Series temporelles - TP1

M1 Mathématiques et finance 2019–2020

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

Instructions : *3 heures. Envoyer un compte-rendu par email.*

Consignes préliminaires : La partie programmation est à réaliser avec Python. Le compte-rendu doit être envoyé *au format pdf* (par exemple réalisé avec un Jupyter ou Ipython notebook) avant la fin du TP sur mon adresse email.

Exercice 1 (Décomposition canonique).

(a) Soit $(\varepsilon_t)_{t\in\mathbb{Z}}$ un bruit blanc gaussien centré réduit. Simuler les séries temporelles $(X_t)_{t=1\ldots500}$ suivantes, représenter-les graphiquement (chronogramme) et superposer à ces graphiques le graphe de la partie déterministe.

1. $X_t = \varepsilon_t$
2. $X_t = 2\sin(\pi t/100) + \varepsilon_t$
3. $X_t = \sqrt{t/10} - 4 + 2\sin(\pi t/100) + \varepsilon_t$
4. $X_t = \frac{t}{200}\sin(\pi t/100)\ \varepsilon_t$

(b) Identifier la tendance, la saisonnalité et le bruit/résidu dans chacune des séries précédentes. Comment distinguer à l'oeil nu un modèle additif d'un multiplicatif ?

Exercice 2 (CAC40). On s'intéresse au jeu de données `EuStockMarkets` disponible sous R qui vient avec l'information : *Contains the daily closing prices of major European stock indices: Germany DAX (Ibis), Switzerland SMI, France CAC, and UK FTSE. The data are sampled in business time, i.e., weekends and holidays are omitted.*

On l'importe sous Python avec la commande :

```
EuStockMarkets = statsmodels.api.datasets.get_rdataset('EuStockMarkets').data
```

(a) Après avoir examiné ce jeu de données, extraire la série temporelle du CAC40 et donner une représentation graphique. Quel modèle de décomposition (additif ou multiplicatif) proposez-vous pour cette série ? Décomposer cette série avec `statsmodels.tsa.seasonal import seasonal_decompose()` et représenter graphiquement chacune des composantes.

(b) Extraire la tendance m_t et superposer son graphe à celui de la série temporelle.

(c) Que vaut la somme de la saisonnalité sur une période ?

(d) Extraire la série des bruits résiduels et étudier cette série (moyenne, variance, histogramme, etc). Dessiner la fonction d'autocorrélation empirique du bruit résiduel à l'aide de `statsmodels.tsa.stattools.acf()` – *Attention aux valeurs* `NAN`.

Exercice 3 (Un modèle autoregressif). Etant donné $(\varepsilon_t)_{t\in\mathbb{Z}} \sim \mathrm{BB}(0,\sigma^2)$, on veut définir un processus $(X_t)_{t\in\mathbb{Z}}$ qui satisfait l'équation

$$X_t = \theta X_{t-1} + \varepsilon_t$$

pour tout $t \in \mathbb{Z}$ avec un paramètre $\theta \in \mathbb{R}$. C'est un premier exemple de processus autorégressif.

(a) Donner une condition sur θ pour que

$$X_t := \sum_{k=0}^{\infty} \theta^k \varepsilon_{t-k} \in L^2$$

pour tout $t \in \mathbb{Z}$. On supposera maintenant cette condition réalisée.

(b) Montrer que ce processus satisfait $X_t = \theta X_{t-1} + \varepsilon_t$. Montrer que c'est un processus stationnaire et calculer sa fonction d'autocorrelation.

(c) Montrer que si un autre processus $(X'_t)_{t\in\mathbb{Z}}$ satisfait $X'_t = \theta X'_{t-1} + \varepsilon_t$ pour tout $t \in \mathbb{Z}$ alors $X_t = X'_t$ pour tout $t \in \mathbb{Z}$.

(d) (Simulation) On définit maintenant $(\tilde{X}_t)_{t\in\mathbb{N}}$ par $\tilde{X}_0 := 0$ et $\tilde{X}_t = \theta \tilde{X}_{t-1} + \varepsilon_t$ pour tout $t \in \mathbb{N}$. Montrer que pour tout $t \in \mathbb{N}$ et $\delta > 0$,

$$\mathbb{P}(|X_t - \tilde{X}_t| \geqslant \delta) \leqslant \frac{\sigma^2}{(1-\theta^2)\delta^2}\,\theta^{2t}.$$

En déduire une façon approchée de simuler $(X_t)_{t\in\mathbb{Z}}$ et l'implémenter. Dessiner la fonction d'autocorrélation empirique de cette série temporelle.