Séries chronologiques : recueil d'exercices

R. S. Stoica

Université Lille 1
Laboratoire Paul Painlevé
59655 Villeneuve d'Ascq Cedex, France

radu.stoica@math.univ-lille1.fr

Septembre, 2009

1 Introduction

Les exercices de ce recueil ont servi pendant trois ans, comme support pour les heures des travaux dirigés et des travaux pratiques qui complétaient le cours “Prévisions dans les séries chronologiques”, donné aux étudiants en dernière année de master pro par Marie-Claude Viano, puis par Thomas Simon. Marie-Claude Viano a beaucoup amélioré la qualité des énoncés du point de vue de la rigueur mathématique, de la logique et de l’orthographe. Je tiens la remercier pour tous ses efforts.

Je serai également reconnaissant au lecteur qui aura bien la gentillesse de me signaler les eventuelles fautes trouvées dans le texte. Ceci, malgré mon orgueil un peu touché, ne fera qu’améliorer la qualité de ce document.

Ce recueil regroupe des exercices qui se trouvent majoritairement dans [1, 3, 2]. Le choix de ces exercices a été fait en fonction des objectifs du cours : familiariser l’étudiant avec les séries chronologiques, détecter et travailler avec les tendances d’une série temporelle, manipuler les modèles de type ARMA, utiliser la fonction de corrélation et de corrélation partielle en vue de la modélisation et finalement, effectuer et évaluer une prédiction qui s’appuie sur un modèle attaché à la série observée.

Le logiciel utilisé pendant les séances des travaux pratiques a été $\mathbb{R}$, disponible à partir du site `http://www.r-project.org/`. Les données réelles sur lesquelles nous avons travaillé sont accessibles à partir du site `http://www.statsci.org/`.

Pour terminer, je voudrais remercier également tous les collègues qui mettent en libre accès sur leurs pages Web leurs ressources pédagogiques.

2 Exercices pour les travaux dirigés

Exercice 1. Soit X et Y deux variables aléatoires avec $\mathbb{E}[Y] = \mu$ et $\mathbb{E}[Y^2] < \infty$.
a) Montrez que la constante $c = \mu$ minimise $\mathbb{E}[(Y - c)^2]$.
b) Déduisez-en que la variable aléatoire $f(X)$ qui minimise $\mathbb{E}[(Y - f(X))^2 | X]$ est $f(X) = \mathbb{E}[Y|X]$.
c) Prouvez que la variable aléatoire qui minimise $\mathbb{E}[(Y - f(X))^2]$ est aussi $f(X) = \mathbb{E}[Y|X]$.

Exercice 2. Soit la série définie par $X_n = Z_n + \theta Z_{n-2}$, avec n entier, $\{Z_n\} \sim BB(0, \sigma^2)$ et θ une constante réelle positive.
a) On note $\rho_X(h) = Cov(X_n X_{n-h})$. Calculez $\rho_X(h)$.
b) Calculez la variance de $(X_1 + X_2 + X_3 + X_4)/4$ pour $\theta = 0.8$ et $\sigma^2 = 1$. Refaites ce calcul pour $\theta = -0.8$.

Exercice 3. Soit la série définie par $X_n = \phi X_{n-2} + Z_n$, avec n entier, $\{Z_n\} \sim BB(0, \sigma^2)$ et $|\phi| < 1$. $\{X_n\}$ est stationnaire. $\{Z_n\}$ et $\{X_m\}$ ne sont pas corrélées pour $m < n$.
a) Calculez $\rho_X(h)$.
b) Calculez la variance de $(X_1 + X_2 + X_3 + X_4)/4$ pour $\phi = 0.9$ et $\sigma^2 = 1$. Refaites ce calcul pour $\phi = -0.9$.

Exercice 4. Sur l'ensemble des suites, on définit l'opérateur différence première ∇ par $(\nabla x)_n = x_n - x_{n-1}$. Si $x_n = \sum_{k=0}^{p} c_k n^k$ avec $n \in \mathbb{V}$, montrez que $(\nabla x)_n$ est un polynôme de degré $p - 1$ et que $(\nabla^{p+1} x)_n = 0$.

Exercice 5. Soient $\{x_1, x_2, \ldots, x_n\}$ les valeurs observées d'une série temporelle et $\hat{\rho}_n(h)$ la fonction d'auto-corrélation empirique. Prouvez que :
a) Si $x_t = a + bt$, où a et $b \neq 0$ sont des constantes, montrez que pour h fixé

$$\lim_{n \to \infty} \hat{\rho}_n(h) = 1$$

b) Si $x_t = c \cos(\omega t)$ où $c \neq 0$ et $\omega \in (-\pi, \pi]$ sont des constantes, montrez que pour h fixée

$$\lim_{n \to \infty} \hat{\rho}_n(h) = cos(\omega h)$$

Rappels peut être utiles ...
• $\mathbb{E}[Y] = \mathbb{E}\left[\mathbb{E}[Y|X]\right]$
• $Cov(aX + bY, cU + dV) = acCov(X, U) + adCov(X, V) + bcCov(Y, U) + bdCov(Y, V)$
• $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$

Exercice 6. L'opération qui à la suite u_n associe la suite $v_n = \sum_{j=-q}^{q} a_j u_{n-j}$ s'appelle filtrage. Soit le filtre (moyenne mobile) donné par les coefficients $a_j = (2q+1)^{-1}$, $-q \leq j \leq q$.
a) Si $m_n = c_0 + c_1 n$, montrez que $\sum_{j=-q}^{q} a_j m_{n-j} = m_n$.
b) Si Z_n , $n = 0, \pm 1, \pm 2, \ldots$ sont des v.a. indépendantes de moyenne nulle et variance σ^2, montrez que la moyenne mobile $A_n = \sum_{j=-q}^{q} a_j Z_{n-j}$ est “petite” pour q grand, c'est à dire $\mathbb{E}[A_n] = 0$ et $Var[A_n] = \frac{\sigma^2}{2q+1}$.

Exercice 7. A et B sont des variables aléatoires non-corrélées et ω est une fréquence de valeur fixée dans l'intervalle $[0, \pi]$. Montrez que le processus $X_t = A \cos \omega t + B \sin \omega t$ est stationnaire et calculez son espérance et sa fonction d'auto-covariance.

Exercice 8.
a) Calculez la fonction d'auto-covariance de la série temporelle $X_t = Z_t + 0.3Z_{t-1} - 0.4Z_{t-2}$ avec $\{Z_t\} \sim BB(0, 1)$.
b) Calculez la fonction d'auto-covariance de la série temporelle $Y_t = W_t - 1.2W_{t-1} - 1.6W_{t-2}$ avec $\{W_t\} \sim BB(0, 0.25)$.
c) Comparez les résultats obtenus.

Exercice 9. Trouvez des séries temporelles ayant comme fonction d'auto-covariance les fonctions suivantes :
a) $\sigma(h) = 1, \quad h = 0, \pm 1, \pm 2, \ldots$
b) $\sigma(h) = (-1)^{|h|}, \quad h = 0, \pm 1, \pm 2, \ldots$
c)$\sigma(h) = 1 + \cos\left(\frac{\pi h}{2}\right) + \cos\left(\frac{\pi h}{4}\right), \quad h = 0, \pm 1, \pm 2, \ldots$
d)

$$\sigma(h) = \begin{cases} 1 & \text{si } h = 0 \\ 0.4 & \text{si } h = \pm 1 \\ 0 & \text{sinon} \end{cases}$$

Exercice 10. Soit $X_1, X_2, \ldots$ une série stationnaire de moyenne μ et fonction d'auto-corrélation $\rho(\cdot)$. Montrez que le meilleur prédicteur de X_{n+h} de la forme $aX_n + b$ est obtenu pour les valeurs $a = \rho(h)$ et $b = \mu(1 - \rho(h))$.

Exercice 11. Montrez que l'équation auto-régressive

$$X_t = \phi X_{t-1} + Z_t, \qquad t = 0, \pm 1, \pm 2, \ldots$$

avec $Z_t \sim BB(0, \sigma^2)$ et $|\phi| = 1$ n'a pas de solution stationnaire.

Exercice 12. Montrez que

$$\frac{1}{1 - \phi z} = -\sum_{j=1}^{\infty} \phi^{-j} z^{-j}$$

pour $|\phi| > 1$ et $|z| > 1$.

Exercice 13. Soit $\{Y_t\}$ la série temporelle formée par l'addition d'un modèle $AR(1)$ et d'un bruit

$$Y_t = X_t + W_t$$

où $X_t = \phi X_{t-1} + Z_t$, $\{Z_t\} \sim BB(0, \sigma_z^2)$, $|\phi| < 1$ et X_s et Z_t non-corrélées pour $s < t$. Nous avons également $\{W_t\} \sim BB(0, \sigma_w^2)$ et $\mathbb{E}[W_s Z_t] = 0$ pour tous les s, t.
a) Montrez que $\{Y_t\}$ est stationnaire et calculez sa fonction de covariance.
b) Montrez que la série $U_t = Y_t - \phi Y_{t-1}$ peut être répresentée par un modèle $MA(1)$.
c) Est-ce que d'après le point précédent nous pouvons dire que $\{Y_t\}$ est un modèle $ARMA(1,1)$? Dans le cas affirmatif, trouvez ses paramètres en sachant $\phi = 0.5$, $\sigma_w^2 = 0.25$ et $\sigma_z^2 = 1$.

Exercice 14. Est-ce que les modèles ARMA considérés sont bien définis ?
a) $X_t + 0.2X_{t-1} - 0.48X_{t-2} = Z_t$
b) $X_t + 1.9X_{t-1} + 0.88X_{t-2} = Z_t + 0.2Z_{t-1} + 0.7Z_{t-2}$
c) $X_t + 0.6X_{t-1} = Z_t + 1.2Z_{t-1}$
d) $X_t + 1.8X_{t-1} + 0.81X_{t-2} = Z_t$
e) $X_t + 1.6X_{t-1} = Z_t - 0.4Z_{t-1} + 0.04Z_{t-2}$
Ici $\{Z_t\} \sim BB(0, \sigma^2)$.

Exercice 15. Calculez la fonction de corrélation de la série temporelle suivante :

$$X_t + 0.2X_{t-1} - 0.48X_{t-2} = Z_t$$

(contentez vous de calculer les premières valeurs)

Exercice 16. Calculez la fonction de corrélation du processus

$$X_t = 0.8X_{t-3} + Z_t$$

avec $\{Z_t\} \sim BB(0, \sigma^2)$.

Exercice 17. Montrez que pour $h = 2$ la valeur du coéfficient d'autocorrélation partielle $\alpha(h)$ du modèle $MA(1)$

$$X_t = Z_t + \theta Z_{t-1}, \quad t = 0, \pm 1, \pm 2, \ldots$$

avec $\{Z_t\} \sim BB(0, \sigma^2)$ est

$$\alpha(2) = -\frac{\theta^2}{1 + \theta^2 + \theta^4}.$$

Exercice 18. Soit $\{X_t\}$, $t = 0, \pm 1, \pm 2, \ldots$ une série stationnaire définie par

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t$$

avec $\{Z_t\} \sim BB(0, \sigma^2)$ et $\mathbb{E}[Z_t X_s] = 0$ pour chaque $s < t$.
a) Montrez que le meilleur prédicteur linéaire de X_{n+1} en fonction de $1, X_1, \ldots, X_n$ est

$$\widehat{X}_{n+1} = \phi_1 X_n + \ldots + \phi_p X_{n+1-p}$$

(on suppose $n > p$).
b) Quelle est l'erreur quadratique moyenne de $\widehat{X}_{n+1}$?

Exercice 19. Soit le modèle $AR(2)$

$$X_n - aX_{n-1} - a^2 X_{n-2} = \epsilon_n, \quad \forall n$$

avec $\{\epsilon_n\} \sim BB(0, \sigma^2)$.
a) Donner une condition suffisante sur a pour qu'il existe un processus stationnaire centré (X_n).
b) Donner les expressions $\sigma(0), \rho(1)$ et $\rho(2)$ en fonction de deux paramètres a et σ^2, et en déduire une expression simple de a en fonction des autocorrélations ...
c) Préciser les valeurs, en fonction de a, des coefficients $r(h)$, $(h = 1, 2, \ldots)$.

3 Exercices pour les travaux pratiques

Exercice 1. Simuler une série temporelle formée de variables normales i.i.d. de paramètres :
a) $\mu = 0$ et $\sigma^2 = 1$,
b) $\mu = 0$ et $\sigma^2 = 4$;
c) Regardez l'auto-corrélation et la covariance des simulations obtenues ; interprétez.
Indication : help(plot), help(rnorm), help(acf), windows() or x11() et aussi par(mfrow=c(2,1))

Exercice 2. Fabriquer des données simulées :
a) Dessinez la droite $m_1 = at + b$ et la sinusoïde $s_{50} = 250 \sin(\frac{2\pi t}{T})$ avec $a = 0.5$, $b = 1.0$ et $T = 50$. On prendra $t \in [0, 999]$.
b) Rajoutez un bruit $\mathcal{N}(0, 10)$ à la droite et à la sinusoïde, respectivement ; regardez les fonction d'auto-corrélation respectives en faisant varier l'écart type du bruit ; interprétez.
c) Rajoutez la droite, la sinusoïde, puis le bruit ; regardez les fonctions d'auto-corrélation respectives ; interprétez.
Indication : help(sin)

Exercice 3. Test du signe pour la vérification des résidus.
Pour la série du bruit $\{\epsilon_n\}$, soit S le nombre de valeurs de i telles que $\epsilon_i > \epsilon_{i-1}$ avec $i = 2, \ldots, n$. Il est possible de montrer que pour une séquence i.i.d. nous avons $\mathbb{E}[S] = \frac{n-1}{2}$et $Var[S] = \frac{n+1}{12}$. Quand n prend des valeurs suffisaient grandes, nous pouvons considérer que S suit une loi normale $\mathcal{N}(\mu_S, \sigma_S^2)$ avec $\mu_S = \mathbb{E}[S]$ et $\sigma_S^2 = Var[S]$.

a) Dans ce contexte, construisez un test statistique qui vérifie si les résidus présentent ou non une tendance. L'hypothèse d'absence de tendance dans les résidus est rejetée si $|S - \mu_S|/\sigma_S > \phi_{1-\alpha/2}$. α est le niveau de confiance du test.
b) Est-t-il possible de tromper ce test ?

Exercice 4. Enlever la tendance :
L'opérateur différence d'ordre p est défini par $(\nabla_p x)_n = x_n - x_{n-p}$. Appliquer cet opérateur pour plusieurs valeurs de p afin d'éliminer la tendance

et/ou la saisonalité des séries précédentes. Analysez l'auto-corrélation des résultats obtenus.
Indication : help(diff)

Exercice 5. Lire la série qui se trouve dans le fichier redwine.dat et analysez la à la lumière de ce que vous avez vu dans les applications précédentes.

Indication : help(read.table)
Pour lire les données vous pouvez utiliser :

rw=read.table("redwine.dat",header=FALSE)
data=rw[,1]

Exercice 6. Soit $\{Z_t\} \sim \mathcal{N}(0, 100^2)$ et les séries temporelles suivantes définies comme suit :

$$Y_t^{(1)} = 1 + 0.5t + Z_t, \tag{1}$$

$$Y_t^{(2)} = 250 \sin\left(\frac{2\pi t}{50}\right) + Z_t, \tag{2}$$

et

$$Y_t^{(3)} = 1 + 0.5t + 250 \sin\left(\frac{2\pi t}{50}\right) + Z_t. \tag{3}$$

a) Simulez les trois séries pour $t \in [0, 999]$.
b) Effectuez un filtrage (moyenne mobile) des trois séries avec et sans bruit. Les coéfficients du filtre sont données par le code R suivant :

```
d=50
a=c(0.5,rep(1,(d-1)),0.5)/d
```

Pour filtrer une série temporelle Y_t, la commande R *filter* peut être utilisée comme suit :

```
y.f=filter(y,a,circular=TRUE)
```

c) Comparez les résultats de vos simulations avec les résultats théoriques obtenus pendant le TD.

Exercice 7.
a) Simuler la série $X_t = A\cos(\omega t) + B\sin(\omega t)$ avec $t = 0, 1, \ldots, 199$, $\omega = 2\pi/10$ et A et B deux variables aléatoires indépendantes de moyenne nulle et variance unité ;
b) Analyser la fonction de covariance empirique pour plusieurs réalisations de la série.

Exercice 8. Simuler les modèles MA suivants et regarder leur fonction de covariance empirique :
a) $X_t = Z_t + 0.3Z_{t-1} - 0.4Z_{t-2}$ avec $\{Z_t\} \sim BB(0, 1)$,
b) $Y_t = W_t - 1.2W_{t-1} - 1.6W_{t-2}$ avec $\{W_t\} \sim BB(0, 0.25)$.
c) Comparer les résultats obtenus avec ceux de TD.

Exercice 9. Tracer les fonctions de covariance théorique suivantes et faire la comparaison avec les fonction de covariance empirique obtenues en simulant les séries temporelles associées (numériquement et graphiquement) :
a) $\sigma(h) = 1$ avec $h = 0, \pm1, \pm2, \ldots,$
b) $\sigma(h) = (-1)^h$ avec $h = 0, \pm1, \pm2, \ldots,$
c) $\sigma(h) = 1 + \cos\frac{\pi h}{2} + \cos\frac{\pi h}{4},$
d)

$$\sigma(h) = \begin{cases} 1 & \text{si } h = 0 \\ 0.4 & \text{si } h = \pm 1 \\ 0 & \text{sinon} \end{cases}$$

Exercice 10. Simuler la trajectoire d'un modèle AR(1) : $X_t = \phi X_{t-1} + Z_t$ avec $\{Z_t\} \sim BB(0, 1)$.
a) $\phi = 0.9$
b) $\phi = -0.9$
c) $\phi = 1.001$
d) $\phi = -1.001$
Regarder les fonctions de covariance et les réalisations du modèle en fonction de différentes conditions initiales. Comparer aussi avec les valeurs théoriques et avec le résultat de l'*Exercice 11* fait en TD.

Exercice 11. Soit les modèles ARMA suivants :

a) $X_t + 0.2X_{t-1} - 0.48X_{t-2} = Z_t$
b) $X_t = Z_t + 0.3Z_{t-1} - 0.4Z_{t-2}$
c) $X_t - 0.5X_{t-1} = Z_t + 0.4Z_{t-1}$
avec $\{Z_t\} \sim BB(0,1)$. Pour chacun de ces modèles vous devez faire :
- simuler le modèle
- calculer/afficher sa fonction de corrélation
- calculer/afficher sa fonction de corrélation partielle
- refaire ces calculs/affichages pour la réalisation du modèle
- comparer les résultats théoriques avec les résultats obtenus par simulation en faisant évoluer le nombre de simulations. Qu'est-ce que vous observez de particulier par rapport à la "forme" de ces fonctions pour chacun de ces modèles ? Est-ce que l'on peut mettre en relation les paramètres des modèles et les courbes observées ?

Exercice 12. Soit le modèle $Y_t = X_t + W_t$ avec X_t un modèle AR(1) et $\{W_t\} \sim BB(0, 0.25)$. Le modèle AR(1) est donné par $X_t = 0.5X_{t-1} + Z_t$ avec $\{Z_t\} \sim BB(0,1)$.
a) Étudier la fonction de corrélation et la fonction de corrélation partielle de X_t ;
b) Simuler Y_t. Analysez sa fonction de corrélation et la fonction de corrélation partielle ;
c) Simuler $U_t = Y_t - 0.5Y_{t-1}$ par deux méthodes et analyser sa fonction de corrélation et sa fonction de corrélation partielle. La première méthode est directe. La deuxième nécessite l'identification d'un modèle MA(1) ;
d) Simuler Y_t en tant que ARMA(1,1).

Exercice 13.
a) Étudier la fonction de corrélation et la corrélation partielle du modèle $X_t = 0.8X_{t-3} + Z_t$, avec $\{Z_t\} \sim BB(0,1)$.
b) Estimer les paramètres de ce processus sur une simulation de longueur n. Regarder les résultats obtenus en fonction de n. Les résultats obtenus sont les coefficients du modèle et les résidus ...
c) Prédire les 10 premières valeurs à partir de n en utilisant le modèle obtenu au point b). Tracer les intervalles de confiance. Étudier les erreurs de prédiction. Comparer avec les données ... A-t-on intérêt de prédire avec un horizon grand ?

Exercice 14. Soit le modèle ARMA(2,3) défini par $X_t - X_{t-1} + 0.24X_{t-2} = Z_t + 0.4Z_{t-1} + 0.2Z_{t-2} + 0.1Z_{t-3}$ avec $\{Z_t\} \sim BB(0,1)$.
a) Étudier sa fonction de corrélation et sa fonction de corrélation partielle.
b) Estimer les paramètres du processus utilisant une simulation de longueur n du modèle.
c) Prédire les 10 premières valeurs à partir de n en utilisant le modèle obtenu au point b) et analyser les résultats.

Exercice 15.
a) Lire la série "oshorts.dat". Prédire les 7 premières valeurs à partir de $n = 50$.
b) Lire la série "sunspots.dat". Prédire les 10 premières valeurs à partir de $n = 100$.
Dans les deux cas il faut justifier le choix du modèle et analyser les résultats.

Exercice 16. Soit la série $Y_t = 0.5t + 1 + \epsilon_t$ avec $t = 0, 1, \ldots, 200$ et $\epsilon_t \sim BB(0, 100)$.
a) Analyser et modéliser la série obtenue.
b) Effectuer la prédiction des premières $n = 5, 10, \ldots, 100$ valeurs.
c) Calculer et afficher les intervalles de confiance.

Exercice 17. Soit la série $Y_t = 250\sin(\frac{2\pi t}{5}) + \epsilon_t$ avec $t = 0, 1, \ldots, 200$ et $\epsilon_t \sim BB(0, 100)$. Mêmes questions qu'à l'exercice précédent.

Exercice 18. Soit la série $Y_t = 10\sin(\frac{2\pi t}{5}) + 0.5t + 1 + \epsilon_t$ avec $t = 0, 1, \ldots, 200$ et $\epsilon_t \sim BB(0, 4)$. Mêmes questions qu'à l'exercice précédent.

Exercice 19. Lire la série "airpass.dat". Mêmes questions qu'à l'exercice précédent.

Annexe A: idées pour les solutions aux exercices des travaux pratiques

```
#
# Exercice 1
#
x11()
par(mfrow=c(2,1))
eps1=rnorm(1000,0,1)
plot(eps1,type="l")
acf(eps1)

x11()
par(mfrow=c(3,1))
eps2=rnorm(1000,1,2)
plot(eps2,type="l")
acf(eps2,lag=50)
acf(eps2,lag=50,type="covariance")

#
# Exercice 2
#
t=c(0:999)
eps=rnorm(length(t),0,100)

a=0.5
b=1.0
m1=a*t+b
y1=m1+eps

x11()
par(mfrow=c(3,1))
plot(t,m1,type="l")
plot(t,y1,type="l")
acf(y1,lag=300)


c=0.1
s50=250*sin(2*pi*t/50)
```

```
y2=s50+eps

x11()
par(mfrow=c(3,1))
plot(t,s50,type="l")
plot(t,y2,type="l")
acf(y2,lag=300)


m1s50=m1+s50
y3=m1s50+eps

x11()
par(mfrow=c(3,1))
plot(t,m1s50,type="l")
plot(t,y3,type="l")
acf(y3,lag=300)

#
# Exercice 3
#
sign.test = function(x,alpha=0.05)
{
  # taille echantillon
  n = length(x)

  # parametres loi
  ms = (n-1)/2
  sds = sqrt((n+1)/12)

  # construction test
  S = sum(x[2:n]>x[1:(n-1)])
  S1 = (S - ms)/sds

  # interval de confiance ou region de rejet1
  I1 = sds*qnorm(alpha/2)+ms
  I2 = sds*qnorm(1-(alpha/2))+ms

  # retourner la p-value
  list(p.value=pnorm(S1),int1 = I1,int2 = I2)
```

```
}

#
# Exercice 4
#
# travail sur le bruit
res=diff(eps,25)
x11()
par(mfrow=c(2,1))
plot(res,type="l")
acf(res,lag=300)

# travail sur y1
res=diff(y1,1)
x11()
par(mfrow=c(2,1))
plot(res,type="l")
acf(res,lag=300)

# travail sur y2
res=diff(y2,50)
x11()
par(mfrow=c(2,1))
plot(res,type="l")
acf(res,lag=300)

# travail sur y3
res1=diff(y3,50)
res=diff(res1,1)
x11()
par(mfrow=c(2,1))
plot(res,type="l")
acf(res,lag=300)

#
# Exercice 5
#
rw=read.table("../../../DATA/redwine.dat",header=FALSE)
data=rw[,1]
```

```
x11()
par(mfrow=c(2,1))
plot(data,type="l")
acf(data,lag=100)

res=diff(data,12)
x11()
par(mfrow=c(2,1))
plot(res,type="l")
acf(res,lag=300)

#
# Exercice 6
#
# produire les donnee simulees
t=c(0:999)
eps=rnorm(length(t),0,100)

a=0.5
b=1.0
m1=a*t+b
y1=m1+eps

s50=250*sin(2*pi*t/50)
y2=s50+eps

m1s50=m1+s50
y3=m1s50+eps

x11()
par(mfrow=c(3,3))
plot(t,m1,type="l")
plot(t,y1,type="l")
acf(y1,lag=300)

plot(t,s50,type="l")
plot(t,y2,type="l")
acf(y2,lag=300)
```

```
plot(t,m1s50,type="l")
plot(t,y3,type="l")
acf(y3,lag=300)

# construire le filtre ...
# si d est pair
d=50
a=c(0.5,rep(1,(d-1)),0.5)/d

# si d est impair filtrage par moyenne mobile
#d=51
#a=rep(1/d,d)

# eliminer la tendance ...
# travail sur bruit
eps.f=filter(eps,a,circular=TRUE)

x11()
par(mfrow=c(2,3))
plot(t,eps,type="l")
acf(eps,lag=300,type="covariance")
acf(eps,lag=300)
plot(t,eps.f,type="l")
acf(eps.f,lag=300,type="covariance")
acf(eps.f,lag=300)

# travail sur m1 et y1
m1.f=filter(m1,a,circular=TRUE)
y1.f=filter(y1,a,circular=TRUE)

x11()
par(mfrow=c(4,3))

plot(m1,type="l")
acf(m1,lag=300,type="covariance")
acf(m1,lag=300)
plot(m1.f,type="l")
acf(m1.f,lag=300,type="covariance")
acf(m1.f,lag=300)
```

```
plot(y1,type="l")
acf(y1,lag=300,type="covariance")
acf(y1,lag=300)
plot(y1.f,type="l")
acf(y1.f,lag=300,type="covariance")
acf(y1.f,lag=300)

# travail sur s50 et y2
s50.f=filter(s50,a,circular=TRUE)
y2.f=filter(y2,a,circular=TRUE)

x11()
par(mfrow=c(4,3))

plot(s50,type="l")
acf(s50,lag=300,type="covariance")
acf(s50,lag=300)
plot(s50.f,type="l")
acf(s50.f,lag=300,type="covariance")
acf(s50.f,lag=300)

plot(y2,type="l")
acf(y2,lag=300,type="covariance")
acf(y2,lag=300)
plot(y2.f,type="l")
acf(y2.f,lag=300,type="covariance")
acf(y2.f,lag=300)

# travail sur m1s50 et y3
m1s50.f=filter(m1s50,a,circular=TRUE)
y3.f=filter(y3,a,circular=TRUE)

x11()
par(mfrow=c(4,3))

plot(m1s50,type="l")
acf(m1s50,lag=300,type="covariance")
acf(m1s50,lag=300)
plot(m1s50.f,type="l")
acf(m1s50.f,lag=300,type="covariance")
```

```
acf(m1s50.f,lag=300)

plot(y3,type="l")
acf(y3,lag=300,type="covariance")
acf(y3,lag=300)
plot(y3.f,type="l")
acf(y3.f,lag=300,type="covariance")
acf(y3.f,lag=300)

#
# Exercice 7
#
t=c(0:199)
a=rnorm(1,0,1)
b=rnorm(1,0,1)
x=a*cos((2*pi/10)*t)+b*sin((2*pi/10)*t)
x11()
par(mfrow=c(2,1))
plot(x,type="l")
acf(x,lag=200,type="covariance")

#
# Exercice 8
#
z=rnorm(1000,0,1)
x=filter(z,filter=c(1,0.3,-0.4),method="convolution",circular=TRUE)
x11()
par(mfrow=c(3,1))
plot(x,type="l")
sigma_x=acf(x,lag=10,type="covariance")
sigma_x[0:3]
rho_x=acf(x,lag=10)
rho_x[0:3]

w=rnorm(1000,0,0.5)
y=filter(w,filter=c(1,-1.2,-1.6),method="convolution",circular=TRUE)
x11()
par(mfrow=c(3,1))
plot(y,type="l")
sigma_y=acf(y,lag=10,type="covariance")
```

```
sigma_y[0:3]
rho_y=acf(y,lag=10)
rho_y[0:3]

#
# Exercice 9
#
ha=c(0:99)
sha=rep(1,length(ha))
xa=rnorm(1,0,1)*rep(1,1000)
x11()
par(mfrow=c(2,1))
acf(xa,lag=100,type="covariance")
plot(ha,sha,col="blue")

hb=c(0:99)
shb=rep(c(1,-1),length(hb)/2)
xb=rnorm(1,0,1)*rep(c(1,-1),500)
x11()
par(mfrow=c(2,1))
acf(xb,lag=100,type="covariance")
plot(hb,shb,col="blue")

hc=c(0:199)
shc=1+cos(0.5*pi*hc)+cos(0.25*pi*hc)
a=rnorm(1,0,1)
b=rnorm(1,0,1)
c=rnorm(1,0,1)
z=rnorm(1,0,1)
t=c(0:999)
xc=z*rep(1,1000)+a*cos((0.5*pi)*t)+b*sin((0.5*pi)*t)
+c*cos((0.25*pi)*t)+d*sin((0.25*pi)*t)
x11()
par(mfrow=c(2,1))
acf(xc,lag=200,type="covariance")
plot(hc,shc,col="blue",type="l")

hd=c(0:9)
shd=c(c(1,0.4),rep(0,8))
theta1=2
```

```
zd1=rnorm(1000,0,sqrt(0.2))
yd1=filter(zd1,filter=c(1,theta1),method="convolution",circular=TRUE)
theta2=0.5
zd2=rnorm(1000,0,sqrt(0.8))
yd2=filter(zd2,filter=c(1,theta2),method="convolution",circular=TRUE)
par(mfrow=c(3,1))
acf(yd1,lag=10,type="covariance")
acf(yd2,lag=10,type="covariance")
plot(hd,shd,col="blue")

#
# Exercice 10
#
M=1000
phi=0.9
z=rnorm(M,0,1)
x0=0.2
x=c(x0,rep(0,M-1))

for(i in 2 : M)
  {
    x[i]=phi*x[i-1]+z[i]
  }


x11()
par(mfrow=c(2,1))
plot(x,col="blue",type="l")
acf(x,lag=20,type="covariance")

#
# Exercice 11
#
acf=ARMAacf(ar=c(-0.2,0.48),ma=0,25)
pacf=ARMAacf(ar=c(-0.2,0.48),ma=0,25,pacf=T)
ar2=arima.sim(n=100,list(order=c(2,0,0,),ar=c(-0.2,0.48)),sd=1)
x11()
par(mfrow=c(4,1))
plot(acf,type="h",xlab="Lag")
abline(h=0)
```

```
plot(pacf,type="h",xlab="Lag")
abline(h=0)
acf(ar2,lag=25)
pacf(ar2,lag=25)

acf=ARMAacf(ar=0,ma=c(0.3,-0.4),25)
pacf=ARMAacf(ar=0,ma=c(0.3,-0.4),25,pacf=T)
ma2=arima.sim(n=1000,list(order=c(0,0,2,),ma=c(0.3,-0.4)),sd=1)
x11()
par(mfrow=c(4,1))
plot(acf,type="h",xlab="Lag")
abline(h=0)
plot(pacf,type="h",xlab="Lag")
abline(h=0)
acf(ma2,lag=25)
pacf(ma2,lag=25)

acf=ARMAacf(ar=0.5,ma=0.4,25)
pacf=ARMAacf(ar=0.5,ma=0.4,25,pacf=T)
arma11=arima.sim(n=200,list(order=c(1,0,1),ar=0.5,ma=0.4),sd=1)
x11()
par(mfrow=c(4,1))
plot(acf,type="h",xlab="Lag")
abline(h=0)
plot(pacf,type="h",xlab="Lag")
abline(h=0)
acf(arma11,lag=25)
pacf(arma11,lag=25)

#
# Exercice 12
#
acf=ARMAacf(ar=0.5,ma=0,25)
pacf=ARMAacf(ar=0.5,ma=0,25,pacf=T)
x11()
par(mfrow=c(3,1))
plot(acf,type="h",xlab="Lag")
abline(h=0)
plot(pacf,type="h",xlab="Lag")
abline(h=0)
```

```
x=arima.sim(n=1000,list(order=c(1,0,0),ar=0.5),sd=1)
w=rnorm(1000,0,1)
y=x+w
x11()
par(mfrow=c(3,1))
plot(y,type="l",col="blue")
acf(y,lag=25)
pacf(y,lag=25)

# differences ...
y2=y[2:1000]
y1=y[1:999]
u1=y2-y1
x11()
par(mfrow=c(3,1))
plot(u1,type="l",col="blue")
acf(u1,lag=25)
pacf(u1,lag=25)

# trouver le MA(1) : utiliser les tds et les tps ...
# theta_u = 0.1 et sigma2_u = 1.25
acf=ARMAacf(ar=0,ma=-0.1,25)
pacf=ARMAacf(ar=0,ma=-0.1,25,pacf=T)
ma1=arima.sim(n=1000,list(order=c(0,0,1),ma=-0.1),sd=sqrt(1.25))
x11()
par(mfrow=c(2,2))
plot(acf,type="h",xlab="Lag")
abline(h=0)
plot(pacf,type="h",xlab="Lag")
abline(h=0)
acf(ma1,lag=25)
pacf(ma1,lag=25)

acf=ARMAacf(ar=0.5,ma=-0.1,25)
pacf=ARMAacf(ar=0.5,ma=-0.1,25,pacf=T)
y3=arima.sim(n=1000,list(order=c(1,0,1),ar=0.5,ma=-0.1),sd=sqrt(1.25))
x11()
par(mfrow=c(2,2))
plot(acf,type="h",xlab="Lag")
```

```
abline(h=0)
plot(pacf,type="h",xlab="Lag")
abline(h=0)
acf(y3,lag=25)
pacf(y3,lag=25)


#
# Exercice 13
#
acf=ARMAacf(ar=c(0,0,0.8),ma=0,25)
pacf=ARMAacf(ar=c(0,0,0.8),ma=0,25,pacf=T)
ar3=arima.sim(n=1000,list(order=c(3,0,0),ar=c(0.0,0.0,0.8)),sd=1)
x11()
par(mfrow=c(2,2))
plot(acf,type="h",xlab="Lag")
abline(h=0)
plot(pacf,type="h",xlab="Lag")
abline(h=0)
acf(ar3,lag=25)
pacf(ar3,lag=25)

y=ar3[1:100]
estim=arima(y,order=c(3,0,0))
estim$coef
x11()
par(mfrow=c(2,1))
plot(estim$resid,type="l")
acf(estim$resid,lag=50)

p=predict(estim,n.ahead=50)
x11()
plot(y,type="o",col="green",xlim=c(0,120))
lines(p$pred,type="o",col="red")
lines(p$pred+1.96*p$se,col="blue",lty="dashed")
lines(p$pred-1.96*p$se,col="blue",lty="dashed")

#
# Exercice 14
#
```

```
acf=ARMAacf(ar=c(1,-0.24),ma=c(0.4,0.2,0.1),25)
pacf=ARMAacf(ar=c(1,-0.24),ma=c(0.4,0.2,0.1),25,pacf=T)
arma23=arima.sim(n=1000,list(order=c(2,0,3),ar=c(1,-0.24),
                 ma=c(0.4,0.2,0.1)),sd=1)
x11()
par(mfrow=c(2,2))
plot(acf,type="h",xlab="Lag")
abline(h=0)
plot(pacf,type="h",xlab="Lag")
abline(h=0)
acf(arma23,lag=25)
pacf(arma23,lag=25)

y=arma23[1:940]
estim=arima(y[1:900],order=c(2,0,3))
estim$coef
x11()
par(mfrow=c(2,1))
plot(estim$resid,type="l")
acf(estim$resid,lag=50)

p=predict(estim,n.ahead=20)
x11()
plot(y,type="o",col="green",xlim=c(880,930))
lines(p$pred,type="o",col="red")
lines(p$pred+1.96*p$se,col="blue",lty="dashed")
lines(p$pred-1.96*p$se,col="blue",lty="dashed")

#
# Exercice 15
#
# oshorts.dat
rw=read.table("../DATA TP/oshorts.dat",header=FALSE)
data=rw[,1]

y=data[1:50]

x11()
par(mfrow=c(3,1))
plot(y,type="o")
```

```
abline(h=0)
acf(y,lag=25)
pacf(y,lag=25)

estim=arima(y,order=c(0,0,1))
estim$coef
x11()
par(mfrow=c(2,1))
plot(estim$resid,type="l")
acf(estim$resid,lag=50)

p=predict(estim,n.ahead=7)
x11()
plot(data,type="o",col="green",xlim=c(1,60))
lines(p$pred,type="o",col="red")
lines(p$pred+1.96*p$se,col="blue",lty="dashed")
lines(p$pred-1.96*p$se,col="blue",lty="dashed")


# sunspots.dat
rw=read.table("../DATA TP/sunspots.dat",header=FALSE)
data=rw[,1]

data=data-mean(data)

y=data[1:100]

x11()
par(mfrow=c(3,1))
plot(y,type="o")
abline(h=0)
acf(y,lag=25)
pacf(y,lag=25)

estim=arima(y,order=c(2,0,0))
estim$coef
x11()
par(mfrow=c(2,1))
plot(estim$resid,type="l")
acf(estim$resid,lag=50)
```

```
p=predict(estim,n.ahead=8)
x11()
plot(data[1:120],type="o",col="green",xlim=c(1,120))
lines(p$pred,type="o",col="red")
lines(p$pred+1.96*p$se,col="blue",lty="dashed")
lines(p$pred-1.96*p$se,col="blue",lty="dashed")

#
# Exercice 16
#
t=c(0:200)
eps=rnorm(length(t),0,100)

a=0.5
b=1.0
m1=a*t+b
y1=m1+eps
x1=diff(y1,1)

x11()
par(mfrow=c(4,1))
plot(y1,type="l",col="blue")
plot(x1,type="l",col="blue")
acf(x1,lag=10)
pacf(x1,lag=10)

sarima(y1,0,1,1)
sarima.for(y1,100,0,1,1)

#
# Exercice 17
#
s5=250*sin(2*pi*t/5)
y2=s5+eps
x2=diff(y2,5)

x11()
par(mfrow=c(3,1))
plot(y2,type="l",col="blue")
```

```
acf(y2,lag=40)
pacf(y2,lag=40)

x11()
par(mfrow=c(3,1))
plot(x2,type="l",col="blue")
acf(x2,lag=40)
pacf(x2,lag=40)

sarima(y2,0,0,5,0,1,0,5)
sarima.for(y2,100,0,0,5,0,1,0,5)

#
# Exercice 18
#
eps=rnorm(length(t),0,4)
s5=10*sin(2*pi*t/5)
y3=m1+s5+eps

x3=diff(y3,1)
x31=diff(x3,5)

x11()
par(mfrow=c(3,1))
plot(y3,type="l",col="blue")
acf(y3,lag=40)
pacf(y3,lag=40)

x11()
par(mfrow=c(3,1))
plot(x3,type="l",col="blue")
acf(x3,lag=40)
pacf(x3,lag=40)

x11()
par(mfrow=c(3,1))
plot(x31,type="l",col="blue")
acf(x31,lag=40)
pacf(x31,lag=40)
```

```
sarima(y3,0,1,5,0,1,0,5)
sarima.for(y3,100,0,1,5,0,1,0,5)

#
# Exercice 19
#
rw=read.table("../../../DATA/airpass.dat",header=FALSE)
data=rw[,1]

x11()
par(mfrow=c(3,1))
plot(data,type="l",col="blue")
acf(data,lag=40)
pacf(data,lag=40)

y=log(data)

x11()
par(mfrow=c(3,1))
plot(y,type="l",col="blue")
acf(y,lag=40)
pacf(y,lag=40)

dy1=diff(y,1)
dy112=diff(dy1,12)

x11()
par(mfrow=c(3,1))
plot(dy1,type="l",col="blue")
acf(dy1,lag=40)
pacf(dy1,lag=40)

x11()
par(mfrow=c(3,1))
plot(dy112,type="l",col="blue")
acf(dy112,lag=40)
pacf(dy112,lag=40)

sarima(y,12,1,0,0,1,0,12)
sarima.for(y,100,12,1,0,0,1,0,12)
```

```
#
# R code by R.H. Shumway and D.S. Stoffer
#
sarima.for=function(xdata,nahead,p,d,q,P=0,D=0,Q=0,S=-1){
  data=as.ts(xdata)
  constant=1:length(data)
  xmean=matrix(1,length(data),1)
  if (d>0)
  fitit=arima(data, order=c(p,d,q), seasonal=list(order=c(P,D,Q), period=S),
             xreg=constant,include.mean=F)
  if (d<.00001)
  fitit=arima(data, order=c(p,d,q), seasonal=list(order=c(P,D,Q), period=S),
             xreg=xmean,include.mean=F)
  if (d+D>1)
  fitit=arima(data, order=c(p,d,q), seasonal=list(order=c(P,D,Q), period=S))
#
  if (d>0)
   nureg=(length(data)+1):(length(data)+nahead)
  if (d<.00001)
   nureg=matrix(1,length(nahead),1)
  if (d+D>1)
   nureg=NULL
  fore=predict(fitit, n.ahead=nahead, newxreg=nureg)

#
# graph:
#
  U = fore$pred + 2*fore$se
  L = fore$pred - 2*fore$se
  minx=min(data,L)
  maxx=max(data,U)
  x11()
  ts.plot(data,fore$pred,col=1:2, ylim=c(minx,maxx),
         ylab=deparse(substitute(xdata)))
  lines(fore$pred, col="red", type="p")
  lines(U, col="blue", lty="dashed")
  lines(L, col="blue", lty="dashed")
#
  return(fore)
```

```
}

#
# R code by R.H. Shumway and D.S. Stoffer
#
sarima=function(data,p,d,q,P=0,D=0,Q=0,S=-1){
  n=length(data)
  constant=1:n
  xmean=matrix(1,n,1)
  if (d>0)
  fitit=arima(data, order=c(p,d,q), seasonal=list(order=c(P,D,Q), period=S),
              xreg=constant,include.mean=F)
  if (d<.00001)
  fitit=arima(data, order=c(p,d,q), seasonal=list(order=c(P,D,Q), period=S),
              xreg=xmean,include.mean=F)
  if (d+D>1)
  fitit=arima(data, order=c(p,d,q), seasonal=list(order=c(P,D,Q), period=S))
  if (S < 0) goof=20 else goof=3*S

  x11()
  par(mfrow=c(2,1))
  plot(fitit$resid,type="l",main="Standardized residuals",ylab=" ")
  acf(fitit$resid,lag=50,main="ACF of residuals", ylab=" ")


  #tsdiag(fitit,gof.lag=goof)
  #k=length(fitit$coef)
  #BIC=log(fitit$sigma2)+(k*log(n)/n)
  #AICc=log(fitit$sigma2)+((n+k)/(n-k-2))
  #AIC=log(fitit$sigma2)+((n+2*k)/n)
  #list(fit=fitit, AIC=AIC, AICc=AICc, BIC=BIC)

  list(fit=fitit)
}
```

References

[1] P. J. Brockwell and R. A. Davis. *Introduction to time series and forecasting*. Springer-Verlag, New-York, 1999.

[2] G. Oppenheim, A. Philippe, and M. C. Viano. *Cours de séries temporelles*. Cours polycopié, 2008.

[3] R. H. Shumway and D. S. Stoffer. *Time series analysis and its application with R examples*. Springer Texts in Statistics, 2006.