en déduit

$$\hat{\varphi} = \hat{\rho}_X(1) = \frac{N}{N-1} \frac{\sum_{t=2}^{N} X_t X_{t-1}}{\sum_{t=1}^{N} X_t^2}.$$

(b) Par définition, $\tilde{\varepsilon}_1 = X_1$. Pour $t \geq 2$,

$$\begin{aligned} \tilde{\varepsilon}_t &= X_t - P^{\perp}_{V^2(1,X_1,\ldots,X_{t-1})}(X_t) \\ &= X_t - P^{\perp}_{V^2(1,X_1,\ldots,X_{t-1})}(\varepsilon_t) - \varphi P^{\perp}_{V^2(1,X_1,\ldots,X_{t-1})}(X_{t-1}) \\ &= X_t - 0 - \varphi X_{t-1} = \varepsilon_t \end{aligned}$$

car $\varepsilon_t \perp V^2(1, X_1, \ldots, X_{t-1})$: en effet, ε est le bruit blanc d'innovation associé à X car on a supposé $|\varphi| < 1$, ce qui implique que la représentation ARMA est causale et inversible.

(c) Notons $\delta_t = \mathrm{var}(\tilde{\varepsilon}_t)/\sigma^2$. L'estimateur des moindres carrés ordinaires de φ est la valeur $\hat{\varphi}_{mco}$ qui minimise

$$\varphi \mapsto \frac{1}{N}\sum_{t=1}^{N}\frac{\tilde{\varepsilon}_t^2}{\delta_t}$$

où $\tilde{\varepsilon}_t$ est vu comme fonction de φ, $X_1, \ldots, X_t$. Or $\tilde{\varepsilon}_1 = X_1$, $\delta_1 = \mathrm{var}(X_1)/\sigma^2 = 1/(1-\varphi^2)$ selon un calcul habituel, $\tilde{\varepsilon}_t = X_t - \varphi X_{t-1}$ et $\delta_t = 1$ quel que soit $t > 1$. On en déduit que

$$\sum_{t=1}^{N}\frac{\tilde{\varepsilon}_t^2}{\delta_t} = (1-\varphi^2)X_1^2 + \sum_{t=2}^{N}(X_t - \varphi X_{t-1})^2 = \sum_{t=1}^{N}X_t^2 - 2\varphi\sum_{t=2}^{N}X_tX_{t-1} + \varphi^2\sum_{t=2}^{N-1}X_t^2$$

Il s'agit donc d'un polynôme du second degré en φ, qui atteint son minimum en

$$\hat{\varphi}_{mco} = \frac{\sum_{t=2}^{N}X_tX_{t-1}}{\sum_{t=2}^{N-1}X_t^2}$$

On en déduit l'estimateur de σ^2 :

$$\begin{aligned}
\hat{\sigma}^2_{mco} &= \frac{1}{N}\sum_{t=1}^{N}\frac{\tilde{\varepsilon}_t^2}{\delta_t}\Big|_{\varphi=\hat{\varphi}_{mco}} \\
&= \frac{1}{N}\left(\sum_{t=1}^{N}X_t^2 - \frac{\left(\sum_{t=2}^{N}X_tX_{t-1}\right)^2}{\sum_{t=2}^{N-1}X_t^2}\right)
\end{aligned}$$

(d) L'estimateur des moindres carrés conditionnels de φ est la valeur $\hat{\varphi}_{mcc}$ qui minimise

$$\varphi \mapsto \frac{1}{N}\sum_{t=1}^{N}\varepsilon_t^2$$

avec $\varepsilon_t = X_t - \varphi X_{t-1}$, et en supposant $X_u = 0$ quel que soit $u \leq 0$. D'où

$$\sum_{t=1}^{N}\varepsilon_t^2 = X_1^2 + \sum_{t=2}^{N}(X_t - \varphi X_{t-1})^2 = \sum_{t=1}^{N}X_t^2 - 2\varphi\sum_{t=2}^{N}X_tX_{t-1} + \varphi^2\sum_{t=1}^{N-1}X_t^2$$

Il s'agit d'un polynôme en φ qui atteint son minimum en

$$\hat{\varphi}_{mcc} = \frac{\sum_{t=2}^{N}X_tX_{t-1}}{\sum_{t=1}^{N-1}X_t^2}$$

Comparaison de $\hat{\varphi}_{mco}$ et de $\hat{\varphi}_{mcc}$:

$$\hat{\varphi}_{mco} - \hat{\varphi}_{mcc} = \sum_{t=2}^{N}X_tX_{t-1}\frac{X_1^2}{\sum_{t=2}^{N-1}X_t^2\sum_{t=1}^{N-1}X_t^2} \simeq N\gamma_X(1)\frac{X_1^2}{N^2\gamma_X(0)^2} = \frac{X_1^2}{N}\frac{\gamma_X(1)}{\gamma_X(0)^2}$$

L'écart est donc d'ordre $1/N$.

(e) L'estimateur du maximum de vraisemblance de φ dans le cas gaussien est la valeur $\hat{\varphi}_{mv}$ qui minimise

$$\varphi \mapsto \frac{1}{N}\sum_{t=1}^{N}\frac{\tilde{\varepsilon}_t^2}{\delta_t}+\frac{1}{N}\left(\prod_{t=1}^{N}\delta_t\right)^{1/N} = \frac{1}{N}\left(\sum_{t=1}^{N}X_t^2 - 2\varphi\sum_{t=2}^{N}X_tX_{t-1} + \varphi^2\sum_{t=2}^{N-1}X_t^2 + (1-\varphi^2)^{1/N}\right)$$

Le point où la fonction précédente s'annule correspond à la valeur $\hat{\varphi}_{mv}$. Pour comparer $\hat{\varphi}_{mco}$ et $\hat{\varphi}_{mv}$, on utilise les inégalités suivantes :

$$\begin{aligned}
&\frac{1}{N}\left(\sum_{t=1}^{N}X_t^2 - 2\hat{\varphi}_{mco}\sum_{t=2}^{N}X_tX_{t-1} + \hat{\varphi}_{mco}^2\sum_{t=2}^{N-1}X_t^2\right)\\
&\leq \quad \frac{1}{N}\left(\sum_{t=1}^{N}X_t^2 - 2\hat{\varphi}_{mv}\sum_{t=2}^{N}X_tX_{t-1} + \varphi_{mv}^2\sum_{t=2}^{N-1}X_t^2\right) \text{ car } \hat{\varphi}_{mco} \text{ est minimisant}\\
&< \quad \frac{1}{N}\left(\sum_{t=1}^{N}X_t^2 - 2\varphi\sum_{t=2}^{N}X_tX_{t-1} + \varphi^2\sum_{t=2}^{N-1}X_t^2 + (1-\varphi^2)^{1/N}\right) \text{ car } 0 < (1-\varphi^2)^{1/N}\\
&< \quad \frac{1}{N}\left(\sum_{t=1}^{N}X_t^2 - 2\varphi\sum_{t=2}^{N}X_tX_{t-1} + \varphi^2\sum_{t=2}^{N-1}X_t^2 + 1\right) \text{ car } (1-\varphi^2)^{1/N} < 1
\end{aligned}$$

En mettant ces polynômes du second degré sous forme canonique, et en utilisant le fait que $\hat{\varphi}_{mco} = \frac{\sum_{t=2}^{N} X_tX_{t-1}}{\sum_{t=2}^{N-1} X_t^2}$, on en déduit

$$0 < (\hat{\varphi}_{mv} - \hat{\varphi}_{mco})^2 < \frac{1}{\sum_{t=2}^{N-1} X_t^2}$$

D'où le résultat attendu

$$|\hat{\varphi}_{mv} - \hat{\varphi}_{mco}| < \frac{1}{\sqrt{\sum_{t=2}^{N-1} X_t^2}}$$

Notons que ce majorant est de l'ordre de $1/\sqrt{N\gamma_X(0)}$.

2. Établissons une équation de récurrence, ou fonction de prévision : quel que soit $h > 0$,

$$\begin{aligned}
\check{X}_{N+h} &= P^{\perp}_{V^2(X_1,\ldots,X_N)}(X_{N+h})\\
&= P^{\perp}_{V^2(X_1,\ldots,X_N)}(\sigma_{N+h} + \varphi X_{N+h-1})\\
&= 0 + \varphi P^{\perp}_{V^2(X_1,\ldots,X_N)}(X_{N+h-1})
\end{aligned}$$

car $\varepsilon_{N+h} \perp V^2(X_1,\ldots,X_N)$ car ε est le bruit blanc d'innovation.. D'où, par récurrence,

$$\check{X}_{N+h} = \varphi^h P^{\perp}_{V^2(X_1,\ldots,X_N)}(X_N) = \varphi^h X_N$$

Ici, X_N est l'unique valeur pivôtale. Pour calculer l'erreur de prédiction, on utilise la représentation MA(∞) de X :

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

D'où

$$X_{N+h} - \check{X}_{N+h} = \sum_{u=0}^{+\infty} \varphi^u \varepsilon_{N+h-u} - \sum_{u=h}^{+\infty} \varphi^u \varepsilon_{N+h-u} = \sum_{u=0}^{h-1} \varphi^u \varepsilon_{N+h-u}$$

On en déduit

$$\operatorname{var}\left(X_{N+h} - \check{X}_{N+h}\right) = \sigma^2 \sum_{u=0}^{h-1} \varphi^{2u} = \sigma^2 \frac{\varphi^{2h}-1}{\varphi^2-1}.$$

3. (a) Quels que soient s et t,

$$\begin{aligned} \mathbb{E}[Y_t] &= \mathbb{E}[X_{-t}] = 0 \\ \gamma_Y(s,t) &= \operatorname{cov}(Y_s, Y_t) = \operatorname{cov}(X_{-s}, X_{-t}) = \gamma_X(|-s-(-t)|) = \gamma_X(|t-s|) = \gamma_X(s,t) \end{aligned}$$

En particulier, quel que soit h, $\gamma_Y(s+h,t+h) = \gamma_X(s+h,t+h) = \gamma_X(s,t) = \gamma_Y(s,t)$. Le processus Y est donc de moyenne constante, et de fonction d'autocovariance invariante par translation temporelle : c'est donc un processus faiblement stationnaire.

(b) Le processus Y ayant mêmes caractéristiques du second ordre que X, il est de même nature que X (au sens faible) : c'est donc un processus AR(1) vérifiant $Y_t = \varepsilon'_t + \varphi Y_{t-1}$, avec ε'_t bruit blanc faible d'innovation de variance σ^2. D'où

$$\begin{aligned} \varepsilon'_t &= Y_t - \varphi Y_{t-1} = X_{-t} - \varphi X_{-t+1} \\ &= -\varphi\varepsilon_{-t+1} + (1+\varphi^2)X_{-t} = -\varphi\varepsilon_{-t+1} + (1+\varphi^2)\sum_{u\geq 0}\varphi^u \varepsilon_{-t-u} \end{aligned}$$

(c)

$$\begin{aligned} \mathbb{E}[\varepsilon_t] &= 1\mathbb{P}(\varepsilon_t = 1) - 1\mathbb{P}(\varepsilon_t = -1) = 0 \\ \operatorname{var}(\varepsilon_t) &= \mathbb{E}\left[\varepsilon_t^2\right] = 1^2\mathbb{P}(\varepsilon_t = 1) + (-1)^2\mathbb{P}(\varepsilon_t = -1) = 1 \end{aligned}$$

(d) Les calculs faits précédemment pour X sont valables pour Y, puisque ces deux processus ont mêmes caractéristiques du second ordre. L'erreur de prédiction au pas h vaut donc

$$\operatorname{var}\left(Y_{N+h} - \check{Y}_{N+h}\right) = \frac{1-\frac{1}{3^h}}{1-\frac{1}{3}}$$

(e) On a

$$X_t = \sum_{u\geq 0} \frac{1}{3^u}\varepsilon_{t-u}$$

avec ε_s ne prenant que 1 et -1 comme valeurs. On en déduit que la partie entière de

$$3^{h-1}Y_N = 3^{h-1}X_{-N} = 3^{h-1}\sum_{u\geq 0}\frac{1}{3^u}\varepsilon_{-N-u} = \sum_{u\geq -h+1}\frac{1}{3^u}\varepsilon_{-N-h-u+1}$$

vaut $\sum_{u=-h+1}^{0} \frac{1}{3^u}\varepsilon_{-N-h-u+1}$, et que la partie fractionnaire restante vaut

$$\sum_{u\geq +1} \frac{1}{3^u}\varepsilon_{-N-h-u+1} = \frac{1}{3}\sum_{u\geq 0} \frac{1}{3^u}\varepsilon_{-N-h-u} = \frac{1}{3}X_{-N-h} = \frac{1}{3}Y_{N+h}$$

Cela montre que Y_{N+h} est fonction de Y_N.

(f) Du résultat qui précède, on peut évidemment déduire que le meilleur prédicteur de Y_{N+h} sachant $Y_1, \ldots, Y_N$ au sens des moindres carrés vaut

$$\mathbb{E}[Y_{N+h}|Y_1, \ldots, Y_N] = Y_{N+h}$$

puisque Y_{N+h} est fonction de $Y_1, \ldots, Y_N$. L'erreur de prédiction est alors nulle. Il n'y a pas de contradiction avec les résultats précédents, car $\check{Y}_{N+h}$ est le meilleur prédicteur de Y_{N+h} sachant $Y_1, \ldots, Y_N$ au sens des moindres carrés *parmi les combinaisons linéaires de* $Y_1, \ldots, Y_N$. Or Y_{N+h} est certes fonction de $Y_1, \ldots, Y_N$, mais n'en est pas une fonction linéaire.