Séries Temporelles 1 (2A) - ENSAI

TP 1 : Premiers pas et quelques expériences numériques

Octobre 2023

Exercice 1

On étudie la série temporelle `nottem` des températures mensuelles enregistrées à Nottingham Castle entre janvier 1920 et décembre 1939. Cette série est disponible sous R :

```
notten.data <- nottem
summary(notten.data)
```

1.- Regarder la structure de ces données, puis tracer la série à l'aide de la fonction `plot.ts`. Observe-t-on une tendance ? une saisonnalité ?

2.- On tente de la désaisonnaliser par régression sur les fonctions trigonométriques, $t \to \cos\left(\frac{2\pi j}{12}t\right)$ pour $j = 1, 2, \ldots, 6$ et $t \to \sin\left(\frac{2\pi j}{12}t\right)$ pour $j = 1, \ldots, 5$. Pour cela, on crée une base trigonométrique à l'aide du code suivant :

```
freq <- matrix(1:6, nrow=1)/12
time <- matrix(1:length(notten.data), ncol = 1)
B <- cbind(cos(2*pi*time%*%freq), sin(2*pi*time%*%freq))[,-12]
B <- as.data.frame(B)
colnames(B) <- c('cos1','cos2','cos3','cos4','cos5','cos6','sin1','sin2','sin3','sin4','sin5')
#library(knitr)
#knitr::kable(head(B)) #Si vous voulez regarder les premières 6 lignes du tableau B
```

3.- Effectuer la régression sur la base créée en utilisant la fonction `lm` puis faire un résumé (`summary`) des résultats obtenus.

4.- Recommencer en ne gardant que les éléments significatifs de la base.

5.- Créer une série temporelle contenant la partie saisonnière obtenue sur cette base puis une contenant les résidus de la régression. On fera appel aux commandes `$fitted` et `$residuals`.

6.- Tracer les trois séries dans une même fenêtre.

7.- Filtrer la série de départ pour éliminer la saisonnalité en utilisant le filtre de différenciation saisonnière via la fonction `diff`. Superposer la série ainsi obtenue avec les résidus précédents. Commenter.

8.- Même question en utilisant cette fois la moyenne mobile arithmétique d'ordre 13 modifiée $M_6^*(B)$. On utilisera la fonction `filter` qui prend en argument la série à filter et un vecteur contenant les coefficients θ_i de la moyenne mobile à appliquer.

9.- Pour terminer, utiliser la fonction `decompose`. Que fait-elle? Faire un graphe de sa sortie. Commenter.

Exercice 2

Soient $(X_t)_t$ et $(Y_t)_t$ deux processus MA définis pour tout $t \in \mathbb{Z}$ par

$$X_t = 3 + \epsilon_t + 2\epsilon_{t-1} \qquad et \qquad Y_t = \eta_t - \frac{1}{2}\eta_{t-1} + \frac{1}{3}\eta_{t-2}$$

où $(\epsilon_t)_t$ et $(\eta_t)_t$ sont deux bruits blancs indépendants de variance respective $\sigma_\epsilon^2 = 1$ et $\sigma_\eta^2 = 2$. On pose $Z_t = X_t + Y_t$.

1.- Simuler une trajectoire de X, de Y et de Z, de taille $n = 1000$.

2.- A l'aide des fonctions `acf` et `pacf`, tracer les fonctions d'autocorrélation et d'autocorrélation partielle empirique de chacun des trois processus.

Exercice 3

Soit le processus AR(1) défini, pour tout $t \in \mathbb{Z}$, par

$$X_t = \mu + \phi X_{t-1} + \epsilon_t$$

où $(\epsilon_t)_t$ est un brut blanc de variance σ^2.

1.- Simuler une trajectoire de ce processus de taille $n = 1000$ pour $\mu = 1$, $\sigma^2 = 0.5$ et $\phi = 0.45$, puis pour $\phi = 0.95$ et enfin pour $\phi = 1$.

2.- Pour chaque valeur de ϕ représenter graphiquement la trajectoire simulée et commentez-les.

3.- A l'aide de la fonction `acf`, tracer la fonction d'autocorrelation empirique du processus. Dans quels cas est-il stationnaire ?

4.- Dans les cas stationnaires, estimer pour chaque $k \in \{10, \ldots, n\}$, μ, ϕ et σ^2 à partir de la trajectoire $x_1, \ldots, x_k$ (on utilisera la moyenne empirique et les équations de Yule-walker). Comparer graphiquement les estimations avec les vraies valeurs des paramètres en fonction de la valeur de k.

5.- Pour chaque $k \in \{10, \ldots, n-1\}$, proposer un estimateur de la prévision de X_{k+1} à partir de l'observation de la trajectoire $x_1, \ldots, x_k$. Superposer la vraie trajectoire et sa prévision en fonction de k. Que constatez-vous selon la valeur de ϕ ?

Exercice 4

Soit $(X_t)_t$ un processus MA(2) de paramètres $m = 3$, $\theta_1 = \frac{1}{2}$ et $\theta_2 = -\frac{1}{3}$.

1.- Simuler une trajectoire pour $t = 1, \ldots, 300$.

2.- Utiliser la fonction `arima` pour estimer les paramètres m, θ_1 et θ_2 par maximum de vraisemblance à partir des 200 premières observations.

3.- Ecrire un algorithme pour prédire les valeurs de X_t pour $t = 200, 201, \ldots, 300$ et les comparer graphiquement avec les vraies valeurs observées.