SÉRIES TEMPORELLES
(Éléments de correction)

UNIVERSITÉ DU MANS (EXAMEN, L3)

EXERCICE 1

(1) Le polynôme retard auto-régressif est :

$$\Phi(L) = 1 - \frac{3}{4}L + \frac{1}{8}L^2.$$

Cherchons ses racines : $\frac{1}{8}L^2 - \frac{3}{4}L + 1 = 0$, soit $L^2 - 6L + 8 = 0$, dont les racines sont $L_1 = 2$ et $L_2 = 4$. Toutes deux sont strictement supérieures à 1 en module, le processus est donc asymptotiquement stationnaire au second ordre. On a d'ailleurs la factorisation :

$$\Phi(L) = \left(1 - \frac{1}{2}L\right)\left(1 - \frac{1}{4}L\right).$$

(2) Sous stationnarité, $\mathbb{E}[Y_t] = \mu$ pour tout t, et l'application de l'espérance à l'équation du processus donne :

$$\mu = \frac{3/8}{1 - 3/4 + 1/8} = \frac{3/8}{3/8} = 1.$$

(3) Notons $\tilde{Y}_t = Y_t - \mu$, qui vérifie :

$$\tilde{Y}_t = \frac{3}{4}\tilde{Y}_{t-1} - \frac{1}{8}\tilde{Y}_{t-2} + \varepsilon_t.$$

En multipliant cette équation par $\tilde{Y}_{t-h}$ et en prenant l'espérance, on obtient pour tout $h \geq 1$ les équations de Yule-Walker :

$$\gamma(h) = \frac{3}{4}\gamma(h-1) - \frac{1}{8}\gamma(h-2), \quad h \geq 1$$

(en utilisant $\gamma(-h) = \gamma(h)$). Pour $h = 0$, on obtient en outre, en multipliant par $\tilde{Y}_t$:

$$\gamma(0) = \frac{3}{4}\gamma(1) - \frac{1}{8}\gamma(2) + \sigma_\varepsilon^2.$$

Pour $h = 1$ **:** $\gamma(1) = \frac{3}{4}\gamma(0) - \frac{1}{8}\gamma(1)$, donc $\frac{9}{8}\gamma(1) = \frac{3}{4}\gamma(0)$, d'où :

$$\gamma(1) = \frac{2}{3}\gamma(0).$$

Pour $h = 2$ **:** $\gamma(2) = \frac{3}{4}\gamma(1) - \frac{1}{8}\gamma(0) = \frac{3}{4} \cdot \frac{2}{3}\gamma(0) - \frac{1}{8}\gamma(0) = \frac{3}{8}\gamma(0)$, soit :

$$\gamma(2) = \frac{3}{8}\gamma(0).$$

Pour $h = 0$ **:** en substituant ces deux relations,

$$\gamma(0) = \frac{3}{4} \cdot \frac{2}{3}\gamma(0) - \frac{1}{8} \cdot \frac{3}{8}\gamma(0) + \sigma_\varepsilon^2,$$

soit $\gamma(0)\left(1 - \frac{1}{2} + \frac{3}{64}\right) = \sigma_\varepsilon^2$, c'est-à-dire $\frac{35}{64}\gamma(0) = \sigma_\varepsilon^2$. Finalement :

$$\gamma(0) = \frac{64}{35}\sigma_\varepsilon^2, \quad \gamma(1) = \frac{128}{105}\sigma_\varepsilon^2, \quad \gamma(2) = \frac{24}{35}\sigma_\varepsilon^2.$$

(4) La relation de récurrence est, pour tout $h \geq 1$:

$$\gamma(h) = \frac{3}{4}\gamma(h-1) - \frac{1}{8}\gamma(h-2).$$

L'équation caractéristique associée est $r^2 - \frac{3}{4}r + \frac{1}{8} = 0$, dont les racines sont $r_1 = \frac{1}{2}$ et $r_2 = \frac{1}{4}$ (inverses des racines L_1, L_2 trouvées à la question 1). La solution générale est donc :

$$\gamma(h) = A\left(\frac{1}{2}\right)^h + B\left(\frac{1}{4}\right)^h,$$

les constantes A et B étant déterminées par les conditions initiales $\gamma(0) = \frac{64}{35}\sigma_\varepsilon^2$ et $\gamma(1) = \frac{128}{105}\sigma_\varepsilon^2$. La résolution du système linéaire $A + B = \frac{64}{35}$, $\frac{A}{2} + \frac{B}{4} = \frac{128}{105}$ donne $A = \frac{64}{21}\sigma_\varepsilon^2$ et $B = -\frac{128}{105}\sigma_\varepsilon^2$. On a donc finalement :

$$\gamma(h) = \frac{64}{21}\left(\frac{1}{2}\right)^h \sigma_\varepsilon^2 - \frac{128}{105}\left(\frac{1}{4}\right)^h \sigma_\varepsilon^2.$$

On note que $\gamma(h)$ tend vers 0 lorsque $h \to \infty$, ce qui est cohérent avec la stationnarité du processus.

EXERCICE 2

(1) Le polynôme retard auto-régressif est $1 + \frac{1}{2}L$, dont la racine est $L = -2$, supérieure à 1 en module : le processus est asymptotiquement stationnaire au second ordre. Le polynôme retard moyenne mobile est $1 + \frac{1}{4}L$, dont la

racine est $L = -4$, également supérieure à 1 en module : le processus est inversible. Ces deux racines étant distinctes, on travaille bien sur la représentation minimale d'un ARMA(1,1).

(2) Sous stationnarité au second ordre, l'espérance et la fonction d'autocovariance sont invariantes dans le temps : $\mathbb{E}[y_t] = \mu$ et $\mathbb{E}[(y_t - \mu)(y_{t-h} - \mu)] = \gamma(h)$ ne dépendent pas de t.

(3) En appliquant l'espérance à l'équation du processus :

$$\mu = 1 - \frac{1}{2}\mu \Longrightarrow \mu = \frac{2}{3}.$$

(4) Posons $\tilde{y}_t = y_t - \mu$. Alors :

$$\tilde{y}_t = -\frac{1}{2}\tilde{y}_{t-1} + \varepsilon_t + \frac{1}{4}\varepsilon_{t-1}.$$

On a besoin de $\mathbb{E}[\tilde{y}_t\varepsilon_t]$ et $\mathbb{E}[\tilde{y}_t\varepsilon_{t-1}]$. Comme $\tilde{y}_t$ est, par inversion du polynôme retard auto-régressif, une combinaison linéaire des innovations courante et passées, on obtient directement :

$$\mathbb{E}[\tilde{y}_t\varepsilon_t] = \sigma_\varepsilon^2 = 1.$$

Pour $\mathbb{E}[\tilde{y}_t\varepsilon_{t-1}]$, on substitue :

$$\begin{aligned}\mathbb{E}[\tilde{y}_t\varepsilon_{t-1}] &= -\frac{1}{2}\mathbb{E}[\tilde{y}_{t-1}\varepsilon_{t-1}] + 0 + \frac{1}{4}\sigma_\varepsilon^2 \\ &= -\frac{1}{2} + \frac{1}{4} = -\frac{1}{4}.\end{aligned}$$

En multipliant l'équation de $\tilde{y}_t$ par $\tilde{y}_t$ puis par $\tilde{y}_{t-1}$ et en prenant les espérances :

$$\gamma(0) = -\frac{1}{2}\gamma(1) + 1 + \frac{1}{4}\cdot\left(-\frac{1}{4}\right) = -\frac{1}{2}\gamma(1) + \frac{15}{16},$$

$$\gamma(1) = -\frac{1}{2}\gamma(0) + 0 + \frac{1}{4}\cdot 1 = -\frac{1}{2}\gamma(0) + \frac{1}{4}.$$

En substituant la seconde dans la première, on obtient :

$$\gamma(0) = -\frac{1}{2}\left(-\frac{1}{2}\gamma(0) + \frac{1}{4}\right) + \frac{15}{16} = \frac{1}{4}\gamma(0) + \frac{13}{16},$$

soit $\frac{3}{4}\gamma(0) = \frac{13}{16}$, d'où :

$$\gamma(0) = \frac{13}{12}, \quad \gamma(1) = -\frac{1}{2}\cdot\frac{13}{12} + \frac{1}{4} = -\frac{7}{24}.$$

(5) Pour $h \geq 2$, l'innovation ε_{t-1} n'est plus corrélée avec $\tilde{y}_{t-h}$ (qui ne dépend que de $\varepsilon_{t-h}, \varepsilon_{t-h-1}, \dots$). En multipliant l'équation de $\tilde{y}_t$ par $\tilde{y}_{t-h}$ et en prenant l'espérance, il vient :

$$\gamma(h) = -\frac{1}{2}\gamma(h-1), \quad h \geq 2.$$

Par itération, en partant de $\gamma(1) = -\frac{7}{24}$:

$$\gamma(h) = \left(-\frac{1}{2}\right)^{h-1} \gamma(1) = -\frac{7}{24}\left(-\frac{1}{2}\right)^{h-1}, \quad h \geq 1.$$

En particulier $\gamma(2) = \frac{7}{48}$, $\gamma(3) = -\frac{7}{96}$, etc.

(6) L'autocorrélation est $\rho(h) = \gamma(h)/\gamma(0)$. Comme $\gamma(0) > 0$, le signe de $\rho(h)$ est celui de $\gamma(h)$. La récurrence $\gamma(h) = -\frac{1}{2}\gamma(h-1)$ implique que $\gamma(h)$ change de signe à chaque pas pour $h \geq 1$. Comme par ailleurs $\gamma(1) < 0$, on a $\rho(h) > 0$ pour h pair et $\rho(h) < 0$ pour h impair (sans oublier $\rho(0) = 1 > 0$). La fonction d'autocorrélation est donc *alternée*, avec une décroissance géométrique en module de raison $\frac{1}{2}$.

EXERCICE 3

(1) Notons f la densité jointe de l'échantillon. En appliquant le théorème de Bayes de manière itérée :

$$\begin{aligned} f(\mathcal{Y}_T; \boldsymbol{\theta}) = {} & f(y_1, y_2; \boldsymbol{\theta}) \\ & \times \prod_{t=3}^{T} f(y_t \mid y_{t-1}, y_{t-2}, \dots, y_1; \boldsymbol{\theta}). \end{aligned}$$

Or, dans le cas d'un AR(2), y_t ne dépend que de (y_{t-1}, y_{t-2}) une fois conditionné par le passé, et l'innovation ε_t est gaussienne : la loi conditionnelle est $y_t \mid y_{t-1}, y_{t-2} \sim \mathcal{N}(c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2},\, \sigma_\varepsilon^2)$. La vraisemblance exacte s'écrit donc :

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}; \mathcal{Y}_T) = {} & f(y_1, y_2; \boldsymbol{\theta}) \\ & \times \prod_{t=3}^{T} \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} e^{-\frac{(y_t - c - \varphi_1 y_{t-1} - \varphi_2 y_{t-2})^2}{2\sigma_\varepsilon^2}}. \end{aligned}$$

Difficulté par rapport à l'AR(1) : dans le cas d'un AR(1), la loi marginale de l'observation initiale y_1 est simplement la loi stationnaire univariée $\mathcal{N}(\mu, \gamma(0))$.

Ici, il faut écrire la loi *jointe* de (y_1, y_2), qui sous stationnarité est une loi normale bivariée d'espérance $(\mu, \mu)'$ et de matrice de variance-covariance :

$$\Sigma_2 = \begin{pmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{pmatrix},$$

où $\gamma(0)$ et $\gamma(1)$ sont des fonctions *non linéaires* des paramètres $(\varphi_1, \varphi_2, \sigma_\varepsilon^2)$, déterminées par les équations de Yule-Walker. C'est cette dépendance non linéaire qui complique l'écriture (et l'optimisation numérique) de la vraisemblance exacte.

(2) La vraisemblance conditionnelle s'obtient en traitant (y_1, y_2) comme déterministes (« données ») plutôt qu'aléatoires :

$$\mathcal{L}_c(\boldsymbol{\theta}; \mathcal{Y}_T \mid y_1, y_2) = \prod_{t=3}^{T} f(y_t \mid y_{t-1}, y_{t-2}; \boldsymbol{\theta}).$$

Soit explicitement :

$$\mathcal{L}_c = \prod_{t=3}^{T} \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} e^{-\frac{(y_t - c - \varphi_1 y_{t-1} - \varphi_2 y_{t-2})^2}{2\sigma_\varepsilon^2}}.$$

Elle diffère de la vraisemblance exacte par l'*absence du facteur* $f(y_1, y_2; \boldsymbol{\theta})$. Son intérêt pratique est double : (i) elle ne fait intervenir les paramètres que de manière *linéaire* dans la moyenne conditionnelle, ce qui rend l'optimisation beaucoup plus simple ; (ii) la perte d'information ($f(y_1, y_2)$) est négligeable lorsque T est grand, et l'estimateur du maximum de vraisemblance conditionnel reste convergent et asymptotiquement équivalent à celui du maximum de vraisemblance exacte.

(3) La log-vraisemblance conditionnelle est :

$$\begin{aligned} \ln \mathcal{L}_c = & -\frac{T-2}{2} \ln(2\pi\sigma_\varepsilon^2) \\ & - \frac{1}{2\sigma_\varepsilon^2} \sum_{t=3}^{T} (y_t - c - \varphi_1 y_{t-1} - \varphi_2 y_{t-2})^2. \end{aligned}$$

À σ_ε^2 fixé, seul le second terme dépend de $(c, \varphi_1, \varphi_2)$; sa maximisation revient à *minimiser* la somme des carrés des résidus :

$$S(c, \varphi_1, \varphi_2) = \sum_{t=3}^{T} (y_t - c - \varphi_1 y_{t-1} - \varphi_2 y_{t-2})^2.$$

C'est exactement le critère des moindres carrés ordinaires appliqué à la régression de y_t sur la constante, y_{t-1} et y_{t-2}, à partir de $t = 3$. En posant :

$$\boldsymbol{y} = \begin{pmatrix} y_3 \\ y_4 \\ \vdots \\ y_T \end{pmatrix}, \quad \boldsymbol{X} = \begin{pmatrix} 1 & y_2 & y_1 \\ 1 & y_3 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & y_{T-1} & y_{T-2} \end{pmatrix},$$

$\boldsymbol{X}$ étant de dimension $(T-2) \times 3$, l'estimateur s'écrit :

$$\widehat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{c} \\ \hat{\varphi}_1 \\ \hat{\varphi}_2 \end{pmatrix} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}.$$

On retrouve enfin l'estimateur de σ_ε^2 en remplaçant les paramètres par leurs estimations dans la condition de premier ordre par rapport à σ_ε^2 :

$$\widehat{\sigma_\varepsilon^2} = \frac{1}{T-2}\sum_{t=3}^{T}\left(y_t - \hat{c} - \hat{\varphi}_1 y_{t-1} - \hat{\varphi}_2 y_{t-2}\right)^2.$$

EXERCICE 4

(1) L'équation du processus latent peut s'écrire $(1-\rho L)x_t = \eta_t$. En appliquant l'opérateur $(1-\rho L)$ à l'équation de mesure $y_t = x_t + \varepsilon_t$, on obtient :

$$(1-\rho L)y_t = (1-\rho L)x_t + (1-\rho L)\varepsilon_t,$$

soit :

$$(1-\rho L)y_t = \eta_t + \varepsilon_t - \rho\varepsilon_{t-1}.$$

Notons $\xi_t \equiv \eta_t + \varepsilon_t - \rho\varepsilon_{t-1}$ le membre de droite.

(2) Calculons les autocovariances de ξ_t. Comme η et ε sont des bruits blancs indépendants entre eux et à toutes les dates :

$$\gamma_\xi(0) = \mathrm{Var}(\eta_t) + \mathrm{Var}(\varepsilon_t) + \rho^2\mathrm{Var}(\varepsilon_{t-1}) = \sigma_\eta^2 + (1+\rho^2)\sigma_\varepsilon^2.$$

$$\gamma_\xi(1) = \mathbb{E}[(\eta_t + \varepsilon_t - \rho\varepsilon_{t-1})(\eta_{t-1} + \varepsilon_{t-1} - \rho\varepsilon_{t-2})] = -\rho\sigma_\varepsilon^2.$$

Pour $h \geq 2$, tous les croisements font intervenir des innovations à des dates distinctes, donc indépendantes : $\gamma_\xi(h) = 0$.

La fonction d'autocovariance de ξ_t est donc identique à celle d'un MA(1) : $\gamma_\xi(h) = 0$ pour $h \geq 2$, et $\gamma_\xi(0) \neq 0$, $\gamma_\xi(1) \neq 0$ (pour $\rho \neq 0$). On peut donc

écrire $\xi_t = u_t - \theta u_{t-1}$ avec u_t un bruit blanc de variance σ_u^2, où θ et σ_u^2 sont déterminés par :

$$\begin{cases} (1+\theta^2)\sigma_u^2 = \sigma_\eta^2 + (1+\rho^2)\sigma_\varepsilon^2, \\ -\theta\,\sigma_u^2 = -\rho\sigma_\varepsilon^2, \end{cases}$$

soit :

$$\boxed{\begin{cases} (1+\theta^2)\sigma_u^2 = \sigma_\eta^2 + (1+\rho^2)\sigma_\varepsilon^2, \\ \theta\,\sigma_u^2 = \rho\,\sigma_\varepsilon^2. \end{cases}}$$

La condition d'inversibilité $|\theta| < 1$ permet de sélectionner l'unique solution acceptable.

(3) En substituant la représentation MA(1) de ξ_t dans l'équation $(1-\rho L)y_t = \xi_t$, on obtient :

$$y_t = \rho y_{t-1} + u_t - \theta u_{t-1},$$

ce qui est bien un processus ARMA(1,1) de paramètre auto-régressif ρ, de paramètre moyenne mobile θ et d'innovation u_t de variance σ_u^2, ces deux dernières quantités étant déterminées par le système ci-dessus.

(4) La racine du polynôme retard auto-régressif est $1/\rho$, strictement supérieure à 1 en module sous l'hypothèse $|\rho| < 1$: le processus y_t est asymptotiquement stationnaire au second ordre. Par construction (voir la question 2), θ peut être choisi tel que $|\theta| < 1$, ce qui assure également l'inversibilité de la représentation ARMA(1,1).

Remarque. Cet exercice illustre une situation très courante : *une dynamique AR(1) observée avec une erreur de mesure devient un ARMA(1,1)*. Plus généralement, l'observation bruitée d'un AR(p) donne un ARMA(p,p). Ce résultat motive l'usage des représentations état-mesure et du filtre de Kalman, qui permettent de travailler directement avec le modèle d'origine plutôt qu'avec sa forme ARMA réduite.