Correction – TD/TP 4
Modèles MA, AR et ARMA stationnaires

M1 Actuariat – Université Gustave Eiffel

Exercice 1 – Une même dépendance, deux écritures MA(1)

1) Pour $X_t = Z_t + \frac{3}{2}Z_{t-1}$,

$$\gamma_X(0) = 1 + \left(\frac{3}{2}\right)^2 = \frac{13}{4}, \qquad \gamma_X(1) = \frac{3}{2}.$$

Ainsi

$$\rho_X(1) = \frac{\gamma_X(1)}{\gamma_X(0)} = \frac{3/2}{13/4} = \frac{6}{13} \simeq 0.462.$$

2) Pour le second MA(1),

$$\gamma(0) = \tau^2(1+\eta^2), \qquad \gamma(1) = \tau^2\eta.$$

Donc nécessairement

$$\rho(1) = \frac{\eta}{1+\eta^2} = \frac{6}{13}.$$

On résout

$$6\eta^2 - 13\eta + 6 = 0.$$

Le discriminant vaut 25, donc

$$\eta = \frac{13 \pm 5}{12},$$

soit

$$\eta = \frac{3}{2} \qquad \text{ou} \qquad \eta = \frac{2}{3}.$$

3) L'écriture inversible d'un MA(1) vérifie $|\eta| < 1$. On choisit donc

$$\eta = \frac{2}{3}.$$

Pour conserver $\gamma(0) = 13/4$, il faut

$$\tau^2\left(1 + \frac{4}{9}\right) = \frac{13}{4},$$

d'où

$$\tau^2 \frac{13}{9} = \frac{13}{4}, \qquad \boxed{\tau^2 = \frac{9}{4}}.$$

On vérifie alors

$$\gamma(1) = \frac{9}{4} \times \frac{2}{3} = \frac{3}{2}.$$

4) Les deux modèles ont exactement les mêmes autocovariances, bien que leurs coefficients MA soient différents. L'ACF ne permet donc pas de choisir entre eux. La condition d'inversibilité impose une représentation canonique, ici celle avec $\eta = 2/3$.

Exercice 2 – Analyser la dynamique d'un AR(2)

1) Le polynôme autorégressif est

$$\Phi(z) = 1 - 1.1z + 0.3z^2.$$

Il se factorise sous la forme

$$\Phi(z) = (1 - 0.5z)(1 - 0.6z).$$

Ses racines sont 2 et 5/3, toutes deux de module supérieur à 1. Le modèle est donc causal.

2) Avec $X_{-2} = X_{-1} = 0$, $Z_0 = 1$ et aucun autre choc,

$$\begin{aligned}
X_0 &= 1,\\
X_1 &= 1.1,\\
X_2 &= 1.1(1.1) - 0.3(1) = 0.91,\\
X_3 &= 1.1(0.91) - 0.3(1.1) = 0.671,\\
X_4 &= 1.1(0.671) - 0.3(0.91) = 0.4651,\\
X_5 &= 1.1(0.4651) - 0.3(0.671) = 0.31031.
\end{aligned}$$

Le choc ne disparaît pas après une ou deux périodes : il se propage dans le temps, mais son effet décroît progressivement.

3) En écrivant

$$\frac{1}{\Phi(B)} = \sum_{j\geq 0} \psi_j B^j,$$

l'identité

$$(1 - 1.1B + 0.3B^2) \sum_{j\geq 0} \psi_j B^j = 1$$

donne, par identification des coefficients,

$$\psi_0 = 1, \qquad \psi_1 = 1.1,$$

et, pour $j \geq 2$,

$$\psi_j = 1.1\psi_{j-1} - 0.3\psi_{j-2}.$$

Les nombres $\psi_0, \psi_1, \ldots$ sont exactement les réponses successives du système à un choc unitaire isolé. On retrouve donc les valeurs calculées à la question précédente.

4) Comme

$$\frac{1}{(1 - 0.5z)(1 - 0.6z)} = \frac{A}{1 - 0.5z} + \frac{B}{1 - 0.6z},$$

on impose

$$A + B = 1, \qquad 0.6A + 0.5B = 0.$$

On obtient $A = -5$ et $B = 6$. Ainsi

$$\frac{1}{\Phi(z)} = -5 \sum_{j\geq 0} (0.5z)^j + 6 \sum_{j\geq 0} (0.6z)^j,$$

d'où

$$\boxed{\psi_j = -5(0.5)^j + 6(0.6)^j.}$$

Les deux termes tendent vers zéro : l'effet d'un choc ancien s'éteint asymptotiquement.

5) En multipliant
$$X_t = 1.1X_{t-1} - 0.3X_{t-2} + Z_t$$
par X_{t-1} et en prenant l'espérance, le terme Z_tX_{t-1} a une espérance nulle car Z_t est orthogonal au passé. On obtient
$$\gamma(1) = 1.1\gamma(0) - 0.3\gamma(1).$$
Donc
$$1.3\gamma(1) = 1.1\gamma(0), \qquad \boxed{\rho(1) = \frac{1.1}{1.3} \simeq 0.846.}$$
Pour $h \geq 2$, le même raisonnement donne
$$\gamma(h) = 1.1\gamma(h-1) - 0.3\gamma(h-2),$$
donc
$$\rho(h) = 1.1\rho(h-1) - 0.3\rho(h-2).$$
Ainsi
$$\begin{aligned}\rho(2) &\simeq 1.1(0.846) - 0.3 = 0.631,\\ \rho(3) &\simeq 1.1(0.631) - 0.3(0.846) = 0.440,\\ \rho(4) &\simeq 1.1(0.440) - 0.3(0.631) = 0.295.\end{aligned}$$

6) L'ACF doit décroître progressivement : elle ne se coupe pas après un retard fini. Pour un AR(2), la PACF théorique se coupe après le retard 2. C'est précisément la signature que l'on cherchera ensuite sur les données.

Exercice 3 – Remonter d'une dépendance observée vers un modèle

1) Pour
$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t,$$
on multiplie d'abord par X_{t-1}. Comme Z_t est orthogonal au passé,
$$\gamma(1) = \phi_1\gamma(0) + \phi_2\gamma(1).$$
En divisant par $\gamma(0)$,
$$\rho(1) = \phi_1 + \phi_2\rho(1).$$
De même, en multipliant par X_{t-2},
$$\gamma(2) = \phi_1\gamma(1) + \phi_2\gamma(0),$$
donc
$$\rho(2) = \phi_1\rho(1) + \phi_2.$$

2) On remplace les autocorrélations théoriques par les valeurs empiriques :
$$\begin{aligned}0.80 &= \phi_1 + 0.80\phi_2,\\ 0.50 &= 0.80\phi_1 + \phi_2.\end{aligned}$$
La première équation donne $\phi_1 = 0.80 - 0.80\phi_2$. En remplaçant dans la seconde,
$$0.50 = 0.64 - 0.64\phi_2 + \phi_2,$$
d'où
$$0.36\phi_2 = -0.14, \qquad \widehat{\phi}_2 \simeq -0.389.$$
Puis
$$\widehat{\phi}_1 \simeq 0.80 - 0.80(-0.389) \simeq 1.111.$$

3) Le polynôme estimé est approximativement

$$\widehat{\Phi}(z) = 1 - 1.111z + 0.389z^2.$$

Ses racines sont complexes conjuguées, approximativement $1.43 \pm 0.73\,i$, et leur module vaut environ $1.60 > 1$. Le modèle estimé est donc causal.

4) Non. Une ACF empirique est bruitée et plusieurs familles de modèles peuvent produire des décroissances assez proches. La PACF apporte ici une information structurelle supplémentaire : une coupure après le retard 2 est caractéristique d'un AR(2) théorique et rend ce choix beaucoup plus plausible.

Exercice 4 – TP : identifier trois séries sans connaître leur modèle

Les données ont été générées de la manière suivante.

— série A : MA(2), $X_t = Z_t + 0.8Z_{t-1} + 0.6Z_{t-2}$;

— série B : AR(2), $X_t = 1.1X_{t-1} - 0.3X_{t-2} + Z_t$;

— série C : ARMA(1,1), $X_t = 0.65X_{t-1} + Z_t + 0.45Z_{t-1}$.

1) Les chronogrammes montrent des dépendances temporelles différentes, mais ils ne permettent pas de lire directement et sûrement l'ordre du modèle. C'est précisément pourquoi on utilise l'ACF et la PACF.

2) Les valeurs obtenues peuvent fluctuer légèrement selon la convention numérique utilisée, mais sur le fichier fourni on observe approximativement :

	1	2	3	4	5	6	7	8
ACF A	.67	.39	.08	.02	-.04	-.06	-.03	-.02
PACF A	.67	-.10	-.26	.19	-.05	-.10	.12	-.04
ACF B	.80	.51	.26	.06	-.07	-.14	-.18	-.18
PACF B	.80	-.36	-.04	-.07	-.02	-.05	-.05	.02
ACF C	.79	.51	.35	.28	.22	.18	.15	.11
PACF C	.79	-.31	.19	-.01	.00	.05	-.01	-.05

3) L'interprétation attendue est :

Série	Forme de l'ACF	Forme de la PACF	Modèle candidat
A	coupure nette après 2, à l'erreur d'échantillonnage près	décroissance/oscillations	MA(2)
B	décroissance progressive	essentiellement deux premiers retards	AR(2)
C	décroissance progressive	décroissance progressive	ARMA(1,1)

4) Pour la série B, les deux premières autocorrélations empiriques sont approximativement

$$\widehat{\rho}(1) \simeq 0.80, \qquad \widehat{\rho}(2) \simeq 0.51.$$

On résout donc

$$0.80 = \phi_1 + 0.80\phi_2,$$
$$0.51 = 0.80\phi_1 + \phi_2.$$

Cela donne des valeurs proches de

$$\widehat{\phi}_1 \simeq 1.08, \qquad \widehat{\phi}_2 \simeq -0.35,$$

avec des écarts dus au fait que l'on remplace les autocorrélations théoriques par des estimations empiriques sur 320 observations. Les vraies valeurs utilisées pour générer la série sont 1.1 et -0.3.

5) Avec ces valeurs estimées, on calcule les racines de

$$1 - \widehat{\phi}_1 z - \widehat{\phi}_2 z^2.$$

Leurs modules restent supérieurs à 1 : le modèle estimé est causal. Avec les vraies valeurs, les racines sont exactement 2 et 5/3.

6) Pour un MA(2), l'ACF *théorique* est nulle après le retard 2. L'ACF empirique, elle, est calculée sur un échantillon fini : elle n'est presque jamais exactement nulle. De petites barres après le retard 2 sont donc normales et ne contredisent pas le modèle si elles restent compatibles avec la variabilité d'échantillonnage.

7) Pour un ARMA(1,1), ni l'ACF ni la PACF ne se coupent en général après un nombre fini de retards. Les deux décroissent progressivement. On peut donc reconnaître une signature compatible avec un ARMA, mais on ne peut pas lire directement (p, q) comme on lit q sur l'ACF d'un MA(q) ou p sur la PACF d'un AR(p).

Idée à retenir

Le point important de ce TP est le changement de sens du raisonnement : on ne part plus d'un modèle connu pour calculer son ACF/PACF ; on part de ce que l'on observe pour proposer une structure de modèle, puis on vérifie qu'elle est cohérente avec les propriétés théoriques du cours.