Séries temporelles, examen.

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.
2. Écrire X_t sous la forme d'une $MA(\infty)$, calculer $Var(X_0)$.
3. Déterminer la densité spectrale de (X_t).
4. Montrer que $W_t - \alpha Z_{t-1} + Z_t$ est une $MA(1)$.
5. Montrer que Y_t s'écrit sous la forme d'un processus ARMA stationnaire, causal et inversible.
6. Déterminer la densité spectrale de $X_t + Z_t$ puis celle Y_t.

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. $|a_2| < 1$.

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 :

$$A(L)(1 - L)^d(X_t) = B(L)(Z_t). \tag{1}$$

On considère

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

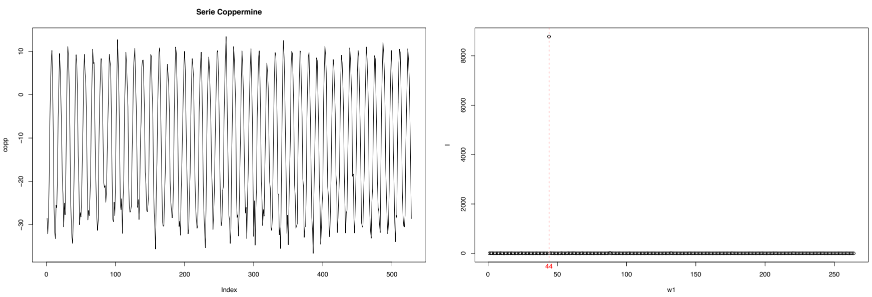
où les Y_0, ..., Y_{d-1} sont des variables aléatoires arbitraires. Montrer que W_t vérifie aussi (1).
Quelle conséquence a cette propriété sur l'étude des processus $ARIMA$?

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.

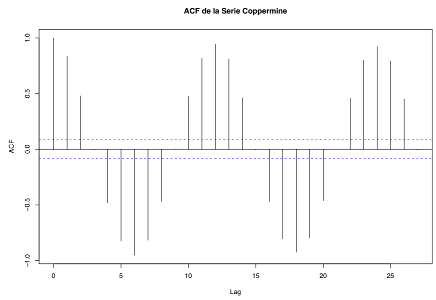
1. Commenter les graphiques et les résultats numériques ci-dessous en répondant notamment aux questions :

- La série coppermine paraît-elle stationnaire ? présente-t-elle une tendance, une saisonalité ? 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 ?

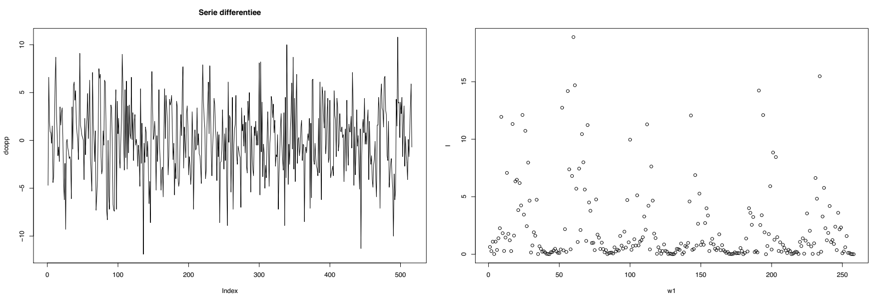
Série Coppermine et le périodogramme



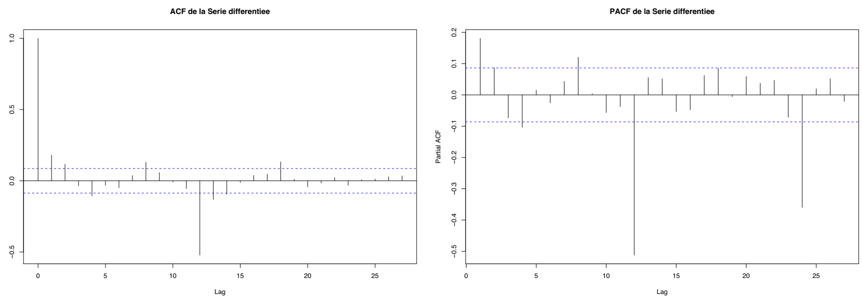
Auto-corrélations de la série Coppermine



Série differenciée et le périodogramme



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



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

p	q	$\widehat{\sigma}^2_\varepsilon$	$\log V$
0	0	13.97857	−1412.654
1	0	13.51929	−1404.051
2	0	13.41893	−1402.136
3	0	13.34461	−1400.712
0	1	13.59923	−1405.567
1	1	13.47151	−1403.140
2	1	13.39776	−1401.731
3	1	13.27229	−1399.319
0	2	13.33839	−1400.594
1	2	13.33833	−1400.593
2	2	13.28581	−1399.581
3	2	12.21091	−1379.305
0	3	13.33825	−1400.591
1	3	12.62783	−1388.189
2	3	11.98282	−1376.493
3	3	12.20879	−1379.260

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

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

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 ?

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.

P	Q	$\widehat{\sigma}^2_\varepsilon$	$\log V$
0	0	11.984368	−1376.526
0	1	6.558542	−1232.809
0	2	6.567932	−1231.538
0	3	6.562049	−1231.499
1	0	9.432254	−1311.159
1	1	7.097006	−1237.768
2	0	7.999000	−1272.788
2	1	7.091900	−1237.582

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 ?

Comparer ces résultats avec ceux obtenus si l'on ajuste la série différenciée Y_t à un $AR(12)$ 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 lorqu'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

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