Projet 3 : Estimation du modèle à volatilité stochastique de Taylor par distance L_2

Isaac GANIYU, Isabelle LACMAGO, Junior JUMBONG

Prof : Salima EL Kolei

13 avril 2024

Outline

Introduction

Contexte et Objectifs

- Découvrir puis appliquer une méthode d'estimation pour le modèle à volatilité stochastique de Taylor alternative à la méthode par Quasi Maximum de Vraisemblance.
- Cette approche s'appuie sur un critère de minimisation L2 et une stratégie de déconvolution permettant de filtrer le bruit des observations grâce à la transformée de Fourier pour retrouver les paramètres du modèle.

Présentation du modèle

modèle

Le modèle SV s'écrit :

$$\begin{cases} R_{i+1} = \exp\left(\frac{X_{i+1}}{2}\right)\xi_{i+1}^{\beta}, & \xi_{i+1} \sim \mathcal{N}\left(0, \sigma_{\xi}^2\right) \\ X_{i+1} = \phi_0 X_i + \eta_{i+1}, & \eta_{i+1} \sim \mathcal{N}\left(0, \sigma_0^2\right) \end{cases}$$

On suppose $|\phi_0| < 1$ pour la stationnarité et l'ergocité du processus Y

$$\begin{cases} Y_{i+1} = \log\left(R_{i+1}^2\right) - \mathbb{E}\left[\log\left(\xi_{i+1}^{2\beta}\right)\right] \\ \varepsilon_{i+1} = \log\left(\xi_{i+1}^2\right) - \beta\mathbb{E}\left[\log\left(\xi_{i+1}^2\right)\right] \end{cases}$$

Le modele $\log -SV$: $\begin{cases} Y_{i+1} = X_{i+1} + \varepsilon_{i+1} \\ X_{i+1} = \phi_0 X_i + \eta_{i+1} \end{cases}$

$\theta_0 = \left(\phi_0, \sigma_0^2\right)$ est le vecteur de paramètre à estimer a partir des observations $(Y_i)_{i \geqslant 1}$.

Méthodes d'estimation

Révue de littérature

- Moments (Taylor 2008) ;
- Moments généralisés : (Melino et Turnbull 1990) ;
- Quasi-maximum de vraisemblance : (Harvey, Ruiz et Shephard 1994) ;
- MCMC : (Chib, Nardari et Shephard 2002) ;

Méthodes d'estimation

Nous utiliserons deux méthodes d'estimation principales des paramètres de notre modèle :

- La méthode par quasi maximum de vraisemblance
- la méthode par minimisation et déconvolution

Pour cette seconde méthode nous utiliserons deux méthodes de calcul de la transformée de Fourrier, ce qui nous donnera en fait trois estimateurs différents. Nous comparerons ensuite ces estimateurs.

Applications

- Epidémiologie ;
- Météorologie ;
- Neuroscience ;
- Ecologie : (Ionides et al. 2011) ;
- Finance : (Johannes, Polson et Stroud 2009)

Estimateur par quasi maximum de vraisemblance

L'approche par quasi maximum de vraisemblance s'effectue en 3 étapes :

- L'estimation du processus de log volatilité par le filtre de Kalman
- Le calcul de la log vraisemblance
- La minimisation de la log vraisemblance

Nos Travaux s'appuient : El Kolei 2013

Filtre de Kalman

Le filtre de Kalman est défini comme suit :
Pour t= 0 :
Échantillonner $x_0 \sim \mathcal{N}(\hat{x}_0, P_0)$ avec $\hat{x}_0 = \frac{\mu}{1-\phi}; \quad P_0 = \frac{\sigma_0^2}{(1-\phi^2)}$.
Pour $t \geq 1$:
Étape de Prédiction :

$$\begin{cases} \widehat{x_t^-} = \mu + \phi \widehat{x}_{t-1} \\ P_t^- = \phi^2 P_{t-1} + \sigma_0^2 \\ \widehat{y_t^-} = \widehat{x_t^-} \end{cases}$$

Étape de Mise à jour

$$\begin{cases} \widehat{x}_t = \widehat{x_t^-} + K_t \left(y_t - \widehat{y_t^-} \right) \\ P_t = (1 - K_t) P_t^- \\ K_t = \frac{P_t^-}{\left(P_t^- + \sigma_\varepsilon^2\right)} \quad , \quad \sigma_\varepsilon^2 = \beta^2 \frac{\pi^2}{2} \end{cases}$$

Log vraisemblance et estimateur

Ensuite, la log vraisemblance est donnée par :

$$l(\theta) = \log L_\theta\left(Y_{1:n}\right) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\log F_t - \frac{1}{2}\sum_{t=1}^{n}\frac{\nu_t^2}{F_t},$$

où ν_t et F_t sont donnés par

$$\nu_t = \left(Y_t - \hat{Y}_t^-\right), \quad F_t = \mathbb{V}_\theta\left[\nu_t\right] = P_t^- + \sigma_\varepsilon^2$$

avec $\hat{Y}_t^- = \mathbb{E}_\theta\left[Y_t \mid Y_{1:t-1}\right]$ et $P_t^- = \mathbb{V}_\theta\left[\left(X_t - \hat{X}_t^-\right)^2\right]$ calculés par le Filtre de Kalman.
L'estimateur de θ_0 est alors définie comme

$$\hat{\theta}_n = \arg\max_{\theta\in\Theta} l(\theta).$$

Critère de minimisation

Cette méthode d'estimation est basée sur la minimisation de la distance L_2 définie par :

$$\|l_\theta - l_{\theta_0}\|_2^2 - \|l_{\theta_0}\|_2^2$$

où la fonction l_θ est donnée par $l_\theta(x) = \phi x f_\theta(x)$ avec f_θ la densité stationnaire de X_i. Ainsi, au point $\theta = \theta_0$ cette distance est minimale et vaut $-\|l_{\theta_0}\|_2^2$.

Critère de minimisation

Ce critère est inaccessible car il dépend de la valeur inconnue θ_0. On considère la version empirique :

$$\mathrm{P}_n m_\theta = \frac{1}{n} \sum_{i=1}^{n} m_\theta\left(\mathrm{Y}_i\right)$$

avec n le nombre d'observations, $\mathrm{Y}_i = (Y_i, Y_{i+1})$ et :

$$m_\theta\left(\mathrm{y}_i\right) : (\theta, \mathrm{y}_i) \in \left(\Theta \times \mathbb{R}^2\right) \mapsto m_\theta\left(\mathrm{y}_i\right) = \|l_\theta\|_2^2 - 2y_{i+1} u^*_{l_\theta}\left(y_i\right),$$

où la fonction u_v est donnée par $u_v(x) = \frac{1}{2\pi} \frac{v^*(-x)}{f^*_\varepsilon(x)}$ et $u^*(t) = \int e^{itx} u(x) dx$ est la transformée de Fourier de la fonction u(t). L'estimateur $\widehat{\theta}_n$ de θ_0 sera obtenu comme

$$\widehat{\theta}_n = \arg\min_{\theta \in \Theta} \mathrm{P}_n m_\theta$$

Démonstration 1

Montrons que $\mathbb{E}\left[m_\theta\left(Y_1\right)\right] = \|l_\theta - l_{\theta_0}\|_2^2 - \|l_{\theta_0}\|_2^2$

$$\begin{aligned} m_\theta\left(Y_1\right) &= \|l_\theta\|_2^2 - 2Y_2 u^*_{l_\theta}\left(Y_1\right) \\ &= \|l_\theta\|_2^2 - 2\left(\phi_0 X_1 + \eta_2 + \varepsilon_2\right) u^*_{l_\theta}\left(Y_1\right) \end{aligned}$$

On sait que $\mathbb{E}\left(\eta_2\right) = \mathbb{E}\left(\varepsilon_2\right) = 0$ et $\eta_2, \varepsilon_2 \perp Y_1$, ainsi on a :

$$\begin{aligned} \mathbb{E}\left[m_\theta\left(Y_1\right)\right] &= \|l_\theta\|_2^2 - 2\phi_0 \mathbb{E}\left[X_1 u^*_{l_\theta}\left(Y_1\right)\right] \\ \mathbb{E}\left[X_1 u^*_{l_\theta}\left(Y_1\right)\right] &= \mathbb{E}\left[X_1 \int e^{iY_1 x} u_{l_\theta}(x)\, dx\right] \\ &= \mathbb{E}\left[X_1 \int e^{iY_1 x}\left(\frac{1}{2\pi}\frac{l^*_\theta(-x)}{f^*_\varepsilon(x)}\right)\, dx\right] \\ &= \frac{1}{2\pi}\int \mathbb{E}\left[X_1 e^{i(X_1+\varepsilon_1)x}\right] \cdot \frac{l^*_\theta(-x)}{f^*_\varepsilon(x)}\, dx, \varepsilon_1 \perp X_1 \end{aligned}$$

Démonstration 1

$$
\begin{aligned}
\mathbb{E}\left[X_1 u^*_{l_\theta}(Y_1)\right] &= \frac{1}{2\pi}\int \frac{\mathbb{E}\left[e^{i\varepsilon_1 x}\right]}{f^*_\varepsilon(x)} \cdot \mathbb{E}\left[X_1 e^{iX_1 x}\right] l^*_\theta(-x)dx \\
&= \frac{1}{2\pi}\mathbb{E}\left[X_1 \int e^{iX_1 x} l^*_\theta(-x)dx\right]; \mathbb{E}\left[e^{i\varepsilon_1 x}\right] = f^*_\varepsilon(x) \\
&= \frac{1}{2\pi}\mathbb{E}\left[X_1 \left(l^*_\theta\right)^*(-X_1)\right] \\
&= \frac{1}{2\pi}\mathbb{E}\left[X_1 \times 2\pi l_\theta(X_1)\right] = \mathbb{E}\left[X_1 \, \mathrm{l}_\theta(X_1)\right] \\
&= \int x\phi x f_{\theta_0}(x) f_\theta(x)dx, \text{ car } l_\theta(x) = \phi x f_\theta(x), f_{X_1}(x) = f_{\theta_0}(x) \\
&= \frac{1}{\phi_0}\left\langle l_\theta, l_{\theta_0}\right\rangle
\end{aligned}
$$

Démonstration 1

Donc, $\mathbb{E}\left[m_\theta\left(Y_1\right)\right] = \left\|l_\theta\right\|_2^2 - 2\left\langle l_\theta, l_{\theta_0}\right\rangle = \left\|l_\theta - l_{\theta_0}\right\|_2^2 - \left\|l_{\theta_0}\right\|^2$

$$\text{car } \left\|l_\theta - l_{\theta_0}\right\|_2^2 = \left\|l_\theta\right\|_2^2 - 2\left\langle l_\theta, l_{\theta_0}\right\rangle + \left\|l_{\theta_0}\right\|_2^2$$

Démonstration 2

Montrons que $\|l_\theta\|_2^2 = \frac{\phi^2\gamma}{4\sqrt{\pi}}, \gamma = \frac{\sigma_0}{1-\phi_0}$

$$l_\theta(x) = \phi x f_\theta(x)$$

f_θ est la densité stationnaire de X. On a :

$$\begin{cases} X = \phi_0 X + \eta \Rightarrow X = \frac{1}{1-\phi_0}\eta \\ \mathbb{E}(X) = 0 \\ \mathrm{Var}(X) = \frac{1}{(1-\phi_0)^2}\,\mathrm{Var}(\eta) = \frac{\sigma_0^2}{(1-\phi_0)^2} = \gamma^2 \end{cases}$$

Donc $l_\theta(x) = \phi x \frac{1}{\sqrt{2\pi\gamma^2}} e^{-\left(\frac{1}{2\gamma^2}x^2\right)}$, et

$$\|l_\theta\|_2^2 = \int |l_\theta(x)|^2\, dx = \frac{\phi^2\gamma}{4\sqrt{\pi}}$$

Critère de minimisation

La démonstration 1 nous permet de justifier le choix du critère empirique car en moyenne il renvoie à la définition de la distance L_2.
$\mathbb{E}\left[m_\theta\left(\mathrm{Y}_1\right)\right]=\|l_\theta-l_{\theta_0}\|_2^2-\|l_{\theta_0}\|_2^2$.
De plus,

$$\begin{aligned} P_n m_\theta &= \frac{1}{n}\sum_{i=1}^{n}\left(\|l_\theta\|_2^2-2Y_{i+1}u_{l_\theta}^*\left(Y_i\right)\right) \\ &= \|l_\theta\|_2^2-\frac{2}{n}\sum_{i=1}^{n}Y_{i+1}u_{l_\theta}^*\left(Y_i\right) \end{aligned}$$

Ainsi, la démonstration 2 nous permet d'avoir une expression simplifiée de $\|l_\theta\|_2^2$. On aboutit au critère :

$$\widehat{\theta}_n=\arg\min_{\theta\in\Theta}\left\{\frac{\phi^2\gamma}{4\sqrt{\pi}}-\frac{2}{n}\sum_{i=1}^{n-1}Y_{i+1}u_{l_\theta}^*\left(Y_i\right)\right\}$$

Fonction de déconvolution

Dans le critère précédant il nous faut deux éléments

- La fonction de déconvolution u_{l_θ} donnée par :

$$u_{l_\theta}(y) = \frac{1}{2\sqrt{\pi}} \left(\frac{-i\phi y \gamma^2 \exp\left(\frac{-y^2}{2}\gamma^2\right)}{\exp(-i\mathcal{E}y) 2^{i\beta y} \Gamma\left(\frac{1}{2} + i\beta y\right)} \right), \quad \mathcal{E} = \beta \mathbb{E}\left[\log\left(\xi_{i+1}^2\right)\right]$$

- La transformée de Fourrier : Nous utilisons dans notre analyse, deux méthodes de calcul différentes de la transformée de Fourrier : Le calcul par intégrale et le calcul par la librairie fft de Python.

Transformée de Fourrier

- $u^*(t) = \int e^{itx} u(x) dx$
- fft : (Balac 2011)

$$u^*(t) = \int e^{itx} u(x)\, dx \to \int_{-T/2}^{T/2} e^{itx} u(x)\, dx$$

$$\int_{-T/2}^{T/2} e^{itx} u(x)\, dx \to \frac{T}{n} e^{i\pi t T} \sum_{k=0}^{n-1} \hat{u}_k e^{\frac{i\pi t T k}{n}}$$

Ces deux méthodes de calcul nous fournissent deux estimateurs différents.

Simulation des données

Nous implémentons le modèle log-SV afin d'obtenir les processus (X_i) de log volatilité et (Y_i). $\theta_0 = (0.7, 0.3)$, n $= 1000$, $\beta = \frac{1}{\sqrt{5\pi}}$ et $\mu = 0$

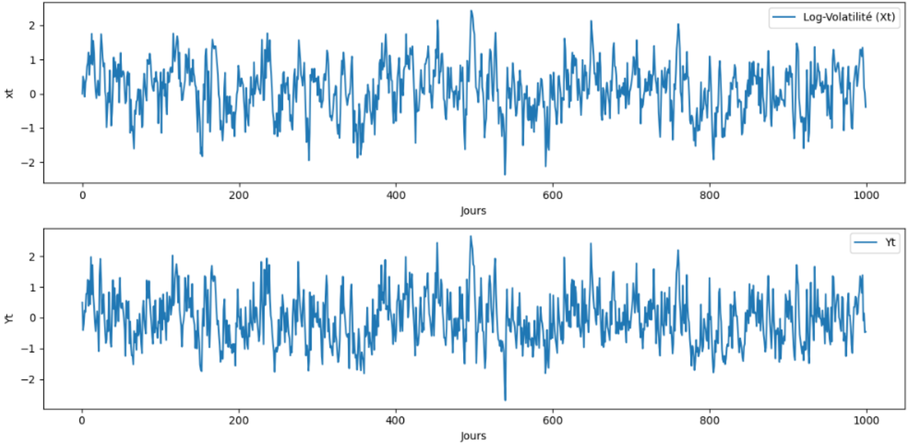

Figure – Processus simulés

Résultats

Estimation de la log-volatilité

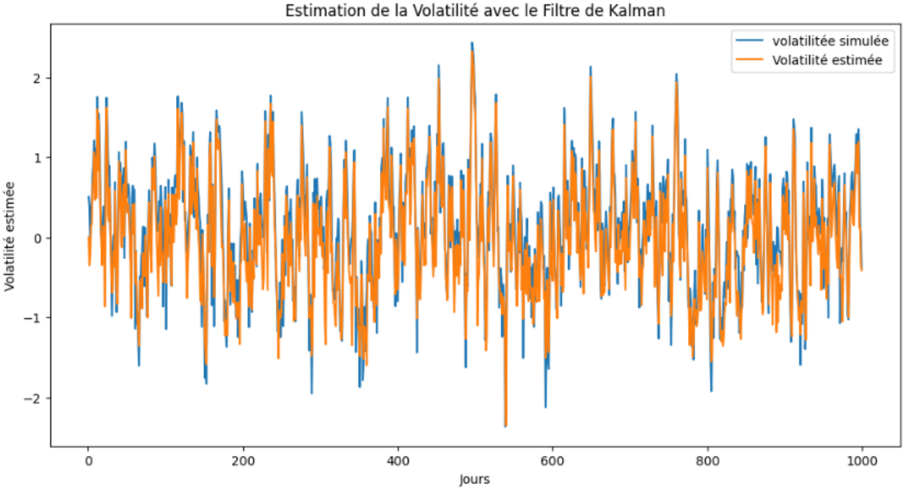

Figure – volatilité simulée vs volatilité estimée

Estimation

Une fois la log volatilité estimée obtenue, nous pouvons calculer la log vraisemblance et la minimiser.

$$\hat{\theta}_n = \arg\max_{\theta \in \Theta} l(\theta)$$

Nous prenons comme valeur initiale pour notre algorithme d'optimisation (0.5, 1) et nous spécifions que $\theta_0 = (\phi_0, \sigma_0^2)$ est tel que $|\phi_0| < 1$ et $\sigma_0^2 \geq 0$
Nous obtenons :

$$\phi_0 = \mathbf{0.6748}$$

$$\sigma_0^2 = \mathbf{0.3084}$$

Estimation

$$\widehat{\theta}_n = \arg\min_{\theta \in \Theta} \left\{ \frac{\phi^2 \gamma}{4\sqrt{\pi}} - \frac{2}{n} \sum_{i=1}^{n-1} Y_{i+1} u^*_{l_\theta}(Y_i) \right\}$$

Nous implémentons ce critère avec

$$u_{l_\theta}(y) = \frac{1}{2\sqrt{\pi}} \left(\frac{-i\phi y \gamma^2 \exp\left(\frac{-y^2}{2}\gamma^2\right)}{\exp(-i\mathcal{E}y) 2^{i\beta y} \Gamma\left(\frac{1}{2} + i\beta y\right)} \right)$$

Estimateur par minimimisation et déconvolution

- Transformée de Fourier par intégrale : Dans un premier temps nous appliquons la définition de la transformée de Fourier $u^*(t) = \int e^{itx} u(x)dx$ à la fonction de déconvolution. Pour l'implémentation du calcul de l'intégrale, nous choisissons les bornes -100 et 100.
- Transformée de Fourier par fft : Ensuite, nous appliquons la méthode Fast Fourrier Transform (FFT) de calcul de la transformée de Fourier à la fonction de déconvolution. Pour l'implémentation du calcul nous choisissons T = 100.

Transformée de Fourier par intégrale

On obtient comme estimations :

Initialisation		Estimation	
ϕ	σ_0^2	ϕ	σ_0^2
0.5	0.5	0.651	0.354
0.6	0.4	0.651	0.354
0.7	0.3	0.651	0.354
0.8	0.2	0.651	0.354
0.9	0.1	0.651	0.354
0	1	0.651	0.354
0.4	0.6	0.651	0.354
0.3	0.7	0.651	0.354
0.2	0.8	0.651	0.354
0.1	0.9	0.651	0.354

Table – Estimation des paramètres

Transformée de Fourier par fft

On obtient comme estimations :

Initialisation		Estimation	
ϕ	σ_0^2	ϕ	σ_0^2
0.5	0.5	0.743	0.264
0.6	0.4	0.743	0.264
0.7	0.3	0.743	0.264
0.8	0.2	0.743	0.264
0.9	0.1	0.743	0.264
0	1	0.743	0.264
0.4	0.6	0.743	0.264
0.3	0.7	0.743	0.264
0.2	0.8	0.743	0.264
0.1	0.9	0.743	0.264

Table – Estimation des paramètres

Comparaison des estimateurs contraste

Nous remarquons que les deux estimateurs précédents fournissent le même résultat quel que soit la valeur d'initialisation. De plus, sur cette simulation, l'estimateur par intégrale sous estime ϕ et surestime σ_0^2 tandis que l'estimateur par Fast Fourrier Transform sur estime ϕ et sous estime σ_0^2. Nous verrons par la suite que ce n'est pas toujours le cas.

Comparaison

Pour chaque méthode, nous appliquons la méthode de Monte Carlo et réitérons le processus d'estimation des paramètres. Nous utilisons 200 itérations Monte Carlo en resimulant les processus (Y_i) et (X_i) à chaque itération. La valeur initiale des paramètres est fixée à $(0.69, 0.29)$.
Nous avons ainsi plusieurs estimations des paramètres pour chaque méthode, ce qui nous permettra de les comparer.

	Contrast with quad	Contrast with fft	QML
MSE	0.013951	0.008752	0.001448

Table – MSE des différentes méthodes

Comparaison

Table – Estimations des paramètres

	Quad		FFT		QML	
	ϕ	σ_0^2	ϕ	σ_0^2	ϕ	σ_0^2
Count	200.000	200.000	200.000	200.000	200.000	200.000
Mean	0.781	0.285	0.736	0.278	0.718	0.300
Std	0.050	0.068	0.055	0.063	0.028	0.020
Min	0.643	0.105	0.591	0.109	0.616	0.237
25%	0.745	0.241	0.695	0.233	0.703	0.286
50%	0.779	0.284	0.737	0.279	0.719	0.301
75%	0.813	0.333	0.773	0.319	0.734	0.311
Max	0.924	0.486	0.901	0.495	0.792	0.359

Comparaison

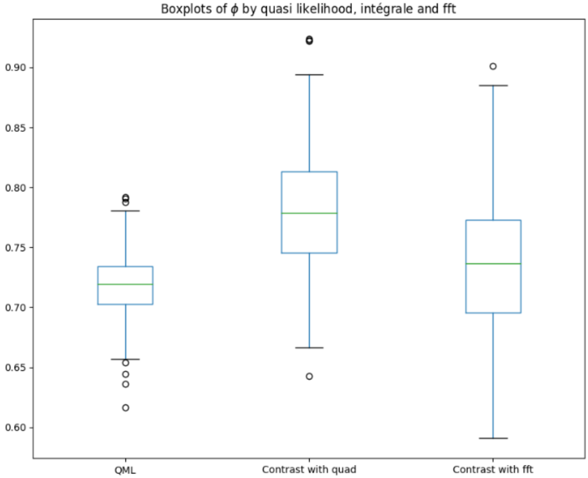

Figure – Boxplots des estimations

Comparaison

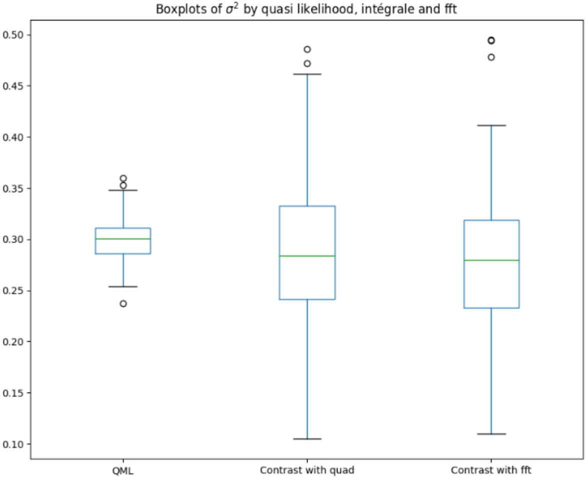

Figure – Boxplots des estimations

Conclusion et limites

Conclusion et limites

- L'estimateur QML semble être le plus performant des trois estimateurs, avec le MSE le plus faible
- Toutefois, nous suspectons que cet estimateur soit avantagé par le fait que la valeur d'initialisation soit proche de la valeur réelle
- L'estimateur contraste par intégrale a pour limite le fait que nous avons choisi des bornes arbitrairement à -100 et 100. Toutefois, les essais montraient que la valeur de la fonction à minimiser variait faiblement en fonction de la borne
- L'estimateur contraste par FFT a pour limite, le choix arbitraire de T.

Application sur les données réelles

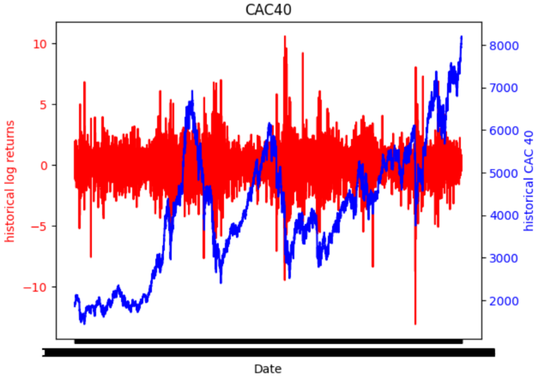

Figure – Données du cac40

Application sur les données réelles

Count	8647
Mean	1.52e-17
Std	1.35
Min	-1.31e+01
25%	-6.63e-01
50%	2.76e-02
75%	7.07e-01
Max	1.05e+01

Table – Statistiques descriptives

Application sur les données réelles

$$R_i = 100 \times \log\left(\frac{S_i}{S_{i-1}}\right) - c$$
$$Y_i = \log(R_i^2) + 1.27$$

Index	Estimations	
	ϕ	σ^2
QML	0.089	5.68
quad	0.15	3.31

Balac, Stéphane (2011). « La transformée de Fourier vue sous l'angle du calcul numérique ». In.

Chib, Siddhartha, Federico Nardari et Neil Shephard (2002). « Markov chain Monte Carlo methods for stochastic volatility models ». In : *Journal of Econometrics* 108.2, p. 281-316.

El Kolei, Salima (2013). « Parametric estimation of hidden stochastic model by contrast minimization and deconvolution : application to the Stochastic Volatility Model ». In : *Metrika* 76.8, p. 1031-1081.

Harvey, Andrew, Esther Ruiz et Neil Shephard (1994). « Multivariate stochastic variance models ». In : *The Review of Economic Studies* 61.2, p. 247-264.

Ionides, Edward L et al. (2011). « Iterated filtering ». In.

Johannes, Michael S, Nicholas G Polson et Jonathan R Stroud (2009). « Optimal filtering of jump diffusions : Extracting latent states from asset prices ». In : *The Review of Financial Studies* 22.7, p. 2759-2799.

Melino, Angelo et Stuart M Turnbull (1990). « Pricing foreign currency options with stochastic volatility ». In : *Journal of econometrics* 45.1-2, p. 239-265.

Taylor, Stephen J (2008). *Modelling financial time series*. world scientific.