Université Paris VII and I

Master 2 M0 2020 − 2021

Analyse des séries financières

Examen final, Février 2021

3h00, sans aucun document sauf une feuille A4 RV

1. Soit $(\varepsilon_t)_{t\in\mathbb{Z}}$ une suite de variables aléatoires gaussiennes centrées indépendantes et identiquement distribuées telles que $\mathbb{E}\left[\varepsilon_0^2\right] = 1$. On s'intéresse à la suite de variables $(X_t)_{t\in\mathbb{Z}}$ définies par récurrence:
$$X_t = \varepsilon_t\,\sigma_t \quad \text{avec} \quad \sigma_t^2 = \omega_0^* + b_1^*\,\sigma_{t-1}^2 \quad \text{pour tout } t \in \mathbb{Z}$$
où $(\omega_0^*, b_1^*) \in]0,\infty[\times[0,\infty[$ un vecteur de paramètres inconnus. Déterminer à quelle condition sur (ω_0^*, b_1^*) il existe une unique solution stationnaire pour $(X_t)_t$. Dans ce cas, déterminer l'autocorrélation de $(X_t)_t$, puis la loi de $(X_1,\dots,X_n)$. Démontrer qu'il n'est pas possible de déterminer un estimateur convergeant de (ω_0^*, b_1^*).

2. On veut modéliser la trajectoire des log-returns $(X_1,\dots,X_n)$ d'un indice financier. Pour cela, on suppose qu'un vrai modèle est un GARCH$(1,1)$ stationnaire d'ordre 2 tel que:
$$X_t = \varepsilon_t\,\sigma_t \quad \text{avec} \quad \sigma_t^2 = \omega_0^* + a_1^*\,X_{t-1}^2 + b_1^*\,\sigma_{t-1}^2 \quad \text{pour tout } t \in \mathbb{Z}$$
où $(\varepsilon_t)_{t\in\mathbb{Z}}$ est une suite de variables aléatoires centrées indépendantes et identiquement distribuées telles que $\mathbb{E}\left[|\varepsilon_0|\right] = 1$ et $\mathbb{E}\left[\varepsilon_0^2\right] < \infty$.
   (a) Est-il possible que $\omega_0^* = 0$? Donner sans preuve l'ensemble $\mathcal{D}$ auquel $(\omega_0^*, a_1^*, b_1^*)$, vecteur de paramètres inconnus, doit appartenir.
   (b) Ecrire en justifiant $(X_t)_t$ sous la forme d'un modèle affine causal.
   (c) On suppose que ε_0 suit une loi de Laplace de paramètre 1, c'est-à-dire que sa densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$ est $f(x) = \frac{1}{2}\,e^{-|x|}$ pour $x \in \mathbb{R}$. Déterminer en fonction des $(X_t,\sigma_t)_{1\le t\le n}$ la vraisemblance de $(X_1,\dots,X_n)$ conditionnellement à $(X_{-k})_{k\in\mathbb{N}}$. En déduire que la log-vraisemblance conditionnelle de $(X_1,\dots,X_n)$ vaut
$$L_{(X_1,\dots,X_n)} = -\sum_{t=1}^n \frac{|X_t|}{\sigma_t} + \log\big(2\,\sigma_t\big)$$
   (d) On suppose que $\theta = (\theta_i)_{1\le i\le 3} = {}^t(\omega_0, a_1, b_1)$ et on veut estimer $\theta^* = {}^t(\omega_0^*, a_1^*, b_1^*)$, à partir de la trajectoire observée $(X_1,\dots,X_n)$. Pour $\theta \in \Theta$, ensemble compact inclus dans $\mathcal{D}$ contenant θ^*, montrer que maximiser la log-vraisemblance conditionnelle en θ revient à minimiser
$$C_n(\theta) = \sum_{t=1}^n q_\theta(t), \quad \text{avec} \quad q_\theta(t) = \frac{|X_t|}{\sigma_t(\theta)} + \log\big(\sigma_t(\theta)\big)$$
   et $\sigma_t^2(\theta) = \omega_0 + a_1\,X_{t-1}^2 + b_1\,\sigma_{t-1}^2(\theta)$ pour $t \in \mathbb{Z}$.

(e) On définit la suite $(\widehat{\sigma}_t(\theta))_{t\geq 0}$ telle que

$$\widehat{\sigma}_0(\theta)=0 \quad \text{et} \quad \widehat{\sigma}_t^2(\theta)=\omega_0+a_1\,X_{t-1}^2+b_1\,\widehat{\sigma}_{t-1}^2(\theta) \quad \text{pour } t\geq 1.$$

Montrer que $\mathbb{E}\big[\sup_{\theta\in\Theta}\big|\sigma_t^2(\theta)-\widehat{\sigma}_t^2(\theta)\big|\big]\leq \mathbb{E}\big[\sup_{\theta\in\Theta}\big|\sigma_0^2(\theta)\big|\big]\,b_1^t$ et $\mathbb{E}\big[\sup_{\theta\in\Theta}\big|\sigma_0^2(\theta)\big|\big]<\infty$.

(f) On note $\widehat{C}_n(\theta)=\displaystyle\sum_{t=1}^n \widehat{q}_t(\theta)$ où $\widehat{q}_\theta(t)=\dfrac{|X_t|}{\widehat{\sigma}_t(\theta)}+\log\big(\widehat{\sigma}_t(\theta)\big)$. Démontrer que

$$\mathbb{E}\big[\sup_{\theta\in\Theta}\big|q_t(\theta)-\widehat{q}_t(\theta)\big|\big]\leq \frac{1}{2\,\underline{\omega}_0^2}\Big(\mathbb{E}\big[\sup_{\theta\in\Theta}\big|\sigma_t^2(\theta)-\widehat{\sigma}_t^2(\theta)\big|\big] \\ +2\sqrt{\mathbb{E}\,[X_0^2]}\,\sqrt{\mathbb{E}\big[\sup_{\theta\in\Theta}\big|\sigma_t^2(\theta)-\widehat{\sigma}_t^2(\theta)\big|\big]}\Big)$$

où l'on note $\underline{\omega}_0=\min\big(\omega_0,\ \theta={}^t(\omega_0,a_1,b_1)\in\Theta\big)$. En déduire qu'il existe $C>0$, indépendante de t, tel que

$$\mathbb{E}\big[\sup_{\theta\in\Theta}\big|q_t(\theta)-\widehat{q}_t(\theta)\big|\big]\leq C\,b_1^{t/2}.$$

(g) Soit $(u_n)_{n\in\mathbb{N}}$ une suite de réels positifs et tel que $\sum_{k\in\mathbb{N}}u_k/k<\infty$. Montrer qu'alors $\frac{1}{n}\sum_{t=1}^n u_t \underset{n\to+\infty}{\longrightarrow} 0$. On pose $S_n=\sum_{t=1}^n\frac{1}{t}\sup_{\theta\in\Theta}\big|q_t(\theta)-\widehat{q}_t(\theta)\big|$. Montrer que pour tout $\varepsilon>0$, $\lim_{p\to\infty}\lim_{q\to\infty}\mathbb{P}\big(|S_p-S_q|\geq\varepsilon\big)=0$. En déduire que $S_n\overset{p.s.}{\underset{n\to+\infty}{\longrightarrow}}0$ puis que

$$\frac{1}{n}\sup_{\theta\in\Theta}\big|\widehat{C}_n(\theta)-C_n(\theta)\big|\overset{p.s.}{\underset{n\to+\infty}{\longrightarrow}}0.$$

(h) Démontrer que $\mathbb{E}\big[\sup_{\theta\in\Theta}\big|q_0(\theta)\big|\big]<\infty$ et en déduire que $\dfrac{1}{n}\sup_{\theta\in\Theta}\big|C_n(\theta)-\mathcal{C}(\theta)\big|\overset{p.s.}{\underset{n\to+\infty}{\longrightarrow}}0$ avec $\mathcal{C}(\theta)=\mathbb{E}\big[q_0(\theta)\big]$.

(i) Montrer que $\mathcal{C}(\theta^*)$ est l'unique minimum de $\mathcal{C}(\theta)$ sur Θ. En déduire que si $\widehat{\theta}_n\in\operatorname*{argmin}_{\theta\in\Theta}\widehat{C}_n(\theta)$ alors $\widehat{\theta}_n\overset{p.s.}{\underset{n\to+\infty}{\longrightarrow}}\theta^*$.

(j) On peut montrer (ne pas le faire!) que $\mathbb{E}\big[\sup_{\theta\in\Theta}\big|\frac{\partial}{\partial\theta_i}q_t(\theta)\big|^2\big]<\infty$ pour $1\leq i\leq 3$, que $F(\theta^*)=\Big(\mathbb{E}\big[\frac{\partial^2}{\partial\theta_i\partial\theta_j}q_t(\theta^*)\big]\Big)_{1\leq i,j\leq 3}$ est une matrice inversible et

$$\sqrt{n}\,\big(\widehat{\theta}_n-\theta^*\big)\overset{\mathcal{L}}{\underset{n\to+\infty}{\longrightarrow}}\mathcal{N}_3\big(0\,,\,F^{-1}(\theta^*)\,G(\theta^*)\,F^{-1}(\theta^*)\big)$$

avec $G(\theta)=\Big(\operatorname{cov}\big(\frac{\partial}{\partial\theta_i}q_t(\theta)\,,\,\frac{\partial}{\partial\theta_j}q_t(\theta)\big)\Big)_{1\leq i,j\leq 3}$. Montrer que

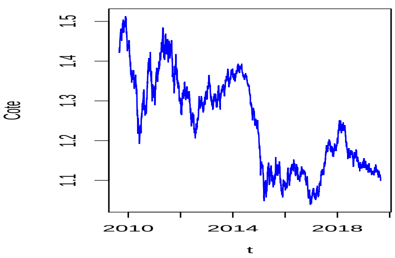
$$F^{-1}(\theta^*)\,G(\theta^*)\,F^{-1}(\theta^*)=\frac{1}{\operatorname{var}(|\xi_0|)}\,\Big(\mathbb{E}\Big[\frac{1}{\sigma_0^2(\theta^*)}\times\frac{\partial}{\partial\theta_i}\sigma_0(\theta^*)\times\frac{\partial}{\partial\theta_j}\sigma_0(\theta^*)\Big]\Big)^{-1}_{1\leq i,j\leq 3}.$$

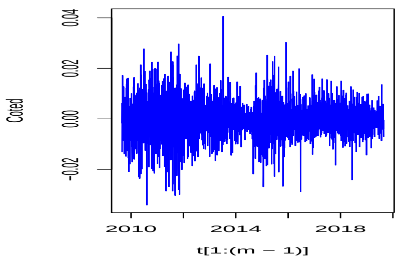
Comment peut-on estimer cette matrice? A quoi cela peut-il servir? Enfin, quel est l'intérêt de cette méthode d'estimation par rapport à l'habituelle estimation par QML?

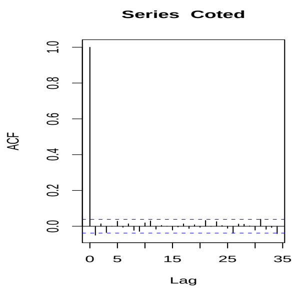
3. On s'intéresse à modéliser la cotation quotidienne en clôture du taux de change Euro/Dollar (variable `Cote`) depuis le 01/09/2009 jusqu'au 01/09/2019. On utilise le logiciel `R` à cet effet.

   (a) Voici les premières commandes effectuées:

```
Cote=as.numeric(ED$Close)
t=as.numeric(ED$Time)
m=length(Cote)
plot(t,Cote,'l',col='blue')
Coted=Cote[2:m]-Cote[1:(m-1)]
ts.plot(Coted)
acf(Coted)
Box.test(Coted, lag = 10, type = c("Ljung-Box"))
```







   Avec pour résultats numériques:

```
Box-Ljung test
data:  Coted
X-squared = 17.585, df = 10, p-value = 0.06239
```

   Question II.1: Que représente `Coded` *par rapport à* `Cote`*? Pourquoi peut-on utliser* `Coded` *plutôt que les log-returns dans ce cadre? Qu'en conclure quant à la série* `Coded`*? Qu'en déduire, dans un premier temps, quant au type de processus qu'est* `Cote`*?*

   (b) On exécute alors:

```
Fit1=garchFit(~garch(1,2),data=Coted,trace=FALSE)
summary(Fit1)
```

   Avec pour résultats numériques:

```
        Estimate  Std. Error  t value Pr(>|t|)
mu     -9.310e-05   1.188e-04   -0.784   0.4332
omega   1.876e-07   9.298e-08    2.017   0.0437 *
alpha1  4.776e-02   6.777e-03    7.048 1.82e-12 ***
beta1   4.743e-02   3.051e-02    1.555   0.1200
beta2   9.009e-01   3.172e-02   28.398  < 2e-16 ***
```

   Question II.2: Expliquer ce qui a été fait. Que conclure de la valeur `< 2e-16`*? Comment sont calculées concrètement les valeurs* `3.051e-02` *et* `1.555`*?*

   (c) On continue alors par:

```
fit3=arima(Coted^2,c(1,0,2))
summary(fit3)
```

On obtient alors:

```
SCoefficients:
         ar1      ma1      ma2  intercept
      0.9967  -0.9329  -0.0411      1e-04
s.e.  0.0019   0.0197   0.0196      0e+00
```

Question II.3: Qu'a-t-on fait ici et pourquoi? Ces résultats vous semblent-ils cohérents avec les précédents?

(d) On tape enfin:

```
AC=as.numeric(unlist(acf(fit3$residuals^2)))
B=m*sum(AC[2:11]^2)
B
1-pchisq(B,10)
```

Et on obtient les réultats:

```
> B
[1] 6.903688
> 1-pchisq(B,10)
[1] 0.7345051
```

Question II.4: Pourquoi a-t-on exécuté ces commandes? Que représentent les 2 valeurs numériques obtenues? Qu'en déduire?