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 du Partiel du 2 avril 2003

Ex 1.

1. Voir cours.
2. (a)

$$\begin{aligned}
\mathbb{P}(\varepsilon_t\varepsilon_{t+1} < 0) &= \mathbb{P}\left((\varepsilon_t < 0, \varepsilon_{t+1} > 0) \cap (\varepsilon_t > 0, \varepsilon_{t+1} < 0)\right) \\
&= \mathbb{P}\left(\varepsilon_t < 0, \varepsilon_{t+1} > 0\right) + \mathbb{P}\left(\varepsilon_t > 0, \varepsilon_{t+1} < 0\right) \\
&\quad \text{car les événements sont disjoints} \\
&= \mathbb{P}\left(\varepsilon_t < 0\right)\mathbb{P}\left(\varepsilon_{t+1} > 0\right) + \mathbb{P}\left(\varepsilon_t > 0\right)\mathbb{P}\left(\varepsilon_{t+1} < 0\right) \\
&\quad \text{car } \varepsilon_t \text{ et } \varepsilon_{t+1} \text{ sont indépendants} \\
&= \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2}
\end{aligned}$$

   (b) Notons $\eta_p = 1_{\varepsilon_{2p-1}\varepsilon_{2p}<0}$. D'après la question précédente, $(\eta_p, p = 1, \ldots, N)$ est une famille de variables de Bernouilli de paramètre $1/2$. Comme elles sont indépendantes, leur somme $T = \sum_{p=1}^{N} \eta_p$ suit une loi binômiale de paramètres $1/2$ et N.

   (c) Soit $X = (X_1, \ldots, X_{2N})$ une famille de variables aléatoires dont x est une réalisation. On souhaite tester H_0 : X est un bruit blanc gaussien, contre H_1 : X n'est pas un bruit blanc gaussien. Sous H_0, $T_X = \sum_{p=1}^{N} 1_{X_{2p-1}X_{2p}<0}$ suit une loi binomiale de paramètre $1/2$ et N. En particulier, il est de moyenne $N/2$ et de variance $N/4$. Il résulte du théorème central limite que pour N grand,

$$\frac{T_X - \frac{N}{2}}{\sqrt{\frac{N}{4}}} \sim \mathcal{N}(0,1)$$

   On en déduit que sous H_0,

$$P\left(\left|\frac{T_X - \frac{N}{2}}{\sqrt{\frac{N}{4}}} > \phi_{1-\alpha/2}\right|\right) \simeq \alpha$$

   avec $\phi_{1-\alpha/2}$ le fractile d'ordre $1 - \alpha/2$ de la loi gaussienne centrée réduite. Le test de zone de rejet

$$\sqrt{\frac{N}{4}} > \phi_{1-\alpha/2}$$

   est donc un test de niveau α de l'hypothèse H_0.

Ex 2.

1. X est un processus AR(1), vérifiant l'équation $\phi(B)(X) = \varepsilon$, avec $\phi(B) = I + \frac{3}{2}B$. Sa représentation est inversible, comme toute représentation d'un processus AR. La transformée en z du filtre linéaire $\phi(B)$ est $\phi(z) = 1+\frac{3}{2}z$. Elle s'annule en $\frac{2}{3}$ de module strictement inférieur à 1. Le filtre est donc bien inversible, mais la représentation n'est pas causale.

   Posons $\tilde{\varepsilon} = \left(I - \frac{2}{3}B\right) \circ \left(I - \frac{3}{2}B\right)^{-1}(\varepsilon)$. D'après une proposition du cours, $\tilde{\varepsilon}$ est un bruit blanc de variance

$$\tilde{\sigma}^2 = \left(\frac{2}{3}\right)^2 \sigma^2$$

   De plus,

$$\begin{aligned}
\tilde{\varepsilon} &= \left(I - \frac{2}{3}B\right) \circ \left(I - \frac{3}{2}B\right)^{-1}(\varepsilon) \\
&= \left(I - \frac{2}{3}B\right) \circ \left(I - \frac{3}{2}B\right)^{-1} \circ \left(I + \frac{3}{2}B\right)(X) \\
&= \left(I - \frac{2}{3}B\right)(X)
\end{aligned}$$

   Cette nouvelle représentation du processus X est causale (et toujours inversible) car la transformée en z de $\left(I - \frac{2}{3}B\right)$ ne s'annule pas sur le disque unité. On en déduit que $\tilde{\varepsilon}$ est le bruit blanc d'innovation associé à X.

   Représentation MA(∞) : le filtre linéaire $\left(I - \frac{2}{3}B\right)$ est inversible ; et son inverse a pour transformée en z

$$\frac{1}{1+\frac{2}{3}z} = \sum_{u\geq 0}\left(-\frac{2}{3}\right)^u z^u$$

   On en déduit

$$\begin{aligned}
\left(I - \frac{2}{3}B\right)^{-1} &= \sum_{u\geq 0}\left(-\frac{2}{3}\right)^u B^u \\
X_t &= \left(I - \frac{2}{3}B\right)^{-1}(\varepsilon)_t = \sum_{u\geq 0}\left(-\frac{2}{3}\right)^u B^u(\varepsilon)_t \\
&= \varepsilon_t + \sum_{u\geq 1}\left(-\frac{2}{3}\right)^u \varepsilon_{t-u}
\end{aligned}$$

   On ne peut déterminer une représentation MA(∞) d'un processus qu'en partant d'une représentation causale (de même qu'on ne peut calculer une représentation AR(∞) qu'en partant d'une représentation inversible). S'il est vrai que $(1+\frac{3}{2}z)^{-1} = \sum_{u\geq 0}\left(-\frac{3}{2}\right)^u z^u$ (pour $|z| < \frac{2}{3}$), il est faux que $(1+\frac{3}{2}B)^{-1} = \sum_{u\geq 0}\left(-\frac{3}{2}\right)^u B^u$ car $\sum_{u\geq 0}\left(-\frac{3}{2}\right)^u B^u$ n'est pas un filtre linéaire ($\sum_{u\geq 0}\left|-\frac{3}{2}\right|^u = +\infty$) et il est faux que $X = \sum_{u\geq 0}\left(-\frac{3}{2}\right)^u \varepsilon_{t-u}$ car cette série n'est pas convergente. En revanche, il est vrai que $(1+\frac{3}{2}B)^{-1} = \sum_{u>0}\left(-\frac{2}{3}\right)^u B^{-u}$ et que $X = \sum_{u>0}\left(-\frac{2}{3}\right)^u \varepsilon_{t+u}$, mais il ne s'agit pas d'une représentation MA(∞).

   Fonction d'autocovariance : dans ce qui suit, on suppose $h \geq 0$.

– Méthode 1 : à partir de la représentation MA(∞).

$$\begin{aligned}
\gamma_X(h) &= \operatorname{cov}(X_0, X_h) = \operatorname{cov}\left(\sum_{u\geq 0}\left(-\frac{2}{3}\right)^u \tilde{\varepsilon}_{-u}, \sum_{u\geq 0}\left(-\frac{2}{3}\right)^u \tilde{\varepsilon}_{h-u}\right)\\
&= \sum_{u\geq 0}\sum_{v\geq 0}\left(-\frac{2}{3}\right)^{u+v} \operatorname{cov}(\tilde{\varepsilon}_{-u}, \tilde{\varepsilon}_{h-v})\\
&= \sum_{u\geq 0}\sum_{v\geq 0}\left(-\frac{2}{3}\right)^{u+v} \tilde{\sigma}^2 1_{-u=h-v}\\
&= \frac{4}{9}\sigma^2 \sum_{u\geq 0}\left(-\frac{2}{3}\right)^{2u+h} \qquad \text{avec } v = u+h\\
&= \frac{4}{9}\left(-\frac{2}{3}\right)^h \frac{1}{1-\left(-\frac{2}{3}\right)^2}\sigma^2\\
&= \frac{5}{4}\left(-\frac{2}{3}\right)^h \sigma^2
\end{aligned}$$

Deux remarques :

– *on n'a pas justifié la possibilité de commuter sommation et covariance, car cela a été vu en cours*

– *trop d'erreurs ont été commises dans la manipulation des indices de sommation : de même que $(a+b)(c+d) \neq ac+bd$, de même*

$$\operatorname{cov}\left(\sum_{u\geq 0}\left(-\frac{2}{3}\right)^u \tilde{\varepsilon}_{-u}, \sum_{u\geq 0}\left(-\frac{2}{3}\right)^u \tilde{\varepsilon}_{h-u}\right) \neq \sum_{u\geq 0}\left(-\frac{2}{3}\right)^{2u} \operatorname{cov}\left(\tilde{\varepsilon}_{-u}, \tilde{\varepsilon}_{h-u}\right).$$

– Méthode 2 : à partir de la représentation causale et inversible.

$$\begin{aligned}
\gamma_X(0) &= \operatorname{var}(X_t) = \operatorname{var}\left(\tilde{\varepsilon}_t - \frac{2}{3}X_{t-1}\right)\\
&= \operatorname{var}(\tilde{\varepsilon}_t) + \frac{4}{9}\operatorname{var}(X_{t-1}) \ \text{ car } \tilde{\varepsilon}\perp 1, X_{t-1}\\
&= \frac{\tilde{\sigma}^2}{1-\frac{4}{9}} = \frac{4}{5}\sigma^2\\
\gamma_X(h+1) &= \operatorname{cov}(X_t, X_{t+h+1}) = \operatorname{cov}\left(X_t, \tilde{\varepsilon}_{t+h+1} - \frac{2}{3}X_{t+h}\right)\\
&= -\frac{2}{3}\operatorname{cov}(X_t, X_{t+h}) \ \text{ car } \tilde{\varepsilon}_{t+h+1}\perp 1, X_t \text{ si } h\geq 0\\
&= -\frac{2}{3}\gamma_X(h) = \left(-\frac{2}{3}\right)^{h+1}\gamma_X(0) = \frac{4}{5}\left(-\frac{2}{3}\right)^{h+1}\sigma^2
\end{aligned}$$

Noter qu'il est ici essentiel d'utiliser le bruit blanc d'innovation (sinon $\tilde{\varepsilon}\perp 1, X_{t-1}$ et $\tilde{\varepsilon}_{t+h+1}\perp 1, X_t$ seraient faux)

2. – Y est un processus du second ordre, car pour tout $t \in \mathbb{Z}$, Y_t est somme de deux variables de carré intégrable, et donc est de carré intégrable.

– quel que soit $t \in \mathbb{Z}$, $\mathbb{E}[Y_t] = \mathbb{E}[X_t] + \mathbb{E}[\varepsilon'_t] = \mathbb{E}[X_t]$. Or X est un processus stationnaire, de moyenne constante (on pourrait vérifier qu'elle est nulle). La fonction moyenne de Y est donc elle-aussi constante.
– quels que soient $s, t \in \mathbb{Z}$, $\gamma_Y(s,t) = \mathrm{cov}(Y_s, Y_t) = \mathrm{cov}(X_s + \varepsilon'_s, X_t + \varepsilon'_t)$. Or X_s et X_t sont fonction du bruit blanc ε, qui est indépendant de ε'. X_s et X_t sont donc indépendants de ε'_s et ε'_t. On en déduit $\gamma_Y(s,t) = \mathrm{cov}(X_s, X_t) + \mathrm{cov}(\varepsilon'_s, \varepsilon'_t) = \gamma_X(|s-t|) + \gamma_{\varepsilon'}(|s-t|)$. La fonction d'autocovariance de Y en (s,t) ne dépend donc que de $|s-t|$; elle est alors invariante par translation temporelle : $\gamma_Y(s,t) = \gamma_Y(s+u, t+u)$ quel que soit u. Ce qui achève de montrer que Y est un processus faiblement stationnaire. En outre,

$$\begin{aligned}\gamma_Y(h) &= \frac{4}{5}\sigma^2 + \sigma'^2 \text{ si } h = 0 \\ &= \frac{4}{5}\left(-\frac{2}{3}\right)^h \sigma_\varepsilon^2 \text{ si } h > 0\end{aligned}$$

Trop souvent il a été écrit que $\gamma_Y(h) = 0$ pour tout $h \neq 0$; or cela est caractéristique d'un bruit blanc, ce que ne sont ni Y, ni Z.

3. Z est un processus ARMA$(1,1)$ vérifiant $\varphi(B)(Z) = \theta(B)(\varepsilon'')$, avec $\varphi(B) = I + \varphi B$ et $\theta(B) = I + \theta B$. Les transformées en z $\varphi(z) = 1 + \varphi z$ et $\theta(z) = 1 + \theta z$ ne s'annulent pas sur le disque unité. La représentation est donc causale et inversible, et ε'' est le bruit blanc d'innovation de Z.

   Pour calculer la fonction d'autocovariance de Z, on peut là encore procéder de deux méthodes :
   – Méthode 1 : à partir de la représentation MA(∞). Le filtre linéaire $\varphi(B)$ est inversible, et la transformée en z de $\varphi(B)^{-1} \circ \theta(B)$ vaut

$$\frac{\theta(z)}{\varphi(z)} = \frac{1+\theta z}{1+\varphi z} = 1 + \frac{(\theta-\varphi)z}{1+\varphi z} = 1 + (\theta - \varphi)\sum_{u\geq 1}(-\varphi)^{u-1}z^u$$

   On en déduit

$$\begin{aligned}\varphi(B)^{-1} \circ \theta(B) &= I + (\theta-\varphi)\sum_{u\geq 1}(-\varphi)^{u-1}B^u \\ Z_t &= \varphi(B)^{-1} \circ \theta(B)(\varepsilon'')_t = \varepsilon''_t + (\theta-\varphi)\sum_{u\geq 1}(-\varphi)^{u-1}\varepsilon''_{t-u}\end{aligned}$$

   D'où le calcul de la fonction d'autocovariance :

$$\begin{aligned}\gamma_Z(0) &= \mathrm{var}(Z_0) = \mathrm{var}\left(\varepsilon''_0 + (\theta-\varphi)\sum_{u\geq 1}(-\varphi)^{u-1}\varepsilon''_{-u}\right) \\ &= \mathrm{var}\left(\varepsilon''_0\right) + (\theta-\varphi)^2\sum_{u\geq 1}(-\varphi)^{2u-2}\,\mathrm{var}\left(\varepsilon''_{-u}\right) \\ &= \left(1 + \frac{(\theta-\varphi)^2}{1-\varphi^2}\right)\sigma''\end{aligned}$$

$$
\begin{aligned}
\gamma_Z(h+1) &= \operatorname{cov}(Z_0, Z_{h+1}) \\
&= \operatorname{cov}\left(\varepsilon''_0 + (\theta-\varphi)\sum_{u\geq 1}(-\varphi)^{u-1}\varepsilon''_{-u}, \varepsilon''_{h+1} + (\theta-\varphi)\sum_{u\geq 1}(-\varphi)^{u-1}\varepsilon''_{h+1-u}\right) \\
&= \operatorname{cov}(\varepsilon''_0, \varepsilon''_{h+1}) + (\theta-\varphi)\sum_{u\geq 1}(-\varphi)^{u-1}\operatorname{cov}\left(\varepsilon''_0, \varepsilon''_{h+1-u}\right) \\
&\quad +(\theta-\varphi)\sum_{u\geq 1}(-\varphi)^{u-1}\operatorname{cov}\left(\varepsilon''_{-u}, \varepsilon''_{h+1}\right) \\
&\quad +(\theta-\varphi)^2\sum_{u\geq 1}\sum_{v\geq 1}(-\varphi)^{u+v-2}\operatorname{cov}(\varepsilon''_{-u}, \varepsilon''_{h+1-v}) \\
&= 0 + (\theta-\varphi)(-\varphi)^h\sigma'' + 0 + (\theta-\varphi)^2\sum_{u\geq 1}(-\varphi)^{2u-2+h+1}\sigma'' \\
&= \left((\theta-\varphi) - \frac{\varphi}{1-\varphi^2}(\theta-\varphi)^2\right)(-\varphi)^h\sigma''
\end{aligned}
$$

– Méthode 2 : à partir de la représentation causale et inversible.

$$
\begin{aligned}
\gamma_Z(0) &= \operatorname{var}(Z_t) = \operatorname{var}\left(\varepsilon''_t + \theta\varepsilon''_{t-1} - \varphi Z_{t-1}\right) \\
&= \operatorname{var}\left(\varepsilon''_t\right) + \theta^2\operatorname{var}\left(\varepsilon''_{t-1}\right) + \varphi^2\operatorname{var}\left(Z_{t-1}\right) - 2\theta\varphi\operatorname{cov}(\varepsilon''_{t-1}, Z_{t-1}) \text{ car } \varepsilon''_t \perp 1, Z_{t-1}, \varepsilon''_{t-1} \\
&= (1+\theta^2)\sigma'' + \varphi^2\gamma_Z(0) - 2\theta\varphi\operatorname{cov}(\varepsilon''_{t-1}, \varepsilon''_{t-1} + \theta\varepsilon''_{t-2} - \varphi Z_{t-2}) \\
&= (1+\theta^2)\sigma'' + \varphi^2\gamma_Z(0) - 2\theta\varphi\operatorname{var}(\varepsilon''_{t-1}) \text{ car } \varepsilon''_{t-1} \perp 1, Z_{t-2}, \varepsilon''_{t-2} \\
&= (1-2\theta\varphi+\theta^2)\sigma'' + \varphi^2\gamma_Z(0) = \frac{1-\theta\varphi+\theta^2}{1-\varphi^2}\sigma''
\end{aligned}
$$

$$
\begin{aligned}
\gamma_Z(1) &= \operatorname{cov}(Z_t, Z_{t+1}) = \operatorname{cov}(Z_t, \varepsilon''_{t+1} + \theta\varepsilon''_t - \varphi Z_t) \\
&= \theta\operatorname{cov}(Z_t, \varepsilon''_t) - \varphi\operatorname{var}(Z_t) \text{ car } \varepsilon''_{t+1} \perp 1, Z_t \\
&= \theta\operatorname{cov}(\varepsilon''_t + \theta\varepsilon''_{t-1} - \theta Z_{t-1}, \varepsilon''_t) - \varphi\gamma_Z(0) \\
&= \theta\operatorname{var}(\varepsilon''_t) - \varphi\gamma_Z(0) \text{ car } \varepsilon''_t \perp 1, Z_{t-1}\varepsilon''_{t-1} \\
&= \left(\theta - \varphi\frac{1-2\theta\varphi+\theta^2}{1-\varphi^2}\right)\sigma''
\end{aligned}
$$

$$
\begin{aligned}
\gamma_Z(h+2) &= \operatorname{cov}(Z_t, Z_{t+h+2}) = \operatorname{cov}(Z_t, \varepsilon''_{t+h+2} + \theta\varepsilon''_{t+h+1} - \varphi Z_{t+h+1}) \\
&= -\varphi\operatorname{cov}(Z_t, Z_{t+h+1}) \text{ car } \varepsilon''_{t+h+2}, \varepsilon''_{t+h+1} \perp 1, Z_t \\
&= -\varphi\gamma_Z(h+1) = (-\varphi)^h\gamma_Z(1)
\end{aligned}
$$

Si on compare les fonctions d'autocovariance de Y et Z, on voit qu'elles sont égales si et seulement si

$$
\begin{aligned}
\varphi &= \frac{2}{3} \\
\left(1 + \frac{(\theta-\varphi)^2}{1-\varphi^2}\right)\sigma'' &= \frac{4}{5}\sigma^2 + \sigma'^2 \\
\left((\theta-\varphi) - \frac{\varphi}{1-\varphi^2}(\theta-\varphi)^2\right)\sigma'' &= \frac{4}{5}
\end{aligned}
$$

Dans ce cas, comme les deux processus ont mêmes caractéristiques du second ordre (fonction moyenne et fonction d'autocovariance), ils sont de même nature, à savoir ARMA$(1,1)$.

Ex 3.

1. (a) $\phi(B)$ conserve la tendance affine si et seulement si

$$\begin{array}{lll} & \forall t \in \mathbb{Z} & \phi(B)(f_\tau)_t = f_\tau(t) \\ \iff & \forall t \in \mathbb{Z} & \alpha B(f_\tau)_t + \beta I(f_\tau)_t + \gamma B^{-1}(f_\tau)_t = f_\tau(t) \\ \iff & \forall t \in \mathbb{Z} & \alpha f_\tau(t-1) + \beta f_\tau(t) + \gamma f_\tau(t+1) = f_\tau(t) \\ \iff & \forall t \in \mathbb{Z} & (\alpha+\beta+\gamma)\, a + (\gamma-\alpha)\, b + (\alpha+\beta+\gamma)\, bt = a + bt \\ \iff & & \left\{ \begin{array}{rcl} (\alpha+\beta+\gamma)\, a + (\gamma-\alpha)\, b & = & a \\ (\alpha+\beta+\gamma)\, b & = & b \end{array} \right. \\ \iff & & \left\{ \begin{array}{rcl} \alpha+\beta+\gamma & = & 1 \\ \gamma-\alpha & = & 0 \end{array} \right. \end{array}$$

(en supposant $b \neq 0$). On en déduit $\gamma = \alpha$ et $\beta = 1 - 2\alpha$.

(b) D'après la question précédente, on choisit l'estimateur de $f_\tau(t)$ de la forme

$$\begin{aligned} \hat{X}_t &= \left(\alpha B + (1-2\alpha) I + \alpha B^{-1}\right)(X)_t \\ &= \left(\alpha B + (1-2\alpha) I + \alpha B^{-1}\right)(f_\tau)_t \\ &\quad + \left(\alpha B + (1-2\alpha) I + \alpha B^{-1}\right)(\varepsilon)_t + \theta\left(\alpha B + (1-2\alpha) I + \alpha B^{-1}\right)(\varepsilon)_{t-1} \\ &= f_\tau(t) + \alpha\varepsilon_{t+1} + (1 - 2\alpha + \theta\alpha)\,\varepsilon_t + (\alpha + \theta(1-2\alpha))\,\varepsilon_{t-1} + \theta\alpha\varepsilon_{t-2} \end{aligned}$$

Le bruit blanc ε étant centré, $\hat{X}_t$ est donc un estimateur sans biais. Calculons son risque :

$$\begin{aligned} \mathbb{E}\left[\left(\hat{X}_t - f_\tau(t)\right)^2\right] &= \mathbb{E}\left[\left(\alpha\varepsilon_{t+1} + (1-2\alpha+\theta\alpha)\,\varepsilon_t + (\alpha+\theta(1-2\alpha))\,\varepsilon_{t-1} + \theta\alpha\varepsilon_{t-2}\right)^2\right] \\ &= \alpha^2 \operatorname{var}(\varepsilon_{t+1}) + (1-2\alpha+\theta\alpha)^2 \operatorname{var}(\varepsilon_t) \\ &\quad + (\alpha+\theta(1-2\alpha))^2 \operatorname{var}(\varepsilon_{t-1}) + \theta^2\alpha^2 \operatorname{var}(\varepsilon_{t-2}) \\ &= \left(1 + \theta^2 - 4\alpha\left(1-\theta+\theta^2\right) + \alpha^2\left(6 - 8\theta + 6\theta^2\right)\right)\sigma^2 \end{aligned}$$

Le minimum de $1+\theta^2+2\alpha\,(1+\theta)+\alpha^2\left(1+(\theta-2)^2+(1-2\theta)^2+\theta^2\right)$ est atteint en

$$\alpha_m = \frac{1-\theta+\theta^2}{3-4\theta+3\theta^2}$$

Le filtre qui conserve la tendance affine et qui permet d'obtenir l'estimateur de risque minimal est donc $\alpha_m B + (1-2\alpha_m) I + \alpha_m B^{-1}$. *Si $\theta = 0$, on retrouve le filtre ordinaire $M_3 = \frac{1}{3}(B + I + B^{-1})$.*

2. (a) Si $\theta = 0$, le modèle est un modèle linéaire ordinaire, de la forme $X = A\beta + \varepsilon$, avec

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \\ \vdots & \vdots \\ 1 & n \end{pmatrix}, \quad \beta = \begin{pmatrix} a \\ b \end{pmatrix} \text{ et } \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Dans ce cas, les estimateurs des moindres carrés de a et b sont

$$\hat{\beta} = \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \left({}^tAA\right)^{-1} {}^tAX.$$

(b) Si $\theta \neq 0$, il s'agit d'un modèle linéaire avec erreurs corrélées. Comme l'on sait, ce modèle est équivalent, à une transformation près, à un modèle linéaire ordinaire, dont les résultats vont donc pouvoir s'appliquer ici.
La fonction d'autocovariance du processus MA(1) $(I + \theta B)(\varepsilon)$ est

$$\begin{aligned}
\gamma_{(I+\theta B)(\varepsilon)}(0) &= \operatorname{var}(\varepsilon_t + \theta\varepsilon_{t-1}) = (1+\theta^2)\sigma^2 \\
\gamma_{(I+\theta B)(\varepsilon)}(1) &= \operatorname{cov}(\varepsilon_t + \theta\varepsilon_{t-1}, \varepsilon_{t+1} + \theta\varepsilon_t) = \theta \operatorname{var}(\varepsilon_t) = \theta\sigma^2 \\
\gamma_{(I+\theta B)(\varepsilon)}(h) &= 0 \text{ si } h > 1
\end{aligned}$$

Les vecteurs aléatoires $(X_1, \ldots, X_n)$ et $(\varepsilon_1 + \theta\varepsilon_0, \ldots, \varepsilon_n + \theta\varepsilon_{n-1})$ ayant même matrice d'autocovariance, on en déduit que $\Gamma_X(n)_{i,j} = \gamma_{(I+\theta B)(\varepsilon)}(|i-j|)$, autrement dit $\Gamma_X(n) = \sigma^2 S_X(n)$, avec

$$S_X(n) = \begin{pmatrix} 1+\theta^2 & \theta & 0 & \cdots & 0 \\ \theta & 1+\theta^2 & \theta & \ddots & \vdots \\ 0 & \theta & 1+\theta^2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \theta \\ 0 & \cdots & 0 & \theta & 1+\theta^2 \end{pmatrix}$$

X est un vecteur gaussien de moyenne $M = A\begin{pmatrix} a \\ b \end{pmatrix}$ et de matrice de covariance $\Gamma_X(n)$. La densité de sa loi par rapport à la mesure de Lebesgue est donc

$$f_{a,b}(x_1, \ldots, x_n) = \frac{1}{\sqrt{(2\pi)^n \det \Gamma_X(n)}} \exp\left(-\frac{{}^t(x-M)\Gamma_X(n)^{-1}(x-M)}{2}\right)$$

avec $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. On recherche la valeur de β pour laquelle cette densité atteint son maximum. Comme elle ne s'annule pas, on peut considérer son logarithme, ce qui revient à rechercher le lieu du minimum de

$$\begin{aligned}
g(\beta) &= {}^t(x - A\beta)\Gamma_X(n)^{-1}(x - A\beta) \\
&= {}^tx\Gamma_X(n)^{-1}x - 2{}^tx\Gamma_X(n)^{-1}A\beta + {}^t\beta{}^tA\Gamma_X(n)^{-1}A\beta
\end{aligned}$$

g est la somme d'une constante, d'une forme linéaire et d'une forme bilinéaire symétrique. Sa différentielle est donc

$$dg(\beta)[H] = -2{}^tx\Gamma_X(n)^{-1}AH + 2{}^t\beta{}^tA\Gamma_X(n)^{-1}AH$$

avec $H \in \mathbb{R}^2$. Cette différentielle est nulle si et seulement si

$$-{}^tx\Gamma_X(n)^{-1}A + {}^t\beta{}^tA\Gamma_X(n)^{-1}A = 0$$

autrement dit si $\beta = \hat{\beta} = \left({}^t A\Gamma_X(n)^{-1}A\right)^{-1} {}^t A\Gamma_X(n)^{-1}x$. g possède donc un seul point critique, qui correspond nécessairement à son minimum car celui-ci est atteint. Les estimateurs du maximum de vraisemblance de a et b sont donc

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \left({}^t A\Gamma_X(n)^{-1}A\right)^{-1} {}^t A\Gamma_X(n)^{-1}X = \left({}^t AS_X(n)^{-1}A\right)^{-1} {}^t AS_X(n)^{-1}X.$$

On peut aussi déterminer l'estimateur $\hat{\sigma}^2_{MV}$ du maximum de vraisemblance de σ^2 : rappelons que la densité vaut

$$f_{a,b}(x_1,\ldots,x_n) = \frac{1}{\sqrt{(2\pi\sigma^2)^n \det S_X(n)}} \exp\left(-\frac{{}^t(x-A\beta)S_X(n)^{-1}(x-A\beta)}{2\sigma^2}\right)$$

En dérivant cette densité par rapport à σ^2, on détermine au point d'annulation une équation qui caractérise $\hat{\sigma}^2_{MV}$. D'où

$$\hat{\sigma}^2_{MV} = \frac{1}{n}{}^t(x-A\hat{\beta})S_X(n)^{-1}(x-A\hat{\beta}),$$

dont on déduit l'estimateur sans biais

$$\hat{\sigma}^2 = \frac{1}{n-2}{}^t(x-A\hat{\beta})S_X(n)^{-1}(x-A\hat{\beta})$$

3. Soit $T = \alpha_1 X_1 + \cdots + \alpha_n X_n$ une combinaison linéaire de $(X_1,\ldots,X_n)$. C'est le meilleur prédicteur linéaire de X_{n+2} si et seulement si parmi toutes les variables aléatoires de sa forme il minimise la distance hilbertienne à X_{n+2}. Or

$$\begin{aligned}
\|X_{n+2}-T\|_2^2 &= \mathbb{E}\left[(X_{n+2}-T)^2\right] \\
&= \mathrm{var}\,(X_{n+2}-T) + (\mathbb{E}[X_{n+2}-T])^2 \\
&= \mathrm{var}(X_{n+2}) + \mathrm{var}(T) + (\mathbb{E}[a+b(n+2)-T])^2 \quad \text{car } T \text{ et } X_{n+2} \text{ sont indépendants} \\
&= \mathrm{var}(X_{n+2}) + \mathbb{E}\left[(a+b(n+2)-T)^2\right]
\end{aligned}$$

Minimiser $\|X_{n+2}-T\|_2$ revient donc à minimiser $\mathbb{E}\left[(a+b(n+2)-T)^2\right]$, autrement dit le risque quadratique de T considéré comme estimateur de $a+b(n+2)$. D'après le théorème de Gauss-Markov, le meilleur estimateur est alors $\hat{a}+\hat{b}(n+2)$, qui est un estimateur sans biais. C'est donc aussi le meilleur prédicteur linéaire de X_{n+2}. Calculons sa variance. La matrice de covariance de $\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \left({}^t AS_X(n)^{-1}A\right)^{-1} {}^t AS_X(n)^{-1}X$ est

$$\begin{aligned}
&\left(\left({}^t A\Gamma_X(n)^{-1}A\right)^{-1} {}^t A\Gamma_X(n)^{-1}\right)\Gamma_X(n)\left(\Gamma_X(n)^{-1}A\left({}^t A\Gamma_X(n)^{-1}A\right)^{-1}\right) \\
&= \left({}^t A\Gamma_X(n)^{-1}A\right)^{-1} = \left({}^t AS_X(n)^{-1}A\right)^{-1}\sigma^2.
\end{aligned}$$

La variance de $\hat{a}+\hat{b}(n+2)$ est alors $\mathrm{var}(\hat{a}+\hat{b}(n+2)) = s^2\sigma^2$, avec

$$s^2 = (1, n+2)\left({}^t AS_X(n)^{-1}A\right)^{-1}\begin{pmatrix} 1 \\ n+2 \end{pmatrix}$$

$X_{n+2}-\hat{a}-\hat{b}(n+2)=a+b(n+2)-\hat{a}-\hat{b}(n+2)+\varepsilon_{n+2}+\theta\varepsilon_{n+1}$ suit une loi gaussienne centrée, de variance $s^2\sigma^2+\mathrm{var}(\varepsilon_{n+2}+\theta\varepsilon_{n+1})=(s^2+1+\theta^2)\sigma^2$, car $(\varepsilon_{n+2},\varepsilon_{n+1})$ est indépendant de $(\hat{a},\hat{b})$.
D'autre part, $\hat{\sigma}^2(n-2)/\sigma^2$ suit une loi du χ_2 à $n-2$ degrés de liberté, et est indépendant de $\hat{\beta}$ et de $(\varepsilon_{n+1},\varepsilon_{n+2})$. On en déduit que

$$\frac{X_{n+2}-\hat{a}-\hat{b}(n+2)}{\sqrt{(s^2+1+\theta^2)\sigma^2}}\frac{1}{\sqrt{\frac{\hat{\sigma}^2}{\sigma^2}}}=\frac{X_{n+2}-\hat{a}-\hat{b}(n+2)}{\sqrt{(s^2+1+\theta^2)\hat{\sigma}^2}}$$

suit une loi de Student à $n-2$ degré de liberté. Soit $t_{0,975}$ le fractile d'ordre $97,5\%$ de cette loi de Student. Alors

$$\begin{aligned}
0,95 &= \mathbb{P}\left(\frac{X_{n+2}-\hat{a}-\hat{b}(n+2)}{\sqrt{(s^2+1+\theta^2)\hat{\sigma}^2}}\in[-t_{0,975},t_{0,975}]\right)\\
&= \mathbb{P}\Big(X_{n+2}\in\Big[\hat{a}+\hat{b}(n+2)-t_{0,975}\sqrt{(s^2+1+\theta^2)\hat{\sigma}^2},\\
&\qquad\qquad \hat{a}+\hat{b}(n+2)+t_{0,975}\sqrt{(s^2+1+\theta^2)\hat{\sigma}^2}\Big]\Big)
\end{aligned}$$

D'où l'intervalle de prévision demandé.

Ex 4. Soit $X=(X_t,t\in\mathbb{Z})$ le processus défini par

$$\forall t\in\mathbb{Z},\quad X_t=\cos\left(\frac{2\pi t}{3}+U\right)$$

où U est une variable aléatoire de loi uniforme sur $[-\pi,\pi]$.

1. Le processus X est du second ordre car les variables X_t sont bornées, donc de carré intégrable. Calculons sa fonction moyenne et sa fonction d'autocovariance :

$$\begin{aligned}
\mu_X(t) &= \mathbb{E}[X_t]=\mathbb{E}\left[\cos\left(\frac{2\pi t}{3}+U\right)\right]\\
&= \int_{-\pi}^{\pi}\cos\left(\frac{2\pi t}{3}+u\right)\frac{du}{2\pi}=\left[\sin\left(\frac{2\pi t}{3}+u\right)\right]_{-\pi}^{\pi}=0
\end{aligned}$$

$$\begin{aligned}
\gamma_X(s,t) &= \mathrm{cov}(X_s,X_t)=\mathbb{E}[X_sX_t]\ \text{ car les variables } X_s,X_t \text{ sont centrées}\\
&= \mathbb{E}\left[\cos\left(\frac{2\pi s}{3}+U\right)\cos\left(\frac{2\pi t}{3}+U\right)\right]=\frac{1}{2}\mathbb{E}\left[\cos\left(\frac{2\pi(s+t)}{3}+2U\right)+\cos\left(\frac{2\pi(s-t)}{3}\right)\right]\\
&= \frac{1}{2}\cos\left(\frac{2\pi(s-t)}{3}\right)+\frac{1}{2}\int_{-\pi}^{\pi}\cos\left(\frac{2\pi(s+t)}{3}+2u\right)\frac{du}{2\pi}=\frac{1}{2}\cos\left(\frac{2\pi(s-t)}{3}\right)+0
\end{aligned}$$

La fonction moyenne est constante, et la fonction d'autocorrélation invariante par translation temporelle : le processus X est donc faiblement stationnaire.

2. Notons $\varepsilon=(\varepsilon_t,t\in\mathbb{Z})$ le bruit blanc d'innovation associé à X. Alors $\varepsilon_t=X_t-P^{\perp}_{V^2(1,X_s,s<t)}(X_t)$. Or on vérifie instantanément que X est un processus périodique de période 3. En particulier, $X_t=X_{t-3}\in V^2(1,X_s,s<t)$. D'où $P^{\perp}_{V^2(1,X_s,s<t)}(X_t)=X_t$ et $\varepsilon_t=0$.

3. Les deux variables aléatoires $\cos\left(\frac{2\pi t}{3}+U\right)$ et $\sin\left(\frac{2\pi t}{3}+U\right)$ sont centrées. En conséquence,

$$\begin{aligned}
\operatorname{cov}\left(\cos\left(\frac{2\pi t}{3}+U\right),\sin\left(\frac{2\pi t}{3}+U\right)\right) &= \mathbb{E}\left[\cos\left(\frac{2\pi t}{3}+U\right)\sin\left(\frac{2\pi t}{3}+U\right)\right]\\
&= \mathbb{E}\left[\frac{1}{2}\sin\left(\frac{4\pi t}{3}+2U\right)\right]\\
&= \int_{-\pi}^{\pi}\sin\left(\frac{4\pi t}{3}+2u\right)\frac{du}{4\pi}=\left[\cos\left(\frac{4\pi t}{3}+2u\right)\right]_{-\pi}^{\pi}\\
&= 0
\end{aligned}$$

Les deux variables sont donc bien décorrélées, et orthogonales. Or

$$\begin{aligned}
X_{t+1} &= \cos\left(\frac{2\pi(t+1)}{3}+U\right)=\cos\left(\frac{2\pi t}{3}+U\right)\cos\left(\frac{2\pi}{3}\right)-\sin\left(\frac{2\pi t}{3}+U\right)\sin\left(\frac{2\pi}{3}\right)\\
&= -\frac{1}{2}\cos\left(\frac{2\pi t}{3}+U\right)-\frac{\sqrt{3}}{2}\sin\left(\frac{2\pi t}{3}+U\right)
\end{aligned}$$

D'où

$$\begin{aligned}
P^{\perp}_{V^2(1,X_t)}(X_{t+1}) &= -\frac{1}{2}P^{\perp}_{V^2(1,X_t)}\left(\cos\left(\frac{2\pi t}{3}+U\right)\right)-\frac{\sqrt{3}}{2}P^{\perp}_{V^2(1,X_t)}\left(\sin\left(\frac{2\pi t}{3}+U\right)\right)\\
&= -\frac{1}{2}\cos\left(\frac{2\pi t}{3}+U\right)
\end{aligned}$$

On en déduit aussi que

$$\begin{aligned}
X_{t+2} &= -\frac{1}{2}\cos\left(\frac{2\pi t}{3}+U\right)+\frac{\sqrt{3}}{2}\sin\left(\frac{2\pi t}{3}+U\right)\\
&= -\cos\left(\frac{2\pi t}{3}+U\right)-\left(-\frac{1}{2}\cos\left(\frac{2\pi t}{3}+U\right)-\frac{\sqrt{3}}{2}\sin\left(\frac{2\pi t}{3}+U\right)\right)\\
&= -X_t-X_{t+1}
\end{aligned}$$

Il en résulte $P^{\perp}_{V^2(1,X_t,X_{t+1})}(X_{t+2})=X_{t+2}$.

4. Notons ρ_X la fonction d'autocorrélation de X.
   - $r_X(1)=\rho_X(1)=\gamma_X(1)/\gamma_X(0)=\cos 1$.
   - Pour calculer $r_X(2)$, notons que $R_X(2)=(\rho_X(|i-j|)^2_{i,j=1}$ est de déterminant égal à $1-\cos^2(1)\neq 0$, et donc est inversible. Dans ce cas, on sait que $r_X(2)=\alpha_2^{(2)}$, avec $P^{\perp}_{V^2(1,X_t,X_{t+1})}(X_{t+2})=\alpha_0^{(2)}+\alpha_1^{(2)}X_{t+1}+\alpha_2^{(2)}X_t$. Or $X_{t+2}=-X_t-X_{t+1}$. D'où $r_X(2)=-1$.
   - Enfin, si $h\geq 3$, alors $P^{\perp}_{V^2(1,X_{t+1},\ldots,X_{t+h-1})}(X_{t+h})=X_{t+h}$, pour les mêmes raisons que vues à la question précédente, et donc, par définition, $r_X(h)=0$.

   Notons enfin qu'il ne s'agit pas du tout d'un processus AR(2), ne serait-ce que parce que l'innovation d'un tel processus n'est pas nulle (sauf si le processus lui-même est nul). Rappelons que dans les hypothèses du théorème caractérisant un processus AR à partir de sa fonction d'autocorrélation partielle, on suppose que sa fonction d'autocorrélation tend vers 0 en l'infini ; ce qui n'est pas le cas ici.