DM2 — M1 finance

22 octobre 2012

Exercice 1 : Description d'une série temporelle

1. Ouvrir la série temporelle `AirPassengers` (stockée sous `R`). Que représente-elle ? À quelle fréquence est-elle échantillonnée ?
2. Tracer le graphique de cette série temporelle.
3. Analyser graphiquement cette série (tendance et composante périodique).
4. Utiliser la fonction `decompose` de `R` pour trouver la tendance et la composante périodique.
5. Pour analyser plus finement cette série, on va devoir faire une régression. Créér un vecteur `Passagers` et un vecteur `t` qui donnent le nombre de passagers et le temps.

Exercice 2 : Un peu de programmation

On veut créer une fonction `reg` qui prenne en argument y et une matrice $M = (1, x_1 - \bar{x}_1, \ldots, x_p - \bar{x}_p)$ et qui calcule $\hat{A}$ telle que la distance entre $y_M = M\hat{A}$ et y soit la plus petite possible.

1. Que vaut M^*M ?
2. Rappeler l'expression du projeté de y sur cet espace.
3. Rappeler l'expression du vecteur de régression $\hat{A}$.
4. Rappeler l'expression des critères de détermination et de détermination ajusté. Dans quel cas utilise-t-on le critère de détermination ? Le critère de détermination ajusté ?
5. Construire une fonction `reg` qui renvoie :
   - le projeté de y sur l'espace $\mathrm{Vect}\,(1, x_1, \ldots, x_p)$
   - le vecteur de régression $\hat{A}$.
   - le critère de détermination R^2
   - le critère de détermination ajusté.

   On stockera ces éléments dans une liste. La fonction `reg` doit aussi afficher $\hat{A}$ et les deux critères de détermination. **Indication:** : la fonction `solve` permet d'inverser une matrice.

Exercice 3 : Régression

On va faire une régression du logarithme du nombre de passagers, `lnP`, en fonction du temps. On utilisera la fonction qu'on vient de construire.

1. faire une régression de `lnP` en fonction de `t` grâce à la fonction `reg`. Tracer la fonction de régression sur le graphe existant.
2. Regarder la fonction de corrélation entre les vecteurs t et t^2. Peut-on utiliser la fonction `reg` pour faire une régression de `lnP` en fonction de t et t^2 ?

   La fonction `lm` de `R` sait faire cette régression. Cependant, notre fonction permet de tracer beaucoup plus rapidement les courbes de régression.
3. On va ruser : on va faire une régression de `lnP` en fonction de $t - 1949$ et $(t - 1949)^2$. Montrer que cette régression revient à faire une régression en fonction de t et t^2. Calculer la nouvelle corrélation. La régression est-elle possible ?
4. Faire la régression. Cette régression est-elle meilleure que la précédente (justifier) ? Si oui, superposer la nouvelle courbe de régression au graphe.

5. Faire de même pour t, t^2 et t^3. Quelle est la meilleure régression ?

On prend maintenant la meilleure des deux régressions précédentes et on ajoute des vecteurs de régression qui prennent en compte la saisonnalité. Pour mieux voir ce qui se passe, on va faire une régression sur $S = ln(P) -$ la meilleure courbe de régression.

6. Quelle est la période T de la série temporelle ?
7. Tracer S en fonction de t.
8. Faire une régression de S par rapport aux vecteurs $\cos(2\pi t/T), \sin(2\pi t/T)$. Superposer la courbe de régression sur le graphique.
9. Une autre façon de faire est de calculer la variation moyenne. Calculer la variation moyenne de S. On appellera cette moyenne S_m (S_m est un vecteur de longueur 12).
   Indication: : à partir du vecteur S, construire une matrice «bien choisie» et utiliser la fonction `rowMeans` ou `colMeans`.
10. Créer une fonction `SP` de période T qui soit égale à la moyenne S_m sur une année.
11. Faire une régression de S en fonction de $SP(t)$. Superposer la courbe de régression au graphe.
    La courbe obtenue est la fonction de période T qui est la plus proche de la courbe S (pour une certaine norme). Cela doit se vérifier sur les coefficients de détermination. Par contre, elle est beaucoup plus irrégulière que le polynôme trigonométrique obtenu précédemment.
12. Faire une régression de `lnP` en fonction de $SP(t)$ et t ou (t et t^2) ou (t et t^2 et t^3) en fonction de ce qui a donné la meilleure régression pour la tendance.
13. Quel est le critère de détermination ajusté ? Est-ce qu'on a une bonne régression ?
14. Tracer `lnP` en fonction de t et superposer la courbe de régression.
15. Quelle courbe de régression approche la série temporelle `AirPassengers` ? Le vérifier graphiquement.

Exercice 4 : Résidus

On s'intéresse maintenant aux résidus de la régression de `lnP`.

1. Calculer les résidus. Quelle est leur variance ? Tracer la fonction d'autocorrélation (on utilisera la fonction `acf` de `R`). Le résultat de cette commande est une liste.
2. Si les résidus sont un bruit blanc de moyenne nulle et de variance σ^2, calculer l'espérance et la variance (théorique) du coefficient d'autocovariance d'odre k $\hat{\sigma}_n(k) = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \varepsilon_{i+k}$.
3. Quelle est la loi asymptotique du coefficient d'autocovariance d'odre k ?
4. On considère le coefficient d'autocorrélation d'odre 1. Proposer un test permettant de décider si les résidus sont un bruit blanc centré ou non (pour σ^2 connu).
5. Calculer la covariance (théorique) entre deux coefficients d'autocovariance $\hat{\sigma}(k)$ et $\hat{\sigma}(k')$, $k > k' \geq 0$.
6. Quelle est la loi asympotique de la somme des carrés des coefficients d'autocovariance ?
7. On considère les 20 premiers coefficients d'autocovariance. Proposer un test permettant de savoir si les résidus sont un bruit blanc centré ou non (avec σ^2 connu).
8. La somme des carrés des coefficients d'autocorrélation suit une loi du χ^2 :

$$Q = n \sum_{k=1}^{p} \hat{\rho}^2(k) \sim \chi^2(p)$$

   Construire un test sous `R` basé sur les 20 premiers coefficients d'autocorrélation permettant de savoir si les résidus sont un bruit blanc ou non. **Indication:** la loi du χ^2 est appelée `chisq` sous `R`.
9. Quelle est la p-valeur de ce test ? Peut-on supposer que les résidus suivent un bruit blanc ?