

Séries temporelles, examen, ÉLÉMENTS DE CORRECTION.

ISFA3, année 2007-2008

Mercredi 2 Avril 2008.

Durée 2h, notes de cours et de TD, calculatrices autorisés.

Vous veillerez à justifier soigneusement vos réponses.

Dans toute la suite, L désigne l'opérateur de décalage $L(X_t) = X_{t-1}$ et $\Delta_d = (1 - L^d)$ l'opérateur de différence d'ordre d , on notera $\Delta_1 = \Delta$.

Exercice 1 On considère $(X_t)_{t \in \mathbb{Z}}$ et $(Y_t)_{t \in \mathbb{Z}}$ définis par :

$$X_t - \alpha X_{t-1} = W_t \text{ et } Y_t - \alpha Y_{t-1} = X_t + Z_t$$

où $|\alpha| < 1$, $(W_t)_{t \in \mathbb{Z}}$ et $(Z_t)_{t \in \mathbb{Z}}$ sont des bruits blancs faibles de variance σ^2 , non corrélés entre eux.

Dans cet exercice, il n'est pas nécessaire de calculer les auto-covariances de (X_t) et (Y_t) .

1. Montrer que $(X_t)_{t \in \mathbb{Z}}$ est faiblement stationnaire.

$(X_t)_{t \in \mathbb{Z}}$ s'écrit comme un processus auto régressif AR(1), de polynôme : $A(z) = 1 - \alpha z$. Ce polynôme a comme racine $z_1 = \frac{1}{\alpha}$, $|z_1| > 1$. Ainsi, il existe une unique solution faiblement stationnaire $(X_t)_{t \in \mathbb{Z}}$.

2. Écrire X_t sous la forme d'une $MA(\infty)$, calculer $Var(X_0)$.

On sait que $(X_t)_{t \in \mathbb{Z}}$ s'écrit sous la forme :

$$X_t = \sum_{i \geq 0} \psi_i W_{t-i},$$

avec les ψ_i tels que

$$\sum_{i \geq 0} \psi_i z^i = \frac{1}{A(z)}.$$

On a

$$\frac{1}{A(z)} = \frac{1}{1 - \alpha z} = \sum_{i \geq 0} (\alpha z)^i.$$

Ainsi,

$$X_t = \sum_{i \geq 0} \alpha^i W_{t-i}.$$

Comme $(W_t)_{t \in \mathbb{Z}}$ est un bruit blanc faible, les W_i sont non corrélés entre eux et on a alors $Var(X_0) = \sigma^2 \sum_{i \geq 0} |\alpha|^{2i} = \frac{\sigma^2}{1 - |\alpha|^2}$.

3. Déterminer la densité spectrale de (X_t) .

La densité spectrale d'un processus filtré $A(L)X_t$ est $|A(e^{iw})|^2 f_X(w)$, la densité spectrale d'un bruit blanc faible de variance σ^2 est $\frac{\sigma^2}{2\pi}$, on a donc :

$$f_X(w) = \frac{\sigma^2}{2\pi |A(e^{iw})|^2} = \frac{\sigma^2}{2\pi (1 - 2\alpha \cos w + \alpha^2)}.$$

4. Montrer que $W_t - \alpha Z_{t-1} + Z_t$ est une $MA(1)$.

Pour montrer qu'un processus $A_t = W_t - \alpha Z_{t-1} + Z_t$ est une $MA(1)$, il suffit de montrer qu'il est faiblement stationnaire et que $\gamma_A(h) = 0$ pour $|h| > 1$. On calcule :

$$\text{cov}(A_{t+h}, A_t) = \text{cov}(W_{t+h} - \alpha Z_{t+h-1} + Z_{t+h}, W_t - \alpha Z_{t-1} + Z_t).$$

Comme $(W_t)_{t \in \mathbb{Z}}$ et $(Z_t)_{t \in \mathbb{Z}}$ sont des bruits blancs faibles non corrélés entre eux, on obtient :

$$\text{cov}(A_{t+h}, A_t) = 0 \text{ si } |h| > 1,$$

$$\text{cov}(A_t, A_t) = \text{Var}(A_t) = \sigma^2(2 + |\alpha|^2).$$

$$\text{cov}(A_{t+1}, A_t) = -\alpha\sigma^2.$$

Ceci montre que les processus $(A_t)_{t \in \mathbb{Z}}$ est faiblement stationnaire et vérifie $\gamma_A(h) = 0$ pour $|h| > 1$, c'est donc une $MA(1)$.

5. Montrer que Y_t s'écrit sous la forme d'un processus ARMA stationnaire, causal et inversible.

On a montré que $W_t - \alpha Z_{t-1} + Z_t$ est une $MA(1)$ qui s'écrit donc $W_t - \alpha Z_{t-1} + Z_t = B_t + aB_{t-1}$ où $(B_t)_{t \in \mathbb{Z}}$ est un bruit blanc faible. On a :

$$\begin{aligned} Y_t - \alpha Y_{t-1} &= X_t + Z_t \\ &= W_t + \alpha X_{t-1} + Z_t \\ &= W_t + \alpha(Y_{t-1} - \alpha Y_{t-2} - Z_{t-1}) + Z_t \\ &= W_t - \alpha Z_{t-1} + Z_t + \alpha Y_{t-1} - \alpha^2 Y_{t-2}. \end{aligned}$$

Ainsi, le processus $(Y_t)_{t \in \mathbb{Z}}$ vérifie :

$$Y_t - 2\alpha Y_{t-1} + \alpha^2 Y_{t-2} = W_t - \alpha Z_{t-1} + Z_t = B_t + aB_{t-1}.$$

Ainsi, $(Y_t)_{t \in \mathbb{Z}}$ est un ARMA(2, 1) qui admet une unique solution stationnaire car le polynôme $P(z) = 1 - 2\alpha z + \alpha^2 z^2$ a toutes ses racines de module 1. Pour montrer que le processus est inversible, il faut vérifier que $|a| < 1$ ce qui découle du calcul de $\text{Var}(A_t)$.

6. Déterminer la densité spectrale de $X_t + Z_t$ puis celle Y_t . Les processus $(X_t)_{t \in \mathbb{Z}}$ et $(Z_t)_{t \in \mathbb{Z}}$ sont non corrélés. On a donc que

$$f_{X+Z}(w) = f_X(w) + f_Z(w) = \frac{\sigma^2}{2\pi} \left[\frac{1}{(1 - 2\alpha \cos w + \alpha^2)} + 1 \right].$$

La densité spectrale de Y_t vérifie :

$$|A(e^{iw})|^2 f_Y(w) = f_X(w) + f_Z(w).$$

Ainsi

$$f_Y(w) = \frac{\sigma^2}{2\pi} \left[\frac{2 - 2\alpha \cos w + \alpha^2}{(1 - 2\alpha \cos w + \alpha^2)^2} \right].$$

Exercice 2 Dans cet exercice, $(Z_t)_{t \in \mathbb{Z}}$ est un bruit blanc faible de variance σ^2 .

Soit $(X_t)_{t \in \mathbb{Z}}$ un processus AR(2) déterminé par le polynôme $\Phi(z) = 1 - a_1 z - a_2 z^2$ ($\Phi(L)(X_t) = Z_t$). Montrer que cette représentation est stationnaire et causale si et seulement si les paramètres a_1 et a_2 vérifient les trois conditions suivantes :

1. $a_1 + a_2 < 1$,

2. $a_2 - a_1 < 1$,
3. $|a_2| < 1$.

Pour que la représentation $\Phi(L)(X_t) = Z_t$ soit stationnaire et causale, il faut et il suffit que les racines du polynôme $\Phi(z)$ soient de module > 1 . On note z_1 et z_2 ces racines. On a :

$$\Phi(z) = 1 - a_1 z - a_2 z^2 = -a_2(z - z_1)(z - z_2).$$

Ainsi, $-a_2 z_1 z_2 = 1$. Le fait que z_1 et z_2 soient de module > 1 implique donc que $|a_2| < 1$. On remarque maintenant que $\Phi(0) = 1$, $\Phi(1) = 1 - a_1 - a_2$, $\Phi(-1) = 1 + a_1 - a_2$. Le fait que z_1 et z_2 soient de module > 1 implique alors que $\Phi(1) > 0$ et $\Phi(-1) > 0$, ce qui montre (1) et (2).

Réciproquement, supposons que les conditions (1), (2) et (3) soient vérifiées. On a alors $|z_1 z_2| > 1$, $\Phi(1) > 0$, $\Phi(-1) > 0$, $\Phi(0) = 1$. Si les racines sont réelles, elles sont donc nécessairement de module > 1 . Si les racines sont complexes, elles sont conjuguées ((1) et (2) \Rightarrow a_1 et a_2 sont réels) donc on même module. Comme $|z_1 z_2| > 1$, on conclue que z_1 et z_2 sont de module > 1 .

Exercice 3 $(Z_t)_{t \in \mathbb{Z}}$ est un bruit blanc faible de variance σ^2 . On suppose que $(X_t)_{t \in \mathbb{Z}}$ est un ARIMA(p, d, q) vérifiant :

$$A(L)(1 - L)^d(X_t) = B(L)(Z_t). \quad (1)$$

On considère

$$W_t = X_t + Y_0 + Y_1 t + \cdots + Y_{d-1} t^{d-1}$$

où les Y_0, \dots, Y_{d-1} sont des variables aléatoires arbitraires. Montrer que W_t vérifie aussi (1).

Quelle conséquence a cette propriété sur l'étude des processus ARIMA ? Soit $W_t = X_t + Y_0 + Y_1 t + \cdots + Y_{d-1} t^{d-1}$. On note $Z_t = Y_0 + Y_1 t + \cdots + Y_{d-1} t^{d-1}$. On montre par récurrence sur $k = 1, \dots, d$ que $(1 - L)^k(Z_t)$ est un polynôme (dont les coefficients sont des variables aléatoires, combinaisons linéaires des Y_i) de degré $(d - 1 - k) \Rightarrow (1 - L)^d(Z_t) \equiv 0$.

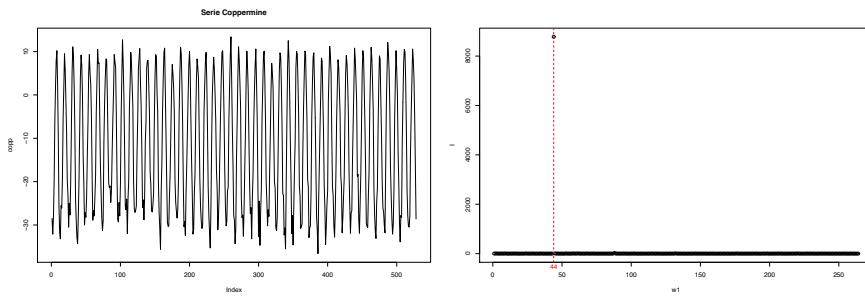
Cette propriété implique qu'un processus de type ARIMA n'est définit qu'à une variable aléatoire du type Z_t près. En particulier, on n'a pas unicité de la solution à l'équation (1), de plus les solutions de (1) ne sont pas en général stationnaires.

Exercice 4 Les graphiques et résultats numériques ci-dessous concernent la série des températures mensuelles relevées à Coppermine - nord du Canada - (**série coppermine**) notée X_t et la série différenciée : $Y_t = \Delta_{12}(X_t)$. La longueur de la série X_t est 528, celle de Y_t est 516.

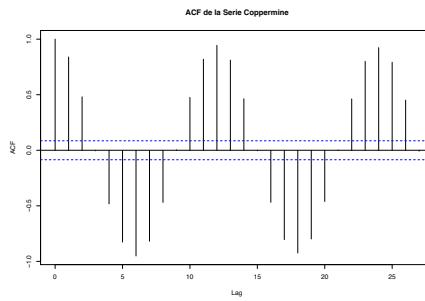
1. Commenter les graphiques et les résultats numériques ci-dessous en répondant notamment aux questions :

- La série coppermine paraît-elle stationnaire ? présente-t-elle une tendance, une saisonnalité ? Mêmes questions pour la série différenciée.
- Justifier le fait que l'on utilise l'opérateur Δ_{12} .
- Quelle modèle peut-on proposer pour la série différenciée Y_t ($AR(p)$, $MA(q)$, $ARMA(p, q)$) ?
- Quel modèle est préférable pour la série Y_t du point de vue du critère :
 - AIC ?
 - BIC ?
 - de l'erreur de prédiction ?

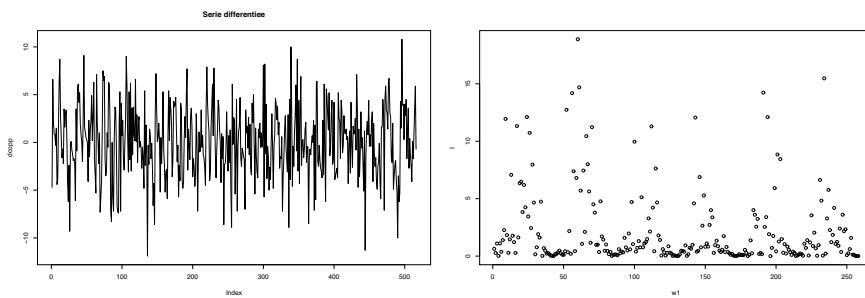
Série Coppermine et le périodogramme



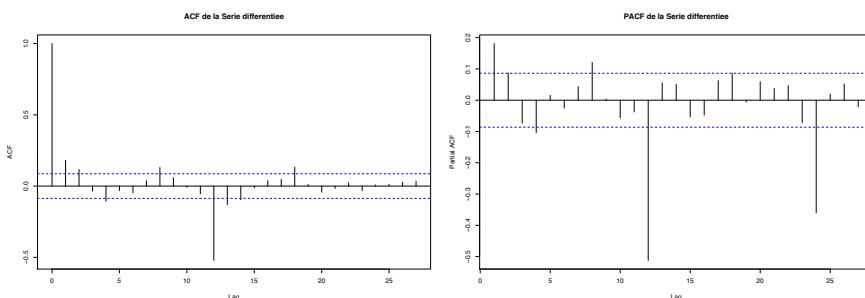
Auto-corrélations de la série Coppermine



Série différenciée et le périodogramme



Auto-corrélations et auto-corrélations partielles de la série différenciée.



Vus les graphiques relatifs à la série Coppermine, celle-ci n'est vraisemblablement pas stationnaire. En effet, le graphe représentant la série elle-même semble indiquer une saisonnalité marquée. Ceci est confirmé par le fait que le périodogramme admet un maximum très largement supérieur aux autres valeurs prises. Ce maximum est atteint en 44, comme la série est de longueur 528, cela signifie qu'il y a probablement une partie saisonnière de période $\frac{528}{44} = 12$. La présence d'une composante périodique est confirmée par le fait que le graphe de l'acp est lui aussi périodique.

La présence de cette composante périodique de période 12 explique que l'on applique l'opérateur Δ_{12} à la série. En effet, l'opérateur Δ_d supprime les composantes saisonnières de période d .

On considère alors les graphes relatifs à la série différentiée $Y_t = \Delta_{12}(X_t)$. Le graphe relatif à la série Y_t elle même ne fait pas apparaître de structure particulière, de même, le périodogramme de la série Y_t ne présente pas de maximum largement supérieur aux autres valuers. Enfin, le graphe de l'acf et de la pacf montrent que les auto-corrélations empiriques et auto-corrélations partielles décroissent rapidement vers 0. On peut donc raisonnablement considérer que (Y_t) est stationnaire. Au vu de l'acf, on pourrait proposer une modélisation de Y_t par une $MA(18)$ ou, si on est moins exigeant sur la confiance, une $MA(12)$ ou une $MA(14)$. Au vu de la pacf, on pourrait proposer de modéliser Y_t par un $AR(12)$. Ces modèles sont cohérence avec les graphes des auto-corrélations empiriques et des auto-corrélations partielles empiriques. Néanmoins, ils comportent beaucoup de paramètres (respectivement 18, 12, 14, 12). Afin de réduire le nombre de paramètres à estimer, on peut proposer des modélisations $ARMA(p, q)$ avec p et $q \leq 3$. On calcule le AIC et le BIC pour chacun des modèles ci-dessous. Rappelons que

$$AIC = -\log V + (p + q) \quad BIC = -\log V + (p + q) \log n$$

où n est la longueur de la série, ici $n = 516$.

Ajustement de la série différenciée à un $ARMA(p, q)$. $\hat{\sigma}_\varepsilon^2$ est la variance estimée de l'innovation du modèle retenu, $\log V$ est le log de la vraisemblance du modèle retenu.

p	q	$\hat{\sigma}_\varepsilon^2$	$\log V$	AIC	BIC
0	0	13.97857	-1412.654	1412.654	1412.654
1	0	13.51929	-1404.051	1405.051	1410.297
2	0	13.41893	-1402.136	1404.136	1414.629
3	0	13.34461	-1400.712	1403.712	1419.450
0	1	13.59923	-1405.567	1406.567	1411.813
1	1	13.47151	-1403.140	1405.140	1415.633
2	1	13.39776	-1401.731	1404.731	1420.469
3	1	13.27229	-1399.319	1403.319	1424.304
0	2	13.33839	-1400.594	1402.594	1413.086
1	2	13.33833	-1400.593	1403.593	1419.331
2	2	13.28581	-1399.581	1403.581	1424.565
3	2	12.21091	-1379.305	1384.305	1410.535
0	3	13.33825	-1400.591	1403.591	1419.329
1	3	12.62783	-1388.189	1392.189	1413.173
2	3	11.98282	-1376.493	1381.493	1407.724
3	3	12.20879	-1379.260	1385.260	1416.736

L'erreur de prédiction d'horizon 1 est donnée par $\hat{\sigma}_\varepsilon^2$. Le tableau ci-dessus fait apparaître que les trois critères : erreur de prédiction d'horizon 1, AIC, BIC sont minimum pour $p = 2$, $q = 3$. Parmi les modèles proposés, on préférera donc le modèle $ARMA(2, 3)$ avec les coefficients estimés par maximum de vraisemblance.

2. On considère maintenant différents modèles SARIMA pour la série Coppermine elle-même. On considère des modèles pouvant s'écrire sous la forme :

$$\Delta_{12}A_p(L)A_P(L^{12})(X_t) = B_q(L)B_Q(L^{12})(\varepsilon_t) \quad (2)$$

où A_k , B_ℓ sont des polynômes de degré respectif k et ℓ , de coefficient constant égal à 1. Combien il y a-t-il de paramètres à estimer pour un tel modèle ?

Le modèle SARIMA ci-dessus à $p + q + P + Q$ paramètres.

Remarque importante Pour le calcul des critères AIC et BIC dans ce cadre, on utilise le nombre de paramètres à estimer à la place de $p + q$ utilisé dans les modèles ARMA.

Le tableau ci-dessous donne la log vraisemblance et la variance estimée des innovations pour des modèles du type (2) pour lesquels $p = 2$, $q = 3$ et différentes valeurs de P et Q .
On calcule le AIC et BIC pour chacun des modèles ci-dessous. Le nombre de paramètres à estimer est $5 + P + Q$.

P	Q	$\hat{\sigma}_\varepsilon^2$	$\log V$	AIC	BIC
0	0	11.984368	-1376.526	1381.526	1407.757
0	1	6.558542	-1232.809	1238.809	1270.285
0	2	6.567932	-1231.538	1238.538	1275.260
0	3	6.562049	-1231.499	1239.499	1281.468
1	0	9.432254	-1311.159	1317.159	1348.636
1	1	7.097006	-1237.768	1244.768	1281.491
2	0	7.999000	-1272.788	1279.788	1316.511
2	1	7.091900	-1237.582	1245.582	1287.551

Quel modèle de type (2) est préférable pour la série X_t du point de vue du critère :

- AIC ?
- BIC ?
- de l'erreur de prédiction ?

C'est le modèle ($p = 2, q = 3, P = 0, Q = 1$) qui minimise l'erreur de prédiction d'horizon 1 et le BIC. Le modèle ($p = 2, q = 3, P = 0, Q = 2$) est légèrement meilleur du point de vue du AIC. On remarque que les modèles SARIMA ci-dessus sont meilleurs que le meilleur des ARMA étudiés précédemment.

Comparer ces résultats avec ceux obtenus si l'on ajuste la série différenciée Y_t à un $MA(12)$ puis à un $ARMA(2, 15)$. Justifier le choix de ces deux modèles pour comparer avec les modèles SARIMA.

On donne ci-dessous les commandes R correspondantes et le résultat obtenu.

Call :

```
arima(x = dcopp, order = c(0, 0, 12))
```

Coefficients :

ma1	ma2	ma3	ma4	ma5	ma6	ma7
0.0696	0.0613	0.0294	0.0271	0.0361	0.0478	0.0530
ma8	ma9	ma10	ma11	ma12	intercept	
0.0524	0.0568	0.0531	0.0794	-0.9174	0.0338	

```
sigma2 estimated as 6.72 : log likelihood = -1240.83.
```

Call :

```
arima(x = dcopp, order = c(2, 0, 15))
```

Coefficients :

ar1	ar2	ma1	ma2	ma3	ma4	ma5	ma6	ma7
0.6742	-0.2482	-0.4840	0.297	-0.0779	0.0167	0.0209	0.0230	0.0148
ma8	ma9	ma10	ma11	ma12	ma13	ma14	ma15	intercept
0.0087	0.0187	0.0254	0.0529	-0.9754	0.4907	-0.2861	0.0733	0.0198

```
sigma2 estimated as 6.405 : log likelihood = -1229.19
```

Les coefficients estimés lorsqu'on ajuste la série X_t à un pour un $SARIMA p = 2, q = 3, P = 0, Q = 1$, puis pour un $SARIMA p = 2, q = 3, P = 0, Q = 2$ sont donnés ci-dessous.

ar1	ar2	ma1	ma2	ma3	sma1
0.8946	0.0731	-0.7139	-0.0485	-0.1728	-0.9638

ar1	ar2	ma1	ma2	ma3	sma1	sma2
0.8725	0.0915	-0.6954	-0.0566	-0.1699	-1.0217	0.0752

On considère un $MA(12)$ car c'est ce qui est suggéré par le graphe d'acf. On considère un $ARMA(2, 15)$ car le meilleur AMRA avec p et $q \leq 3$ est donné pour $p = 2$ et $q = 3$. On essaie de l'améliorer en augmentant q . On aurait pu prendre $q = 12, q = 14, q = 18$ au vu de l'acf.

Pour ces deux nouveaux modèles, on remarque que $\hat{\sigma}_\epsilon^2$ est plus petit que pour les autres ARMA mais du même ordre que les ARIMA étudiés plus haut. On calcule les AIC et BIC correspondant :

pour $AR(12)$: $AIC = 1252.829, BIC = 1315.782$.

pour $ARMA(2, 15)$: $AIC = 1246.19, BIC = 1335.374$.

Du point de vue des critères AIC et BIC, les modèles SARIMA ci-dessus sont préférables. Les deux meilleurs modèles ($Q = 1$ et $Q = 2$) ont nettement moins de paramètres et sont donc plus rapides à calculer.

Le fait que les SARIMA sont préférable se voit aussi dans les tableaux relatifs aux paramètres estimés : pour les modèles $MA(12)$ et $ARMA(2, 15)$, de nombreux paramètres estimés sont dans l'intervalle de confiance $[-\frac{1}{\sqrt{n}} \cdot 1.96, \frac{1}{\sqrt{n}} \cdot 1.96] = [-0.086, 0.086]$ de 0, i.e. peuvent être considérés comme égaux à 0. Bien que l'utilisation de cet intervalle de confiance ne soit pas complètement justifié sans renseignements supplémentaires sur les résidus, cela fait penser que ces modèles ne sont probablement pas pertinents.

3. Finalement, quel modèle proposeriez-vous dans le but d'effectuer des prévisions sur les températures à Coppermine ?

Le modèles SARIMA $p = 2, q = 3, P = 0, Q = 1$ paraît le plus indiqué pour modéliser et faire de prévision sur les températures à Coppermine. On peut noter aussi que le modèle $p = 2, q = 3, P = 0, Q = 2$ minimise le critère AIC et peut aussi être pertinent.