Series temporelles - TP3

M1 Mathématiques et finance 2019–2020

Responsable : Adrien Hardy, email : adrien.hardy@univ-lille.fr

Instructions : *3 heures. Envoyer un compte-rendu par email en fin de séance.*

Exercice 1 (Estimer les paramètres d'un ARMA)

(1) Simuler une trajectoire de longueur 500 de la série :

$$X_t = 1.5X_{t-1} - 0.75X_{t-2} + \varepsilon_t \;, \qquad \varepsilon_t \sim \text{ BB(0,1)}.$$

Cette série est-elle stationnaire, causale, inversible (justifier) ?

(2) Afficher le graphe de son `acf` (autocorrelation function) puis de sa `pacf` (partial autocorrelation function) et commenter.

(3) Demander à votre voisin(e) de simuler en cachette un processus AR(p) ou MA(q) de son choix de longueur 1000 puis essayer de retrouver p ou q.

(4) Simuler une trajectoire de longueur 1000 de l'ARMA(2,1) :

$$X_t = 0.89X_{t-1} - 0.5X_{t-2} + \varepsilon_t - 0.23\varepsilon_{t-1} \;, \qquad \varepsilon_t \sim \text{ BB(0,1)}.$$

Donner son `acf` et `pacf`. Qu'en pensez-vous ? Retrouver les paramètres de ce modèle depuis la simulation. Essayer ensuite de modéliser cette trajectoire par un ARMA(p,q) avec $1 \leqslant p, q \leqslant 4$ et relever à chaque fois la valeur de son AIC et BIC[1] et identifier le modèle d'AIC et/ou BIC minimal.

[1] Si on considère une classe de modèles paramétrés (comme ARMA(p,q)) de paramètres θ (ici $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$) et qu'on note k le nombre de paramètres (ici $k = p + q$), on définit le critère d'information d'Akaike par AIC $:= 2k - 2\max_\theta \log \text{Vraisemblance}(\theta|\text{données})$ et le critère d'information Bayésien par BIC $:= \log(N)k - 2\max_\theta \log \text{Vraisemblance}(\theta|\text{données})$ où N est le nombre d'observations (données). Minimiser un critère AIC ou BIC revient à sélectionner le modèle ayant la vraisemblance la plus grande avec un nombre de paramètres "assez petit", au sens des formules ci-dessus. L'AIC pénalise le nombre de paramètres moins fortement que le BIC (si $N \geqslant 8$).

Exercice 2 (Pétrole)

(1) Importer oil.csv de ma page web sur Python et faire un graphe. Cette série représente le prix hebdomadaire du pétrole brut en dollars par baril entre 2000 et mi-2010.

(2) Calculer la série des log-rendements[2] et afficher son chronogramme, son `acf` et sa `pacf`. Commentez.

(3) Ajuster un modèle AR qui vous paraît convenir et étudier les résidus de cette modélisation. Proposez une prédiction pour la semaine suivant la dernière date de la série et expliquer dans un court rapport pourquoi cette prédiction est pertinente ou non.

(4) Même question avec un modèle ARMA.

Exercice 3 (Algorithme de Durbin-Levinson)

Créer une fonction qui étant donné une réalisation de série temporelle $x_1, \dots, x_n$ et $p \geqslant 1$ renvoie la version empirique du vecteur $(\phi_{1,p}, \dots, \phi_{p,p})$ des coordonnées du prédicteur linéaire optimal $X_t^{*,p}$, ainsi que le risque quadratique (empirique) σ_p^2. Tester votre algorithme sur la série temporelle de la question (1) de l'exercice 1 et donner une estimation des coefficients du modèle.

[2]On rappelle que le log-rendement x_t d'un actif p_t c'est $x_t = \log(p_{t+1}/p_t) = \log(p_{t+1}) - \log(p_t)$.