TD 3 : Chaînes de Markov

M2 Statistique et économétrie
Séries temporelles

2010-11

Objectif : L'objectif de ce TP est de manipuler les outils de modélisation par chaîne de Markov.

Nous allons étudier les données de Azzilini et Bowman (1990) sur des éruptions du Geyser Old Faithful dans le parc du Yellow Stone aux Etats-Unis (voir photo). Il s'agit de données sur la durée des éruptions et les temps d'attente entre 2 éruptions.

Nous allons nous intéresser à la durée des éruptions, mais nous travaillerons avec une forme discrétisée de ces observations. En pratique, on symbolise par 0 les éruptions de courte durée (moins de 3 minutes) et par 1 celles de longue durée (plus de 3 minutes).

1. Analyse descriptive

Pour avoir une idée on peut commencer par visualiser la série temporelle et la distribution de sa loi stationnaire à l'aide des lignes de code suivantes. Commentez les graphiques.

```
library(MASS)
D = faithful
summary(D)
plot(ts(D$eruptions),ylab="Durée des éruptions (s)")%$
points(D$eruptions)
dev.new()
hist(D$eruptions)
```

Dans la suite nous utiliserons un jeu de données plus long et déjà discrétisé. Vous pouvez le télécharger sur la page web du cours : `faithful.txt`. Il n'est pas aisé de tracer la série temporelle sous la forme d'un graphique facile à lire. Nous proposons la représentation suivante.

```
d = read.table("faithful.txt",header=FALSE)
L = length(d) ;
xx = t(matrix(1:L,L,2)) ; yy = t(matrix(c(0*(1:L)),d+1),L,2) ;
txt = which(d==1)
matplot(xx[,txt],yy[,txt],col = "black",tly=1,type="l")
txt = which(d==0)
matlines(xx[,txt],yy[,txt],col = "red",tly=1,type="l")
```

Commentez le graphique obtenu.

2. Modélisation

2.1. Chaîne de Markov d'ordre un

L'idée la plus naturelle pour modéliser cette série temporelle est de la décrire par une chaîne de Markov d'ordre un.

1. Estimer les paramètres de la chaîne de Markov.
2. En déduire la loi stationnaire et la fonction d'autocorrélation (ou d'autocovariance) du modèle (jusqu'à k=8).
3. Comparer ces statistiques avec les statistiques empiriques.
4. Commenter les résultats obtenus.

2.2. Chaîne de Markov d'ordre deux

On a observé dans la question précédente qu'un modèle de chaîne de Markov d'ordre un ne permet pas de reproduire correctement la fonction d'autocorrélation des données. On propose alors d'ajuster un modèle d'ordre deux.

1. Estimer les paramètres de la chaîne de Markov. Pour simplifier les calculs, on peut se ramener à une chaîne de Markov d'ordre un en regroupant les données par couples ; par exemple, regarder les transition entre l'état (0,0) et l'état (0,1). L'espace d'états est alors de dimension 4 : $\{(0,0),(0,1),(1,0),(1,1)\}$. Rq : ici l'état (0,0) n'est jamais observé.
2. En déduire la loi stationnaire et la fonction d'autocorrélation du modèle.
3. Comparer ces statistiques avec les statistiques empiriques.
4. Commenter les résultats obtenus.