Séries Chronologiques

Université d'Angers – M2 Data Science

Fiche de TD

Exercice 1. Soient (ε_t) un bruit blanc gaussien réduit et (ξ_t) la suite définie par

$$\forall\, t \in \mathbb{Z}, \qquad \xi_t = \frac{\varepsilon_t^2 - 1}{\sqrt{2}}.$$

On définit la série chronologique (X_t) par

$$\forall\, t \in \mathbb{Z}, \qquad X_t = \begin{cases} \varepsilon_t & \text{si } t \text{ est pair} \\ \xi_{t-1} & \text{si } t \text{ est impair.} \end{cases}$$

1) Vérifier que (X_t) n'est pas un bruit blanc fort.
2) Montrer que (X_t) forme malgré tout un bruit blanc et identifier sa variance.

Exercice 2. Montrer que le filtrage par moyennes mobiles arithmétiques (standard et modifiée) laisse invariantes les tendances linéaires. Justifier l'importance de cette observation dans la stratégie d'estimation vue en cours pour le modèle additif.

Exercice 3. Soit le processus

$$\forall\, t \in \mathbb{Z}, \qquad X_t = a_0 + a_1\, t + \sin\left(\frac{\pi t}{2}\right) + \varepsilon_t$$

où $a_0, a_1 \in \mathbb{R}$ et (ε_t) est un bruit blanc de variance $\sigma^2 > 0$.

1) Pour $\tau \in \mathbb{N}^*$, on considère l'opérateur

$$M_\tau(B) = \frac{1}{\tau} \sum_{k=0}^{\tau-1} B^k \qquad \text{et l'on pose} \qquad Y_t = M_\tau(B)\, X_t.$$

   Identifier le plus petit τ tel que la série filtrée ne comporte plus de composante périodique, puis montrer que le processus (Y_t) ainsi construit n'est toujours pas stationnaire (sauf si $a_1 = 0$).
2) Montrer que le processus (Z_t) défini par $Z_t = (I - B)\, Y_t$ est stationnaire. Donner sa fonction d'autocovariance γ et d'autocorrélation ρ.

Exercice 4. Soit (ε_t) un bruit blanc fort de variance $\sigma^2 > 0$. On considère les processus définis par

$$\forall\, t \geqslant 1, \qquad S_t = \sum_{k=1}^{t} \varepsilon_k \qquad \text{et} \qquad P_t = \prod_{k=1}^{t} \varepsilon_k.$$

Calculer l'espérance de ces processus et, pour $h \geqslant 0$, calculer $\mathbb{C}\mathrm{ov}(S_t, S_{t+h})$ et $\mathbb{C}\mathrm{ov}(P_t, P_{t+h})$. Que peut-on dire du processus (P_t) lorsque $\sigma^2 = 1$?

Exercice 5. Soit (ε_t) un bruit blanc gaussien de variance $\sigma^2 > 0$. Étudier la stationnarité et, le cas échéant, calculer l'autocovariance γ et l'autocorrélation ρ des processus définis par :

1) $\forall\, t \in \mathbb{Z},\ X_t = \varepsilon_t\, \mathbb{1}_{\{t \text{ est pair}\}} + (1-\varepsilon_t)\, \mathbb{1}_{\{t \text{ est impair}\}}$.
2) $\forall\, t \in \mathbb{Z}, \forall\, \theta \in \mathbb{R},\ Y_t = \cos(\theta\, \varepsilon_t) + \sin(\theta\, \varepsilon_{t-1})$.
3) $\forall\, t \in \mathbb{Z},\ Z_t = \varepsilon_t^2 - 2\, \varepsilon_{t-1}$.

Exercice 6. Soient (X_t) et (Y_t) deux processus stationnaires et mutuellement décorrélés, d'autocovariances respectives γ_X et γ_Y.

1) Montrer que le processus $(X_t + Y_t)$ est stationnaire et exprimer son autocovariance γ_{X+Y} en fonction de γ_X et de γ_Y.
2) Proposer un contre-exemple montrant que, si l'on enlève l'hypothèse d'absence de corrélation entre les processus, alors $(X_t + Y_t)$ n'est pas nécessairement stationnaire.

Exercice 7. On se place dans l'espace L^2 des variables aléatoires réelles de carré intégrable, muni de son produit scalaire usuel. Soient $X, Y \in L^2$ avec $\mathbb{V}(X) > 0$ et $\mathcal{M} = \text{Vect}\{1, X\} = \{a + b\, X \text{ pour } a, b \in \mathbb{R}\}$ l'espace vectoriel formé par les transformations affines de X. On note $P_{\mathcal{M}}$ la projection orthogonale dans L^2 sur $\mathcal{M}$.

1) On pose $P_{\mathcal{M}} Y = \alpha + \beta\, X$. Montrer que

$$\alpha = \mathbb{E}[Y] - \frac{\mathbb{C}\text{ov}(X,Y)}{\mathbb{V}(X)}\, \mathbb{E}[X] \quad \text{et} \quad \beta = \frac{\mathbb{C}\text{ov}(X,Y)}{\mathbb{V}(X)}.$$

2) À quelle méthode statistique classique correspond cette procédure lorsque l'on travaille non plus dans L^2 mais dans $\mathbb{R}^n$ muni de son produit scalaire usuel ?

Exercice 8. Soit l'opérateur $\Phi(B) = I - \frac{3\,B}{4} + \frac{B^2}{8}$.

1) Justifier que $\Phi(B)$ est inversible. On considère son inverse sous la forme

$$\Phi^{-1}(B) = \sum_{k\, \in\, \mathbb{Z}} \psi_k\, B^k.$$

Montrer que la suite de coefficients (ψ_k) satisfait $\psi_k = 0$ pour $k < 0$, $\psi_0 = 1$, $\psi_1 = \frac{3}{4}$ et, pour $k \geqslant 2$, la relation de récurrence

$$\psi_k = \frac{3}{4}\, \psi_{k-1} - \frac{1}{8}\, \psi_{k-2}.$$

2) Montrer que la solution explicite est donnée par

$$\forall\, k \geqslant 0, \qquad \psi_k = \frac{1}{2^{k-1}} - \frac{1}{4^k}.$$

3) On considère le processus AR(2) défini par

$$\forall\, t \in \mathbb{Z}, \qquad \Phi(B)\, X_t = \varepsilon_t$$

où (ε_t) est un bruit blanc de variance $\sigma^2 > 0$.

a) Justifier que le processus est stationnaire. Est-il causal ?
b) Calculer son espérance μ et sa variance $\gamma(0)$ en fonction de σ^2. Donner l'expression de son autocorrélation ρ (sans faire explicitement le calcul).

Exercice 9. On veut montrer dans cet exercice qu'un processus AR(1) de la forme

$$\forall\, t \in \mathbb{Z}, \qquad X_t = \phi\, X_{t-1} + \varepsilon_t$$

n'est pas stationnaire si $|\phi| = 1$. Le processus que l'on étudie est à valeurs réelles, ce qui signifie que $|\phi| = 1 \Leftrightarrow \phi = \pm 1$. On suppose dans la suite de l'exercice que $t \geqslant 1$. On remarque alors que

$$X_t - \phi^{\,t}\, X_0 = \sum_{k=0}^{t-1} \phi^{\,k}\, \varepsilon_{t-k}.$$

Raisonnons par l'absurde et supposons que $(X_t)_{t \,\in\, \mathbb{Z}}$ est un processus stationnaire.

1) Déduire des propriétés de la fonction γ que, dans ce cas, $4\,\gamma(0) \geqslant \sigma^2\, t$.

2) Conclure à la contradiction.

Exercice 10. Soit (X_t) le processus engendré par la relation

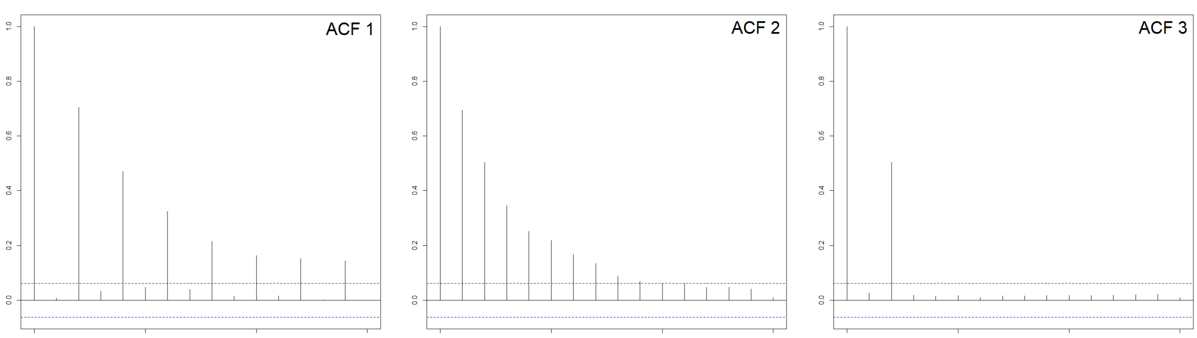
$$\forall\, t \in \mathbb{Z}, \qquad X_t = \phi\, X_{t-2} + \varepsilon_t$$

où $\phi \in \mathbb{R}^*$. Il s'agit donc d'un AR(2) dont le premier paramètre est fixé à 0.

1) Déterminer les contraintes sur ϕ qui garantissent que (X_t) est causal et, dans ce cas, identifier les coefficients (ψ_k) de l'écriture

$$\forall\, t \in \mathbb{Z}, \qquad X_t = \sum_{k=0}^{+\infty} \psi_k\, \varepsilon_{t-k}.$$

2) Parmi les trois ACF empiriques données ci-dessous, indiquer la seule qui pourrait correspondre à un tel processus. Justifier par le calcul.



3) Sous les contraintes de causalité du processus, construire son autocorrélation partielle $\alpha(h)$ pour $h \in \mathbb{N}^*$.

Exercice 11. On considère le processus ARMA(2,1) engendré par la relation

$$\forall\, t \in \mathbb{Z}, \qquad X_t = \frac{1}{12}\, X_{t-1} + \frac{1}{12}\, X_{t-2} + \varepsilon_t + \frac{1}{4}\, \varepsilon_{t-1}.$$

1) Justifier que le processus est causal (et donc stationnaire). En déduire que $\mathbb{E}[X_{t-i}\, \varepsilon_t] = 0$ pour tout $i \geqslant 1$.

2) Montrer que $\mathbb{E}[X_t\, \varepsilon_t] = \sigma^2$ et que $\mathbb{E}[X_t\, \varepsilon_{t-1}] = \frac{\sigma^2}{3}$.

3) On s'intéresse à l'autocovariance du processus.

a) Montrer, par un raisonnement de type Yule-Walker et en utilisant les questions précédentes, que

$$\begin{pmatrix} 12 & -1 & -1 \\ -1 & 11 & 0 \\ -1 & -1 & 12 \end{pmatrix} \begin{pmatrix} \gamma(0) \\ \gamma(1) \\ \gamma(2) \end{pmatrix} = \begin{pmatrix} 13 \\ 3 \\ 0 \end{pmatrix} \sigma^2$$

et que

$$\forall\, h \geqslant 3, \qquad \gamma(h) = \frac{\gamma(h-1)}{12} + \frac{\gamma(h-2)}{12}.$$

b) La résolution du système correspondant aux valeurs initiales (que l'on ne demande pas d'effectuer) donne

$$\begin{pmatrix} \gamma(0) \\ \gamma(1) \\ \gamma(2) \end{pmatrix} = \begin{pmatrix} 12 & -1 & -1 \\ -1 & 11 & 0 \\ -1 & -1 & 12 \end{pmatrix}^{-1} \begin{pmatrix} 13\,\sigma^2 \\ 3\,\sigma^2 \\ 0 \end{pmatrix} = \begin{pmatrix} 9 \\ 3 \\ 1 \end{pmatrix} \frac{\sigma^2}{8}.$$

Vérifier alors que la solution à l'équation de récurrence est donnée par

$$\gamma(h) = \frac{9\,\sigma^2}{8} \left(\frac{1}{3}\right)^h.$$

4) On s'intéresse maintenant à son autocorrélation partielle.

a) Calculer $\alpha(1)$ puis montrer qu'il existe une constante k telle que

$$\begin{pmatrix} \gamma(0) \\ \gamma(1) \\ \vdots \\ \gamma(h-1) \end{pmatrix} = k \begin{pmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(h) \end{pmatrix}.$$

En déduire la valeur de $\alpha(h)$ pour tout $h \geqslant 2$.

b) Le résultat précédent ne contredit-il pas l'écriture ARMA(2,1) du processus ? Pouvez-vous résoudre ce paradoxe (qui n'en est pas un) ?

Exercice 12. Construire un processus stationnaire dont la fonction d'autocorrélation est donnée par

$$\forall\, h \in \mathbb{Z}, \qquad \rho(h) = \begin{cases} 1 & \text{si } h = 0 \\ \frac{2}{5} & \text{si } |h| = 1 \\ 0 & \text{si } |h| > 1 \end{cases}$$

puis déterminer $\alpha(1)$, $\alpha(2)$ et $\alpha(3)$, ses trois premières autocorrélations partielles.

Exercice 13. Soit (ε_t) le processus défini par

$$\forall\, t \in \mathbb{Z}, \qquad \varepsilon_t = \eta_t^2\, \eta_{t-1}$$

où

$$\eta_t \sim \begin{cases} \mathcal{R}(p) & \text{si t est pair} \\ \mathcal{N}(\mu, \sigma^2) & \text{si t est impair} \end{cases}$$

avec $0 \leqslant p \leqslant 1$, $\mu \in \mathbb{R}$ et $\sigma^2 > 0$. On suppose les variables (η_t) indépendantes entre elles. On rappelle que :

- $X \sim \mathcal{R}(p)$ signifie que $\mathbb{P}(X = 1) = p = 1 - \mathbb{P}(X = -1)$.
- Les moments d'ordre 3 et 4 d'une variable $Y \sim \mathcal{N}(\mu, \sigma^2)$ sont donnés par $\mathbb{E}[Y^3] = 3\mu\sigma^2 + \mu^3$ et par $\mathbb{E}[Y^4] = 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4$.

1) Montrer que le processus (ε_t) est centré si et seulement si $p = 1/2$ et $\mu = 0$.

2) On suppose pour la suite que le processus (ε_t) est centré (et donc que $p = 1/2$, $\mu = 0$). Montrer que $\mathbb{C}\mathrm{ov}(\varepsilon_t, \varepsilon_s) = 0$ dès que $t \neq s$.
3) Montrer qu'il existe une et une seule valeur de σ^2 permettant de définir le processus (ε_t) comme un bruit blanc faible (centré). On donnera cette valeur.
4) On admettra maintenant que le processus (ε_t) est un bruit blanc faible centré. A t fixé, montrer que ε_t et ε_{t-1} n'ont pas la même loi.
5) Le processus (ε_t) est-il un bruit blanc fort ? Justifier la réponse.

Exercice 14. Soit le polynôme de degré n défini, pour tout $z \in \mathbb{C}$, par

$$P(z) = 1 - p_1\, z - p_2\, z^2 - \ldots - p_n\, z^n.$$

1) Montrer par l'absurde que la condition $\|p\|_1 = \sum_i |p_i| < 1$ est suffisante pour que P soit causal (c'est-à-dire que $P(z) \neq 0$ pour tout $|z| \leqslant 1$).
2) En déduire une condition suffisante pour vérifier rapidement la causalité d'un processus ARMA(p, q). Montrer qu'elle n'est pas nécessaire.

Exercice 15. On considère dans ce problème le processus MA(1) donné par

$$\forall\, t \in \mathbb{Z}, \qquad X_t = \varepsilon_t + \theta\, \varepsilon_{t-1}$$

avec comme paramètre $|\theta| < 1$. Pour tout $n \geqslant 1$, on note $\Gamma_n = (\gamma(i-j))_{0 \,\leqslant\, i,j \,\leqslant\, n-1}$. La suite $(\Gamma_n)_{n \,\geqslant\, 1}$ est donc formée de matrices symétriques d'ordre $n \times n$.

1) Donner Γ_1 et montrer que Γ_n est tridiagonale pour $n \geqslant 2$.
2) Pour tout $n \geqslant 1$, on pose $d_n = \frac{\det(\Gamma_n)}{\sigma^{2n}}$. Montrer que

$$d_n = \left\{ \begin{array}{ll} 1 + \theta^2 & \text{si } n = 1 \\ 1 + \theta^2 + \theta^4 & \text{si } n = 2 \\ (1 + \theta^2)\, d_{n-1} - \theta^2\, d_{n-2} & \text{si } n > 2. \end{array} \right.$$

3) Vérifier que la solution est donnée par $\forall\, n \geqslant 1,\ d_n = 1 + \theta^2 + \ldots + \theta^{2n}$.
4) En déduire la valeur de

$$\lim_{n \,\to\, +\infty} \frac{\ln(\det(\Gamma_n))}{n}.$$

Exercice 16. Pour un processus observable sur un intervalle des temps $t \in \{1, \ldots, n\}$, considéré comme stationnaire, on va considérer l'estimateur de l'autocovariance

$$\forall\, |h| \leqslant n-1, \qquad \widehat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} \big(X_t - \bar{X}_n\big)\, \big(X_{t+|h|} - \bar{X}_n\big).$$

On appelle $\widehat{\rho}(h)$ l'estimateur de l'autocorrélation $\rho(h)$ du processus construit à partir de $\widehat{\gamma}(h)$, comme dans la Sec. 2.5 du poly. Sous certaines hypothèses que l'on suppose ici vérifiées, on peut montrer que l'estimateur $\widehat{\gamma}(h)$ est consistant et que, lorsque $n \to \infty$,

$$\forall\, h \in \mathbb{Z}, \qquad \sqrt{n}\, \big(\widehat{\rho}(h) - \rho(h)\big) \xrightarrow{\mathcal{L}} \mathcal{N}(0, w_h)$$

où la variance asymptotique est donnée par la *formule de Bartlett* et vaut

$$w_h = \sum_{k=1}^{\infty} \big(\rho(k+h) + \rho(k-h) - 2\, \rho(k)\, \rho(h)\big)^2.$$

a) Soit $(X_t) \sim \mathrm{MA}(q)$. Montrer que $\rho(h) = 0$ pour tout $|h| \geqslant q+1$ et donc que $w_{q+1} = 1+2\,\rho^2(1)+\ldots+2\,\rho^2(q)$.

b) Soit maintenant $(X_t) \sim \mathrm{MA}(q+1)$. Montrer que $\rho(q+1) \neq 0$.

c) Sur la base des observations $(X_1, \ldots, X_n)$, on souhaite mettre en concurrence les modélisations $\mathrm{MA}(q)$ et $\mathrm{MA}(q+1)$ par l'intermédiaire du test de $\mathcal{H}_0$: "$\rho(q+1) = 0$" contre l'alternative $\mathcal{H}_1$: "$\rho(q+1) \neq 0$". Proposer un protocole statistique (de niveau 5%) pour effectuer ce test à l'aide du résultat rappelé ci-dessus. On pourra par exemple étudier le comportement de

$$S_n = \frac{\sqrt{n}\,\widehat{\rho}(q+1)}{\sqrt{\widehat{w}_{q+1}}}$$

sous $\mathcal{H}_0$ et sous $\mathcal{H}_1$ pour en déduire une zone de rejet de $\mathcal{H}_0$ au profit de $\mathcal{H}_1$.

Exercice 17. Donner l'écriture explicite de :

1) $(X_t) \sim \mathrm{ARIMA}(1,1,1)$.
2) $(X_t) \sim \mathrm{ARI}(2,1)$.
3) $(X_t) \sim \mathrm{SARIMA}(2,0,1) \times (1,1,0)_7$.
4) $(X_t) \sim \mathrm{SARMA}(2,0) \times (1,2)_{12}$.
5) $(X_t) \sim \mathrm{I}(2)$.

Exercice 18. Soit le processus (X_t) engendré par la relation

$$\forall\, t \in \mathbb{Z}, \qquad X_t = 2\,r\cos\theta\,X_{t-1} - r^2\,X_{t-2} + \varepsilon_t$$

où (ε_t) est un bruit blanc de variance $\sigma^2 > 0$, $0 < r < 1$ et $0 < \theta < \pi$.

1) Montrer que (ε_t) est l'innovation du processus.
2) Par un raisonnement de type Yule-Walker, montrer que

$$\begin{pmatrix} 1 & -2\,r\cos\theta & r^2 & 0 \\ -2\,r\cos\theta & 1+r^2 & 0 & 0 \\ r^2 & -2\,r\cos\theta & 1 & 0 \\ 0 & r^2 & -2\,r\cos\theta & 1 \end{pmatrix} \begin{pmatrix} \gamma(0) \\ \gamma(1) \\ \gamma(2) \\ \gamma(3) \end{pmatrix} = \begin{pmatrix} \sigma^2 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

3) On admettra (sans chercher à refaire les calculs) que le système de la question 2 conduit à la solution

$$\begin{pmatrix} \gamma(0) \\ \gamma(1) \\ \gamma(2) \\ \gamma(3) \end{pmatrix} = \sigma^2\,k_{r,\,\theta} \begin{pmatrix} 1+r^2 \\ 2\,r\cos\theta \\ r^2\,(4\cos^2\theta - (1+r^2)) \\ 2\,r^3\,\cos\theta\,(4\cos^2\theta - (2+r^2)) \end{pmatrix}$$

où $k_{r,\,\theta} \neq 0$ est une constante commune à chaque composante, qui dépend de r et de θ. Déduire de ces simplifications l'expression de l'autocorrélation partielle α du processus.

4) Justifier sans faire aucun calcul que

$$\frac{\begin{vmatrix} 1+r^2 & 2\,r\cos\theta \\ 2\,r\cos\theta & r^2\,(4\cos^2\theta - (1+r^2)) \end{vmatrix}}{\begin{vmatrix} 1+r^2 & 2\,r\cos\theta \\ 2\,r\cos\theta & 1+r^2 \end{vmatrix}} = -r^2$$

et que

$$\begin{vmatrix} 1+r^2 & 2\,r\cos\theta & 2\,r\cos\theta \\ 2\,r\cos\theta & 1+r^2 & r^2\,(4\cos^2\theta - (1+r^2)) \\ r^2\,(4\cos^2\theta - (1+r^2)) & 2\,r\cos\theta & 2\,r^3\,\cos\theta\,(4\cos^2\theta - (2+r^2)) \end{vmatrix} = 0.$$

5) Donner le meilleur prédicteur linéaire de X_{t+1} et l'erreur quadratique moyenne de cette prédiction.

Exercice 19. Soit (X_t) le processus suivant la représentation

$$\forall\, t \in \mathbb{Z}, \qquad (I - B)\, X_t = (I + \theta\, B)\, \varepsilon_t$$

où (ε_t) un bruit blanc de variance $\sigma^2 > 0$ et $|\theta| < 1$.

1) Quelle est la nature des processus (X_t) et (ΔX_t) ? Sont-ils stationnaires ?

2) Donner le meilleur prédicteur linéaire $\widehat{Y}_{t+h}$ de $Y_{t+h} = \Delta X_{t+h}$ pour tout horizon $h \geqslant 1$.

3) Montrer que le meilleur prédicteur linéaire de $\widehat{X}_{t+h}$ de X_{t+h} est tel que

$$\widehat{X}_{t+h} = \widehat{X}_{t+1}.$$

4) Exprimer l'erreur de prévision à horizon h en fonction de (ε_t). En déduire sa variance.

Exercice 20. Montrer que le krigeage simple reconstruit à l'identique les points de mesures sur lesquels s'appuie l'estimation.

Pour aller plus loin... On se place dans le cadre du processus MA(1) centré, pour lequel on va déterminer la fonction d'autocorrélation partielle α. On considère

$$\forall\, t \in \mathbb{Z}, \qquad X_t = (I + \theta\, B)\, \varepsilon_t$$

avec $0 < |\theta| < 1$, où (ε_t) est un bruit blanc de variance $\sigma^2 > 0$.

1) Donner $\alpha(1)$.

2) Pour $h \geqslant 2$, soit la projection orthogonale de X_{h+1} sur $\mathrm{Vect}\{X_h, \ldots, X_1\}$ donnée par

$$X_{h+1} = \phi_1\, X_h + \ldots + \phi_h\, X_1 + R$$

où $\langle R, X_i \rangle = 0$ pour tout $i = 1, \ldots, h$. Montrer que les coefficients (ϕ_j) satisfont la relation de récurrence

$$\left\{ \begin{array}{l} (1+\theta^2)\, \phi_1 + \theta\, \phi_2 = \theta \\ \theta\, \phi_{j-1} + (1+\theta^2)\, \phi_j + \theta\, \phi_{j+1} = 0 \qquad \text{pour } j \in \{2, \ldots, h-1\} \\ \theta\, \phi_{h-1} + (1+\theta^2)\, \phi_h = 0. \end{array} \right.$$

3) En déduire que l'autocorrélation partielle α du processus MA(1) mis sous forme canonique est

$$\alpha(h) = -\frac{(-\theta)^h (1 - \theta^2)}{1 - \theta^{2(h+1)}}.$$

4) Vérifier que α décroît vers 0 à vitesse géométrique lorsque $h \to +\infty$, comme le prévoit la théorie.

Pour aller encore plus loin... Reprenons la Sec. 3.3 du cours pour tenter de la généraliser *a minima* aux AR(p) pour p quelconque. L'idée est la suivante : trouver les conditions que doivent satisfaire les paramètres $\phi_1, \ldots, \phi_p$ pour que le processus défini par

$$\forall\, n \geqslant p, \qquad X_n = \phi_1\, X_{n-1} + \ldots + \phi_p\, X_{n-p} + \varepsilon_n$$

avec des valeurs initiales $X_{p-1}, \ldots, X_0$ décorrélées du bruit blanc (ε_n) de variance $\sigma^2 > 0$, possède certaines propriétés de stationnarité (on s'intéressera seulement à l'espérance et à la variance asymptotiques).

1) On a vu qu'il était très différent de travailler sur $\mathbb{Z}$ ou sur $\mathbb{N} \cup \mathbb{I}$. Rappeler la réponse la question lorsque l'on définit le processus sur $\mathbb{Z}$.

2) On veut tout d'abord montrer que l'AR(p) est en fait un AR(1) vectoriel. Pour cela, montrer qu'il existe une matrice A et un bruit vectoriel (E_n) tels que

$$\forall\, n \geqslant p, \qquad Y_n = A\, Y_{n-1} + E_n$$

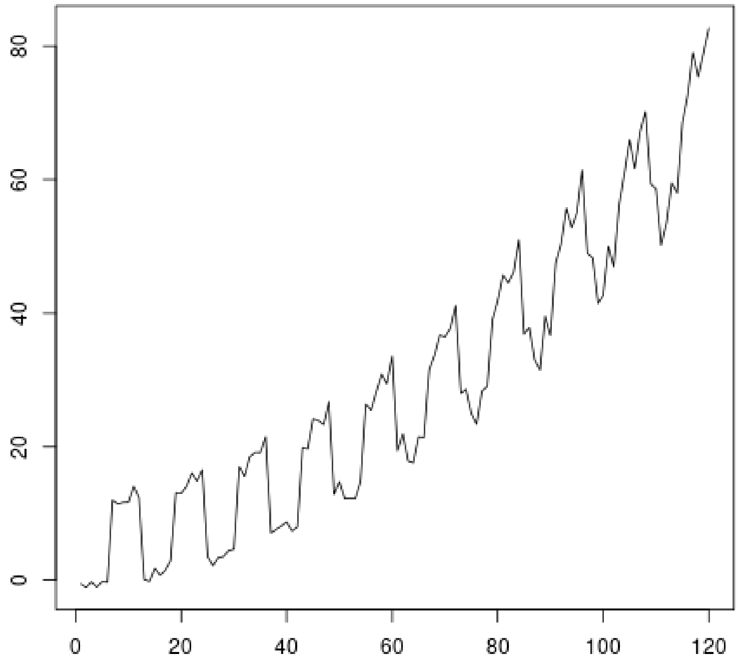
où Y_n est un vecteur colonne p-dimensionnel défini par $Y_n = (X_n, \dots, X_{n-p+1})$. On appelle A la *matrice compagnon* du processus.

3) En déduire une relation valable pour tout $n \geqslant p-1$ entre Y_n, sa valeur initiale Y_{p-1} et le bruit (E_n).

4) En déduire les conditions nécessaires et suffisantes sur les paramètres du modèle garantissant que $\mathbb{E}[Y_n]$ et $\mathbb{V}(Y_n)$ convergent lorsque $n \to +\infty$.

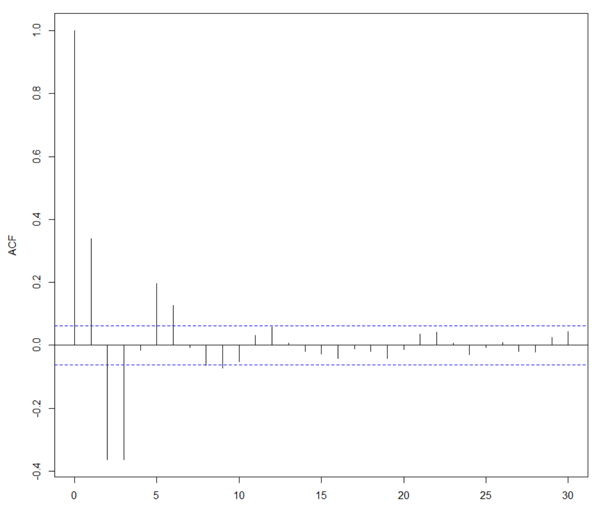
QCM Exam 2022-23.

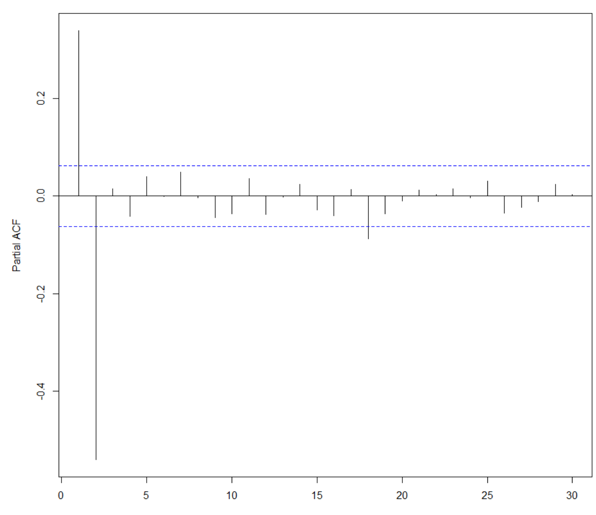
Question 1. Je dois modéliser les données chronologiques représentées ci-dessous.



- [ ] Je commence par réfléchir aux rangs p et q d'une modélisation ARMA(p, q).
- [ ] Un test de stationnarité me semble inutile tellement la réponse est claire.
- [ ] Avant toute chose, vérifions par un test adéquat s'il s'agit ou non d'un bruit blanc.
- [ ] Je peux commencer par estimer puis éliminer une composante déterministe.

Question 2. J'observe les ACF et PACF suivantes pour une série chronologique de taille $n = 200$.





- [ ] Je tente de modéliser la série par un AR(2).

- [ ] Je tente de modéliser la série par un ARMA(6,2).
- [ ] Ces autocorrélations empiriques ne sont pas caractéristiques d'un ARMA.
- [ ] Le 18ème pic de la PACF pourrait me conduire à envisager un modèle saisonnier de période 18.

Question 3. Ma série s'appelle `X`, et j'ai de bonnes raisons de vouloir la modéliser par un SARIMA de période 12. Je demande à `R` quelques tests ADF et KPSS. Les p–values (approximatives) obtenues sont données ci-dessous.

Série	`X`	`diff(X)`	`diff(X,12)`	`diff(diff(X,12))`
ADF	0.562	0.121	$< 10^{-6}$	$< 10^{-6}$
KPSS	$< 10^{-6}$	0.008	0.344	0.122

- [ ] Je commence par tester $(d, D) = (0, 0)$.
- [ ] Le choix $(d, D) = (1, 0)$ est sujet à controverse. Je peux le tester mais une confirmation visuelle s'impose.
- [ ] Il est raisonnable de rejeter l'hypothèse de stationnarité sur `X`.
- [ ] Les choix $(d, D) = (0, 1)$ et $(d, D) = (1, 1)$ semblent pertinents.

Question 4. On considère les résidus d'une modélisation ARMA sur lesquels on veut faire les tests usuels. Relier par une ou des flèche(s) les deux cases de gauche aux deux cases de droite.

On teste la normalité car...	... cela impacte la fiabilité de l'estimation.
On teste la blancheur car...	... le choix des rangs p et q en dépend.

Par ailleurs, parmi les outils suivants, quels sont ceux permettant de vérifier la normalité ?

- [ ] ACF/PACF.
- [ ] Test de Shapiro.
- [ ] QQ-plot.
- [ ] Test KPSS.
- [ ] 95% des valeurs standardisées dans $[\pm 1.96]$.
- [ ] Test de Ljung-Box.
- [ ] Histogramme.
- [ ] Nuage de points entre les valeurs à l'instant $t - 1$ et les valeurs à l'instant t.

Question 5. Ma série s'appelle `X`, je lance la ligne de code suivante sous `R`.

```
> Arima(X, order=c(1,0,0), seasonal=list(order=c(0,1,1), period=6), include.drift=TRUE)
```

- [ ] Je suis en train de demander une estimation AR(1) de période 6.
- [ ] Je vais obtenir l'estimation d'un paramètre moyenne mobile et d'un paramètre autorégressif saisonnier.
- [ ] Si mon modèle est correct, la série que je manipule n'est pas issue d'un processus stationnaire.
- [ ] L'argument `include.drift = TRUE` est totalement inutile ici.

Question 6. Ma série s'appelle `X` et j'ai de bonnes raisons de penser qu'elle est périodique de période 12. Je lance la ligne de code suivante sous `R`.

```
> Res = decompose(ts(X, frequency=12))$random
```

- [ ] J'espère que `Res` ne possède pas d'autocorrélation significative, sinon je dois revoir mon choix de modèle.

- ☐ J'espère que `Res` est périodique de période 12, sinon j'ai mal évalué la périodicité.
- ☐ J'espère que `Res` prend essentiellement des valeurs dans $[\pm 1.96]$.
- ☐ Dans la décomposition additive de `X`, `Res` contient des fluctuations de nature aléatoire.

Question 7. Je demande à `R` une estimation SARIMA de période 12 avec tendance linéaire. J'obtiens la sortie suivante.

```
Coefficients:
         ar1      ar2     ar3     ma1     sar1    sma1    drift
      0.1240  -0.4201  0.2107  0.3222  -0.7807  0.7805   1.1001
s.e.  0.0091   0.0098  0.0553   7e-04   0.0112  0.0054   0.9881
```

- ☐ J'ai demandé un SARIMA de période 12 avec $p = 3$, $q = 1$, $P = 1$, $Q = 1$.
- ☐ Peut-être devrais-je tenter un modèle similaire avec $P = 0$.
- ☐ Peut-être devrais-je tenter un modèle similaire avec `include.drift = FALSE`.
- ☐ J'ai fait attention au préalable à ce que la série que je tente de modéliser est bien stationnaire.

Question 8. Je suis satisfait du triplet (d, D, s) de mon modèle SARIMA, mais je ne sais pas trop comment évaluer les paramètres restants (p, q, P, Q).

- ☐ La commande `auto.arima` permet d'obtenir rapidement un premier modèle de référence.
- ☐ La blancheur du résidu et la significativité des estimateurs sont des bons indicateurs.
- ☐ Je me focalise sur 1 ou 2 critères, selon les besoins de l'étude, et je les optimise.
- ☐ Il suffit d'observer les ACF et PACF de la série différenciée issue du choix (d, D, s).

Question 9. L'intervalle de prédiction...

- ☐ ... est à éviter, cela alourdit les graphiques et n'apporte rien de plus que la prédiction en elle-même.
- ☐ ... dépend de l'horizon de prédiction.
- ☐ ... perd sa propriété de symétrie autour de la prédiction en cas de transformation logarithmique.
- ☐ ... est un gage de qualité lorsqu'il est de grande amplitude.

Question 10. Je réalise un krigeage simple pour déterminer la température du jour à midi dans le Maine-et-Loire à partir d'un ensemble de mesures récoltées sur le territoire. Pour évaluer la qualité du modèle, je compare les températures mesurées avec les températures prédites par le krigeage sur ces emplacements de mesure. Je trouve un MSE $\approx 10^{-16}$ c'est-à-dire, aux approximations numériques près, une erreur nulle.

- ☐ Cela prouve que mon modèle est d'une qualité exceptionnelle.
- ☐ Cela montre qu'il faudrait plutôt envisager un krigeage ordinaire.
- ☐ C'est tout à fait normal mais cela ne me donne aucune information sur la qualité de mon modèle.
- ☐ Une erreur statistique nulle ? Je me suis planté dans mon code.