

Modélisation des séries temporelles

Séance 6

Notions abordées : synthèse de l'ensemble des notions vues en cours. Vous devez savoir mettre en œuvre la méthode de Box et Jenkins pour identifier un modèle (SARIMA, d'intervention, à variable exogène) de manière raisonnée.

1 Exercice 1 : SNCF

Le fichier `sncf.csv` contient le trafic passager en 2^{de} classe de janvier 1963 à décembre 1980. On construit l'objet `ts` associé par :

```
sncf <- read.csv(file="sncf.csv",header=TRUE,sep=";")
sncf <- as.vector(t(as.matrix(sncf[,2:13])))
sncf <- ts(sncf,start = c(1963, 1), frequency = 12)
```

et on examine la chronique par :

```
sncf
```

À l'aide de la méthode de Box et Jenkins, proposez une modélisation SARIMA de cette chronique.

Faites une prévision sur l'année 1981, ainsi qu'une prévision commençant en décembre 1979 superposée aux valeurs mesurées en 1980.

2 Exercice 2 : Busban

La table `busban` donne la chronique mensuelle du trafic passagers de la RATP sur l'ensemble du réseau des bus de la banlieue parisienne de janvier 1984 à décembre 1995 (unité : milliers de passagers).

1. Chargez la table et créez l'objet `ts` adéquat :

```
busban <- read.table(file="busban.txt",sep='\t',header=TRUE)
busban <- ts(busban,start=c(1984,1),frequency=12)
```

2. Dessinez la chronique `busban`.

3. Faites une décomposition par lissage de cette chronique, et tracez le graphe de la chronique corrigée des variations saisonnières. Que peut-on dire de janvier 1987 ?

4. Modélisez la chronique `busban` par un modèle SARIMA selon la méthode de Box-Jenkins.

5. Faites une prévision sur une année.

6. Dessinez la chronique des résidus et observez qu'un des résidus semble aberrant.

7. Proposez un modèle d'intervention permettant de tenir compte de la grève des agents de la RATP en janvier 1987 dont vous quantifierez l'impact. Vous envisagerez une intervention ponctuelle, et une intervention dont l'effet s'atténue au cours du temps.

3 Exercice 3 : SOI / Recruit

La table `soirecruit` contient 453 valeurs mensuelles de deux chroniques à partir de janvier 1950 : `soi` (*Southern Oscillation Index*, mesurant l'influence des phénomènes El Niño et El Niña) et `recruit` (*Recruitment*, nombre de poissons qui éclosent dans une certaine région de l'océan Pacifique).

Plus d'informations sur le SOI ici :

<http://www.bom.gov.au/climate/glossary/soi.shtml>

Il semble raisonnable de penser que le nombre de poissons est lié aux conditions climatiques.

1. Chargez la table, et créez l'objet `ts` adéquat :

```
soirecruit <- read.table("soirecruit.txt", sep='\t', header=T)
X <- ts(soirecruit[,1], start=c(1950,1), frequency=12) # soi
Y <- ts(soirecruit[,2], start=c(1950,1), frequency=12) # recruit
```

2. Dessinez les deux chroniques.
3. À l'aide de la méthode de Box et Jenkins, établissez un modèle « à fonction de transfert » permettant d'expliquer `recruit` en fonction de `soi`. En particulier, identifiez le décalage. Vous remarquerez aussi que `recruit` est négativement corrélé à `soi`, et le signe du terme relatif à `soi` dans le modèle final.

Remarque : les fonctions R ne permettent pas une estimation correcte du modèle (problèmes de convergence numérique), contrairement à ce que l'on pouvait faire avec SAS. Le but est de mettre en œuvre les grandes lignes de la méthode et de discuter les difficultés rencontrées.

Cet exercice est inspiré de : Times Series and Analysis and Its Applications, par R.H. Shumway et D.S. Stoffer, Springer, 2011.