Rough Volatility :Kernel estimation for a Volterra process

Mame Diarra TOURÉ

Internship supervisors : Sergio Pulido and Nicolas Brunel

From May 27 2019 to July 22 2019

ENSIIE

École Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise

1 Square de la Résistance, 91025 EVRY CEDEX

https ://www.ensiie.fr/

Table of contents

Acknowledgement

I would like to thank Mr Nicolas Brunel and Mr Sergio Pulido Nino for giving me this great opportunity to work on such a trailblazing project and for their kindness,patience and availability during this internship.
I would like to thank the whole Ensiie administration who have welcomed me into their offices.
I also would like to thank Mr Vathana Ly-Vath for agreeing to be my tutor for this internship.
Finally, I would like to thank my parents for their trust and unconditional support.

1 Abstract

Rough volatility models are very appealing because of their remarkable fit of both historical and implied volatilities. However, due to the non-Markovian and non-semimartingale nature of the volatility process, there is no simple way to simulate efficiently such models, which makes risk management of derivatives an intricate task. In this paper,We worked on rough volatility models with convolution kernels define by a sum of exponential.We determined a proper way to simulate the model and to estimate its parameters.

2 Introduction

In the derivatives world, log-prices are often modeled as continuous semi-martingales. For a given asset with log-price Y_t ,such a process takes the form

$$dY_t = \mu_t dt + \sigma_t dW_t$$

where μ_t is a drift term and W_t is a one-dimensional Brownian motion. The term σ_t denotes the volatility process and is the most important ingredient of the model. In the Black-Scholes framework, the volatility function is either constant or a deterministic function of time. On the other hand, in so-called stochastic volatility models, the volatility σ_t is modeled as a continuous Brownian semi-martingale. Notable amongst such stochastic volatility models are the Hull and White model, the Heston model, and the SABR model. Whilst stochastic volatility dynamics are more realistic than local volatility dynamics, generated option prices are not consistent with observed European option prices. In terms of the smoothness of the volatility process, the preceding models offer two possibilities : very regular sample paths in the case of Black-Scholes, and volatility trajectories with regularity close to that of Brownian motion for the local and stochastic volatility models. [5] Starting from the stylized fact that volatility is a long memory process, various authors have proposed models that allow for a wider range of regularity for the volatility. In a pioneering paper, Comte and Renault[4] proposed to model log-volatility using fractional Brownian motion (fBM for short), ensuring long memory by choosing the Hurst parameter H $>$ 1/2. Our study is based on the following rough volatility model :

$$\begin{cases} V_t = V_0 + \int_0^t K(t-s)dX_s \\ dX_s = \lambda(\theta - V_s)ds + \sqrt{V_s}dB_s \end{cases} \qquad (1)$$

Where V represent the volatility and B_s is a Brownian motion. Our aim is to estimate the kernel K which we suppose, has the following form :$K(t) = \sum_{i=1}^{m} c_i e^{-x_i t}$.Indeed,due to the introduction of the fractional kernel in the stochastic differential equation, we lose the Markovian and semi-martingale structure. In order to overcome theses difficulties, we approximate these models by simpler ones that we can use in practice.We choosed a sum of exponential because exponential does not explode (reach infininity) at 0 in opposition to $K(t) = \frac{t^{H-\frac{1}{2}}}{\Gamma(H+\frac{1}{2})}$.Relying on existence results of stochastic Volterra equations the paper [1] provide in theorem 3.1 page 7 the strong existence and uniqueness of the model under some general conditions. Thus the approximation is uniquely well-defined. We can therefore deal with simulation, pricing and hedging problems under these multi-factor models by using standard methods developed for stochastic volatility models.

In order to get acquainted with the subject, I started by working on a CIR(Cox-Ingersoll-Ross) process which is the same model than (1) only the kernel K is no longer a sum of exponential but equals to 1. Hence we have the following system

$$\left\{ dV_t = \lambda(\theta - V_t)ds + \sigma\sqrt{V_t}dB_t \right. \qquad (2)$$

I simulated and estimated the parameters(λ, θ, σ) of the process and then moved on to working on a process with a convolution kernel. After simulating the volatility sample from the model, I worked on a proper way to make sure that the simulation was right before dealing with the estimation. At first I supposed that $K(t) = ce^{-xt}$ which give us the following model

$$\begin{cases} V_t = V_0 + \int_0^t ce^{-x(t-s)} dX_s \\ dX_s = \lambda(\theta - V_s)ds + \sqrt{V_s} dB_s \end{cases} \tag{3}$$

and I've estimated the parameters x and c. Then I supposed that $K(t) = c_1 e^{-x_1 t} + c_2 e^{-x_2 t}$

$$\begin{cases} V_t = V_0 + \int_0^t (c_1 e^{-x_1(t-s)} + c_2 e^{-x_2(t-s)}) dX_s \\ dX_s = \lambda(\theta - V_s)ds + \sqrt{V_s} dB_s \end{cases} \tag{4}$$

and I estimated the set of parameters x_1, x_2, c_1, c_2.
Finally, I worked with the kernel $K(t) = c_1 e^{-x_1 t} + c_2 e^{-x_2 t} + c_3 e^{-x_3 t}$

I used the Software R for simulation and estimation purposes.

3 Study of a CIR process

3.1 Definition and parameters of a CIR process

In mathematical finance, the Cox–Ingersoll–Ross model (or CIR model) describes the evolution of interest rates. It is a type of "one factor model" (short rate model) as it describes interest rate movements as driven by only one source of market risk.
The process, initially introduced to model the short interest rate (Cox, Ingersoll and Ross), is now widely used in modelling because they present interesting features like the non-negativity and the mean reversion. Moreover, some standard expectations can be analytically calculated which can be useful especially for calibrating the parameters. Thus, they have also been used in finance to model the stochastic volatility.

The CIR model specifies that the volatility V_t follows the stochastic differential equation, also named the CIR Process :
$dV_t = \lambda(\theta - V_t)\, dt + \sigma\sqrt{V_t}\, dB_t \qquad (5).$
The parameter λ corresponds to the speed of adjustment of θ to the mean and σ to the volatility and the drift term $\lambda(\theta - V_t)$ represents a force pulling the volatility towards its long term mean θ with a speed parameter λ.

3.2 Discretization and simulation formula of a CIR process

In order to simulate samples of volatility we discretize the stochastic differential equation (5). Hence we obtain

$$V_{t_{k+1}} = V_{t_k} + \lambda(\theta - V_{t_k})\Delta t + \sigma\sqrt{V_{t_k}^{+}\Delta t}\mathcal{E}_k$$

where

$$\mathcal{E}_1, ..., \mathcal{E}_N \sim> \mathcal{N}(0,1)$$

We, then have a recurrence formulae that enables us to simulate data of volatility.

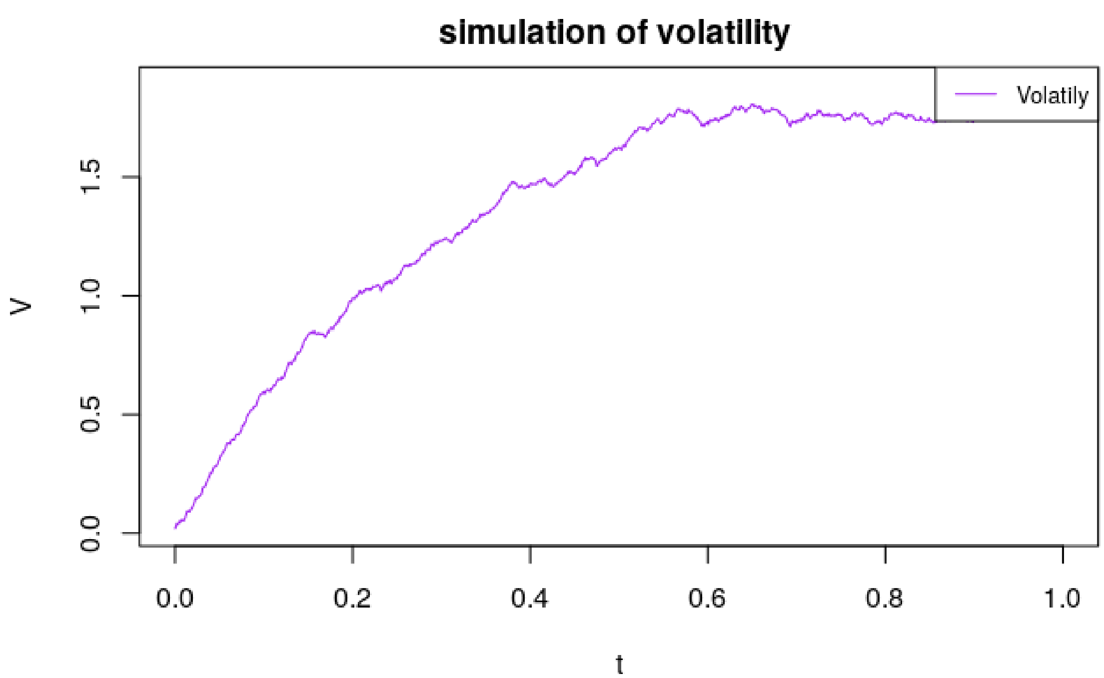

Figure 1 : Plot of CIR volatility for $\lambda = 3$, $\theta = 2$, $\sigma = 0.2$, $V_0 = 0.02$

3.3 Estimation of the parameters

3.3.1 Method 1 : The maximum likelihood estimator

When it comes to estimation the first method that comes to mind is the maximum likelihood function. In statistics, maximum likelihood estimation (MLE) is a method of estimating the parameters by maximizing a likelihood function, so that under the assumed statistical model the observed data is most probable. The point in the parameter space that maximizes the likelihood function is called the maximum likelihood estimate.The logic of maximum likelihood is both intuitive and flexible, and as such the method has become a dominant means of statistical inference. The principle of maximum likelihood provides an unified approach to estimating parameters of the distribution given sample data. Although ML estimators are not in general unbiased, they possess a number of desirable asymptotic properties :
consistency : $\lim\limits_{n \to +\infty} \hat{\theta_n} \overset{P}{=} \theta$
normality : $\hat{\theta_n} \sim> \mathcal{N}(\theta, \Sigma)$, where Σ^{-1} is the Fisher information matrix.
efficiency : $\text{Var}(\hat{\theta_n})$ approaches Cramer-Rao lower bound. In the paper [2] it is demonstrated that the MLE is suitable for such CIR process. The volatility is given by the following equation.

$$V_{t_{k+1}} = V_{t_k} + \lambda(\theta - V_{t_k})\Delta t + \sigma\sqrt{V_{t_k}^{+}\Delta t}\xi_{k+1}$$

The markovianity of the process gives us

$$P(V_{t_k}|V_0, ..., V_{t_{k-1}}) = P(V_{t_k}|V_{t_{k-1}})$$

$$P(V_k|V_{k-1}) \sim> \mathcal{N}(V_{t_{k-1}} + \lambda(\theta - V_{t_{k-1}})\Delta t, \quad \sigma^2 V^+_{t_{k-1}} \Delta t)$$

The likelihood function of a sample of length N is given by

$$L(v_{t_k}, .., v_0) =$$

$$\prod_{k=1}^{N} \frac{P(v_0)}{\sqrt{2\pi\sigma^2 v_{t_{k-1}} \Delta t}} exp(-\frac{1}{2} \frac{(v_{t_k} - (v_{t_{k-1}} + \lambda(\theta - v_{t_{k-1}})\Delta t))^2}{\sigma^2 v_{t_{k-1}} \Delta t}$$

$$L(v_{t_k}, .., v_0) =$$

$$\left(\frac{P(v_0)}{\sqrt{2\pi\sigma^2 \Delta t}}\right)^N \prod_{k=1}^{N} \frac{1}{\sqrt{v_{t_{k-1}}}} exp(-\frac{1}{2} \frac{(v_{t_k} - (v_{t_{k-1}} + \lambda(\theta - v_{t_{k-1}})\Delta t))^2}{\sigma^2 v_{t_{k-1}} \Delta t}$$

and we can also write the log-likelihood function

$$\mathcal{L}(v_{t_k}, .., v_0) =$$

$$cst + \sum_{k=1}^{N} -\frac{1}{2} ln(v_{t_{k-1}}) - \frac{1}{2\sigma^2 \Delta t} \sum_{k=1}^{N} \frac{(v_{t_k} - (v_{t_{k-1}} + \lambda(\theta - v_{t_{k-1}})\Delta t))^2}{v_{t_{k-1}}}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = -\frac{1}{2\sigma^2 \Delta t} \sum_{k=1}^{N} \frac{1}{v_{t_{k-1}}} [(-2(\theta - v_{t_{k-1}})\Delta t (v_{t_k} - (v_{t_{k-1}} + \lambda(\theta - v_{t_{k-1}})\Delta t))$$

The maximum likelihood estimate for the parameter λ
is given by

$$\lambda^* = \frac{\sum_{k=1}^{N} \frac{\theta - v_{t_{k-1}}}{v_{t_{k-1}}} (v_{t_k} - v_{t_{k-1}})}{\Delta t \sum_{k=1}^{N} \frac{(\theta - v_{t_{k-1}})^2}{v_{t_{k-1}}}} \qquad (6)$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = -\frac{1}{2\Delta t \sigma^2} \sum_{k=1}^{N} \frac{-2\Delta t \lambda}{v_{t_{k-1}}} (v_{t_k} - (v_{t_{k-1}} + \lambda(\theta - v_{t_{k-1}})\Delta t))$$

The maximum likelihood estimate for the parameter θ is given by

$$\theta^* = \frac{N(\lambda \Delta t - 1) + \sum_{k=1}^{N} \frac{v_{t_k}}{v_{t_{k-1}}}}{\Delta t \lambda \sum_{k=1}^{N} \frac{1}{v_{t_{k-1}}}} \qquad (7)$$

$$\frac{\partial \mathcal{L}}{\partial \sigma} = \frac{-N}{\sigma} + \frac{1}{\sigma^3 \Delta t} \sum_{k=1}^{N} \frac{(v_{t_k} - (v_{t_{k-1}} + \lambda(\theta - v_{t_{k-1}})\Delta t))^2}{v_{t_{k-1}}}$$

The maximum likelihood estimate for the parameter σ is given by

$$\sigma^* = \sqrt{\frac{\sum_{k=1}^{N} \frac{(v_{t_k} - (v_{t_{k-1}} + \lambda(\theta - v_{t_{k-1}})\Delta t))^2}{v_{t_{k-1}}}}{N \Delta t}} \tag{8}$$

We implemented those formulas and we've seen that the MLE is effectively working

```
[1] "the real value of the parameters are= 3 theta= 2 sigma= 0.2"
[1] "lambda mv = 2.73772374490643"
[1] "theta mv= 2.00790985366912"
[1] "sigma_mv= 0.197755180223805"
```

. Figure 2 : Result of the estimation by MLE for N=10000

```
[1] "the real value of the parameters are= 3 theta= 2 sigma= 0.2"
[1] "lambda mv = 2.91607954501534"
[1] "theta mv= 2.00157436539978"
[1] "sigma_mv= 0.200085469051264"
```

. Figure 3 : Result of the estimation by MLE for N=1.10^6
We also notice that the estimators are, indeed consistent : The bigger N is the closer the estimators are to the real values.

3.3.2 Method 2 : Estimation using the asymptotic distribution of the volatility

Due to the mean reversion(the assumption that a stock's price will tend to move to the average price over time) as time becomes large , the distribution of V(volatility) approaches a gamma distribution with the following density

$$f(v; \lambda, \theta, \sigma) = \frac{\beta^\alpha}{\Gamma(\alpha)} v^{\alpha-1} e^{-\beta r_\infty}$$

$$f(v; \lambda, \theta, \sigma) = \frac{\beta^\alpha}{\Gamma(\alpha)} v^{\alpha-1} e^{-\beta r_\infty},$$

where

$$\beta = 2\lambda/\sigma^2$$

$$\alpha = 2\lambda\theta/\sigma^2.[3]$$

Knowing that the mean of a gamma distribution and its variance, we have the following estimation

$$\hat{\theta} = mean(V)$$

, where mean(V) represents the empirical mean of the volatility V

$$\hat{\lambda} = \frac{\sigma^2\theta}{2 \times var(V)}$$

, Where var(V) represents the empirical variance of the volatility V

$$\hat{\sigma} = \sqrt{\frac{2var(V) \times \lambda}{\theta}}$$

```
[1] "the real parameters lambda= 1.2 theta= 1 sigma= 0.1"
[1] "we have the following estimations : lambda= 1.18003969038654"
[1] "theta= 1.0005429843963 sigma= 0.100842200872043"
```

Figure 4 :Result of estimation using the asymptotic distribution of Volatility

4 Simulation of rough volatility and verification

4.1 Simulation of the volatility when the kernel is approximated by a sum of exponential $K(t) = \sum_{i=1}^{M} c_i e^{-x_i t}$

$$V_t = V_0 + \int_0^t K(t-s) dX_s$$

$$dX_t = \lambda(\theta - V_t) + \sqrt{V_t} dB_t$$

$$V_t = V_0 + \sum_{k=1}^{M} c_i \int_0^t exp(-x_i(t-s)) dX_s$$

let

$$Y_t^{(i)} = \int_0^t exp(-x_i(t-s)) dX_s$$

$$dY_t^{(i)} = (dX_t - x_i Y_t^{(i)})dt = [\lambda(\theta - V_t) - x_i Y_t^{(i)}]dt + \sqrt{V_t} dB_t$$

$$dY_t^{(i)} = [\lambda(\theta - (V_0 + \sum_{k=1}^{M} c_j Y_t^{(j)} - x_i Y_t^{(i)}))]dt + \sqrt{V_0 + \sum_{k=1}^{M} c_j Y_t^{(j)} dB_t}$$

$Y_0^{(i)} = 0 \quad \forall i$

$$Y_{t_{k+1}}^{(i)} = \frac{1}{1 + x_i \Delta t}[Y_{t_k}^{(i)} + (\lambda(\theta - \sum_{k=1}^{M} c_j Y_{t_k}^{(j)} - V_0))\Delta t + \sqrt{V_0 + \sum_{k=1}^{M} c_j Y_{t_k}^{(j)} \sqrt{\Delta t}} \quad \epsilon_{k+1}]$$

$$V_{t_k} = V_0 + \sum_{k=1}^{M} c_i Y_{t_k}^{(i)}$$

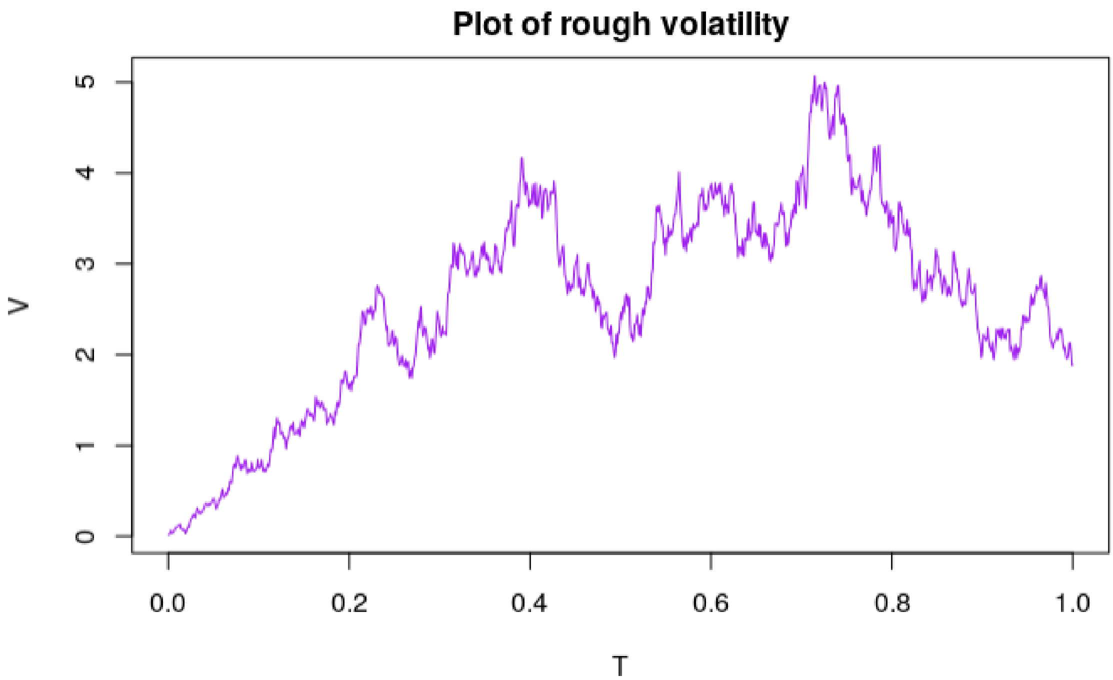

Figure 5 : Simulation of rough volatility for $\lambda = 3$ and $\theta = 2$

4.2 Verification of the simulation

We have

$$V_t = V_0 + \sum_{k=1}^{M} c_i Y_t^{(i)}$$

In order to make sure that we properly simulated the sample of rough volatility, we resolved the following differential equation satisfied by the vector mean of Y_t, $\mathbb{E}[Y_t]$. Then we compare the solution to the empirical mean of our sample.

$$\mathbb{E}[V_t] = V_0 + \sum_{i=1}^{M} c_i \mathbb{E}[Y_t^{(i)}]$$

$$Y_t^{(i)} = \int_0^t exp(-x_i(t-s))[\lambda(\theta - V_s)ds + \sqrt{V_s}dB_s]$$

$$\mathbb{E}[Y_t^{(i))}] = e^{-x_i t} \int_0^t e^{x_i s}[\lambda(\theta - V_0 - \sum_{j=1}^{M} c_i \mathbb{E}[Y_t^{(j)}])ds]$$

$$\frac{d\mathbb{E}[Y_t^{(i)}]}{dt} = -x_i E[Y_t^{(i)}] + \lambda(\theta - V_0 - \sum_{j=1}^{M} c_j \mathbb{E}[Y_t^{(j)}])]$$

$$f(t) = \begin{pmatrix} E[Y_t^{(1)}] \\ \vdots \\ E[Y_t^{(i)}] \\ \vdots \\ E[Y_t^{(M)}] \end{pmatrix}$$

$$f'(t) = \begin{pmatrix} \frac{dE[Y_t^{(1)}]}{dt} \\ \vdots \\ \frac{dE[Y_t^{(i)}]}{dt} \\ \vdots \\ \frac{dE[Y_t^{(M)}]}{dt} \end{pmatrix}$$

we suppose that $Y_t^{(i)}$ are independent and identically distributed. we obtain the following differential equation

$$f'(t) + \beta f(t) = \alpha$$

with

$$\alpha = \lambda(\theta - V_0) \begin{pmatrix} 1 \\ \vdots \\ \vdots \\ 1 \end{pmatrix}$$

and

$$\beta = diag(x_i) + \lambda \begin{pmatrix} \overrightarrow{c} \\ \vdots \\ \vdots \\ \overrightarrow{c} \end{pmatrix}$$

hence

$$f(t) = e^{-\beta t} f(0) + e^{-\beta t} \int_0^t e^{\beta s} \alpha ds$$

since

$$Y_0 = 0 \quad \forall i \quad \in \{1, ..., M\}$$

$$f(0) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\int_0^t e^{\beta s} ds = [\beta^{-1} e^{\beta s}]_0^t = \beta^{-1} e^{\beta t} - \beta^{-1}$$

$$f(t) = e^{-\beta t} \times (\beta^{-1} e^{\beta t} - \beta^{-1}) \times \alpha$$

We've simulated multiple samples of volatility and then we take the mean (for each line) in order to compare it with the solution of the differential equation. When we superimposed the plot of $f(t)$ and the plot of the mean of V we see that the two of them fit perfectly meaning that our simulation is correct
However in order to get the a good fitting we needed a high number of simulation which can mean a slow convergence rate.

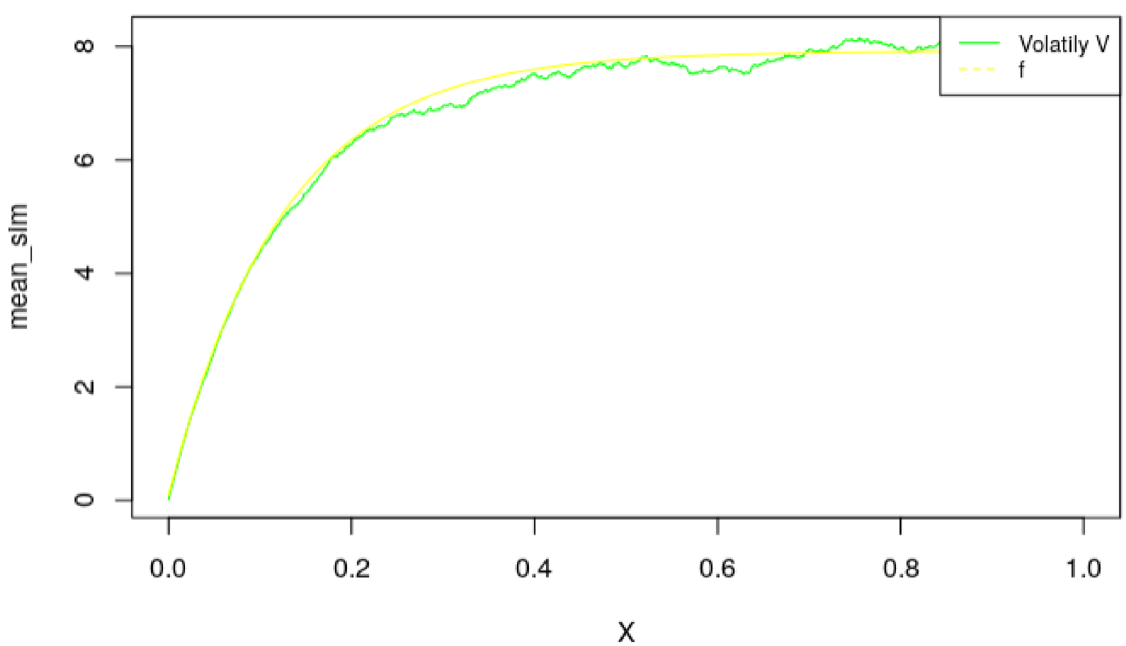

Figure 6 : Plot of mean of volatility and f for 100 simulations

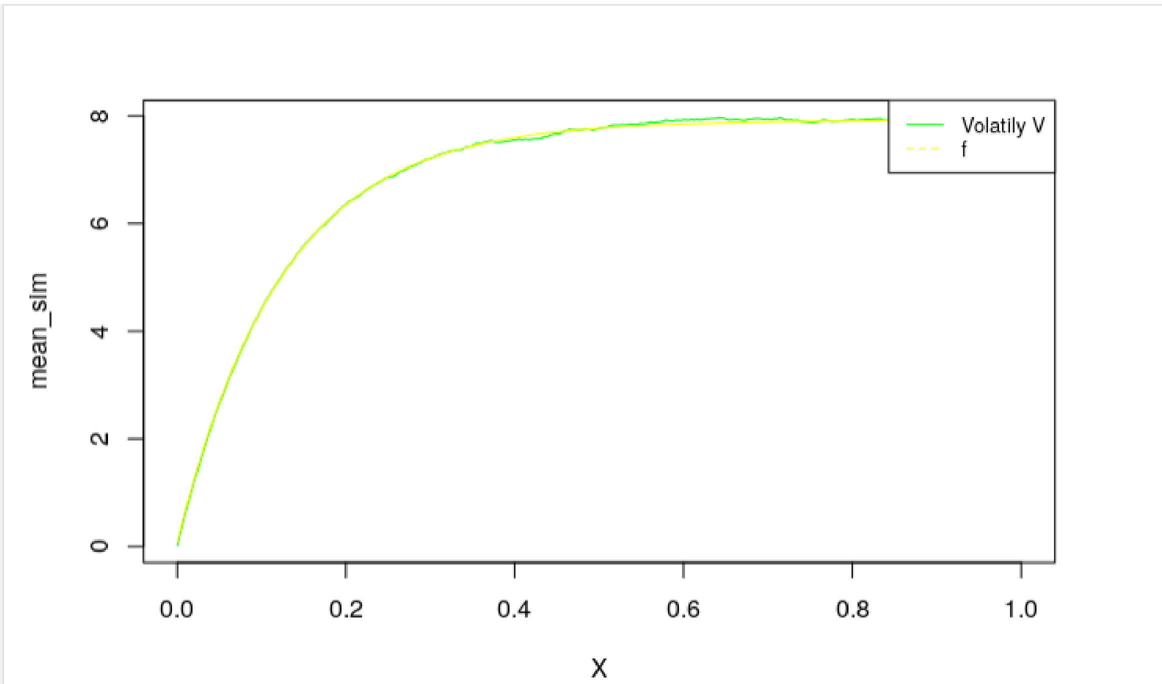

Figure 7 : Plot of mean of volatility and f for 1000 simulations

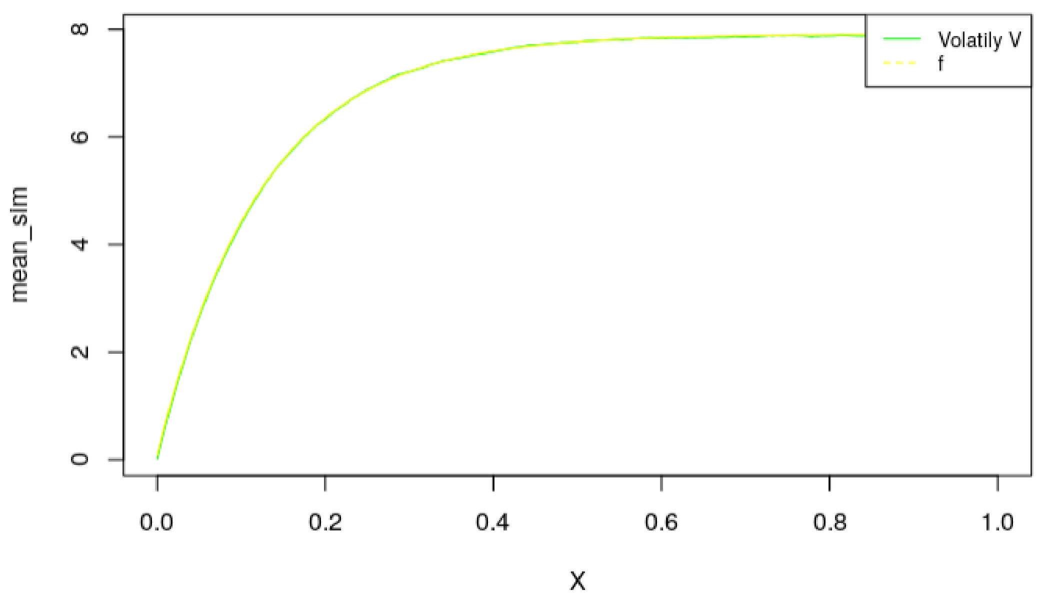

Figure 8 : Plot of mean of volatility and f for 5000 simulations

5 Study of Volatility with a kernel define by $K(t) = ce^{-xt}$

5.1 The differential equation satisfied by the volatility V and it's discretization

When the kernel K is represented by one exponential $K(t) = ce^{-xt}$, the volatility V satisfies the following differential equation :

$$V_t = V_0 + \int_0^t K(t-s)dX_s$$

$$V_t = V_0 + c\int_0^t exp(-x(t-s))dX_s$$

and

$$dX_s = \lambda(\theta - V_s)ds + \sqrt{V_s}dB_s$$

We discretize the previous equatiton and we obtain

$$V_t = V_0 + \sum_{k=0}^{t} ce^{-x(t-s_k)}dX_{s_k}$$

5.2 First attempt : estimation after deconvoluating

Let $F_k = ce^{-x(t-s_k)}$ then we got

$$V_{t_k} - V_0 = (F * dX)_k$$

where $*$ denote the convolution operator hence

$$X_t = (F^{-1} * (V_t - V_0))_t$$

where F^{-1} is the convolution inverse of F

$$F^{-1} * F = F * F^{-1} = 1$$

We Know, thanks to the paper [6] that the resolvent of the kernel K is given by

$$F^{-1}(dt) = c^{-1}(\delta_0(dt) + xdt)$$

that enable us to determine X_t

$$X_t = \frac{x}{c}\int_0^t (V_{t-s} - V_0)ds + \frac{1}{c}(V_t - V_0)$$

after discretization we get

$$X_{t_k} = \frac{x}{c}\Delta t \sum_{i=0}^{k-1}(V_{t_i} - V_0) + \frac{1}{c}(V_{t_k} - V_0)$$

$$U_{t_k} = X_{k+1} - X_{t_k} = \frac{1}{c}(V_{t_{k+1}} - V_{t_k}) + \frac{x}{c}\Delta t(V_{t_k} - V_0)$$

$$U_{t_k} = \frac{1}{c}(V_{t_{k+1}} - V_{t_k}) + \frac{x}{c}\Delta t(V_{t_k} - V_0)$$

Knowing that

$$P(U_t|V_t) --> \mathcal{N}(\lambda\Delta t(\theta - v_t), v_t\Delta t)$$

We can then write the likelihood function

$$L(u_0, ..., u_t) = \prod_{k=0}^{N-1} P(u_{t_k}|v_{t_k}) =$$

$$\prod_{k=1}^{N} \frac{1}{\sqrt{2\pi\Delta t v_{t_{k-1}}}} e^{\frac{-1}{2\Delta t v_{t_{k-1}}}((\frac{1}{c}(v_{t_k} - v_{t_{k-1}}) + \frac{x}{c}\Delta t(v_{t_k-1} - v_0)) - \lambda\Delta t(\theta - v_{t_{k-1}}))^2}$$

and the log-likelihood function

$$\mathcal{L}(u_t, ..., u_0) = \sum_{k=1}^{N} \frac{-1}{2} ln(2\pi\Delta t v_{t_{k-1}}) - \frac{1}{2\Delta t v_{t_{k-1}}}((\frac{1}{c}(v_{t_k} - v_{t_{k-1}}) +$$
$$\frac{x}{c}\Delta t(v_{t_k-1} - v_0)) - \lambda\Delta t(\theta - v_{t_{k-1}}))^2$$

We suppose that the parameters λ and θ are known and we want to estimate x and c.

First let's suppose that we know c and try to find x. In order to do that we differentiated the log-likelihood function, and find its zero

$$\frac{\partial\mathcal{L}}{\partial x} = \sum_{k=1}^{N} -\frac{1}{2\Delta t v_{t_{k-1}} c} 2\Delta t(v_{t_{k-1}} - v_0)(\frac{1}{c}(v_{t_k} - v_{t_{k-1}}) +$$

$$\frac{x}{c}\Delta t(v_{t_k-1} - v_0)) - \lambda\Delta t(\theta - v_{t_{k-1}}))$$

$$\frac{\partial \mathcal{L}}{\partial x} = -\frac{1}{c}(\sum_{k=1}^{N}\frac{1}{v_{t_{k-1}}}(v_{t_{k-1}} - v_0)((v_{t_k} - v_{t_{k-1}}))+$$

$$\frac{x}{c}\Delta t\sum_{k=1}^{N}\frac{1}{v_{t_{k-1}}}(v_{t_{k-1}} - v_0)^2 - \lambda\Delta t(\sum_{k=1}^{N}\frac{1}{v_{t_{k-1}}}(\theta - v_{t_{k-1}})(v_{t_{k-1}} - v_0)))$$

$$\frac{\partial \mathcal{L}}{\partial x} = 0 \quad ===>$$

$$(\frac{1}{c}\sum_{k=1}^{N}\frac{1}{v_{t_{k-1}}}(v_{t_{k-1}} - v_0)((v_{t_k} - v_{t_{k-1}})) - \lambda\Delta t(\sum_{k=1}^{N}\frac{1}{v_{t_{k-1}}}(\theta - v_{t_{k-1}})(v_{t_{k-1}} - v_0))+$$

$$\frac{x}{c}\Delta t\sum_{k=1}^{N}\frac{1}{v_{t_{k-1}}}(v_{t_{k-1}} - v_0)^2 = 0$$

$$x^* = \frac{\lambda c\Delta t\sum_{k=0}^{N}\frac{1}{v_{t_{k-1}}}(\theta - v_{t_{k-1}})(v_{t_{k-1}} - v_0)}{\Delta t\sum_{k=1}^{N}\frac{1}{v_{t_{k-1}}}(v_{t_{k-1}} - v_0)^2} -$$

$$\frac{(\sum_{k=1}^{N}\frac{1}{v_{t_{k-1}}}(v_{t_{k-1}} - v_0)((v_{t_k} - v_{t_{k-1}}))}{\Delta t\sum_{k=1}^{N}\frac{1}{v_{t_{k-1}}}(v_{t_{k-1}} - v_0)^2}$$

Now we suppose that we know x and we do the same method in order to find c

$$\mathcal{L}(u_t, ..., u_0) = \sum_{k=1}^{N}\frac{-1}{2}ln(2\pi\Delta t v_{t_{k-1}}) - \frac{1}{2\Delta t v_{t_{k-1}}}((\frac{1}{c}((v_{t_k} - v_{t_{k-1}})+$$

$$x\Delta t(v_{t_k-1} - v_0)) - \lambda\Delta t(\theta - v_{t_{k-1}}))^2$$

$$\frac{\partial \mathcal{L}}{\partial c} = \sum_{k=1}^{N}\frac{1}{\Delta t v_{t_{k-1}}} \times \frac{((v_{t_k} - v_{t_{k-1}}) + x\Delta t(v_{t_k-1} - v_0))}{c^2}$$

$$\times((\frac{1}{c}((v_{t_k} - v_{t_{k-1}}) + x\Delta t(v_{t_k-1} - v_0)) - \lambda\Delta t(\theta - v_{t_{k-1}}))$$

$$\frac{\partial \mathcal{L}}{\partial c} = 0 \quad ===>$$

$$\frac{1}{c}\sum_{k=1}^{N}\frac{1}{v_{t_{k-1}}}((v_{t_k}-v_{t_{k-1}})+x\Delta t(v_{t_k-1}-v_0))^2$$

$$-\lambda\Delta t\sum_{k=1}^{N}((v_{t_k}-v_{t_{k-1}})+x\Delta t(v_{t_k-1}-v_0))(\theta-v_{t_k-1})=0$$

hence

$$c^*=\frac{\sum_{k=1}^{N}\frac{1}{v_{t_{k-1}}}((v_{t_k}-v_{t_{k-1}})+x\Delta t(v_{t_k-1}-v_0))^2}{\lambda\Delta t\sum_{k=1}^{N}\frac{1}{v_{t_{k-1}}}((v_{t_k}-v_{t_{k-1}})+x\Delta t(v_{t_k-1}-v_0))(\theta-v_{t_k-1})}$$

We've found out that even this method give a good estimator for the parameter x, by deconvoluating we loose some dependance for c which result in a bad estimation of c.

Hence deconvoluationg is not an efficent way to estimate the parameters of the kernel.

5.3 Second attempt : estimation using directly the expression of the volatility V

We know that

$$V_k=V_0+\sum_{j=0}^{k-1}K(h(k-j))(\lambda h(\theta-V_j)+\sqrt{hVj}\xi_j$$

$$V_k=V_0+\sum_{j=0}^{k-1}ce^{-x(h(k-j))}(\lambda h(\theta-V_j)+\sum_{j=0}^{k-1}ce^{-x(h(k-j))}\sqrt{hV_j}\xi_j$$

$$P(V_k|V_{j<k})\sim>\mathcal{N}(V_0+\sum_{j=0}^{k-1}ce^{-x(h(k-j))}(\lambda h(\theta-V_j)),\sum_{j=0}^{k-1}ce^{-x(h(k-j))}\sqrt{hV_j})$$

let

$$\rho=e^{-x}$$

$$a_{k,j}=h(k-j)$$

$$b_j=\lambda h(\theta-V_j)$$

$$P(V_k|V_{j<k}) ===> \mathcal{N}(V_0 + \sum_{j=0}^{k-1} c\rho^{a_{k,j}} b_j, \sum_{j=0}^{k-1} c^2\rho^{2a_{k,j}} hV_j)$$

We can write the likelihood function

$$L = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi \sum_{j=0}^{k-1} c^2\rho^{2a_{k,j}} hV_j}} e^{\frac{-1}{2\sum_{j=0}^{k-1} c^2\rho^{2a_{k,j}} hV_j}(V_k - V_0 - \sum_{j=0}^{k-1} c\rho^{a_{k,j}} b_j)^2}$$

and the loglikelihood function

$$\mathcal{L} = \sum_{k=1}^{N} \frac{-1}{2} ln(2\pi \sum_{j=0}^{k-1} c^2\rho^{2a_{k,j}} hV_j) - \frac{1}{2\sum_{j=0}^{k-1} c^2\rho^{2a_{k,j}} hV_j}(V_k - V_0 - \sum_{j=0}^{k-1} c\rho^{a_{k,j}} b_j)^2$$

In this case there is no explicit formula for x^*. In order to find a good estimator, we used the function optimize of the package stats.

In opposition for the parameter c we were able to find an explicit estimator c^* by using the explicit formula or V_t

$$V_k = V_0 + \sum_{j=0}^{k-1} K(h(k-j))(\lambda h(\theta - V_j) + \sqrt{hVj}\xi_j$$

$$V_k = V_0 + \sum_{j=0}^{k-1} ce^{-x(h(k-j))}(\lambda h(\theta - V_j) + \sum_{j=0}^{k-1} ce^{-x(h(k-j))}\sqrt{hV_j}\xi_j$$

$$P(V_k|V_{k-1}) ===> \mathcal{N}(V_0 + \sum_{j=0}^{k-1} ce^{-x(h(k-j))}(\lambda h(\theta - V_j)), \sum_{j=0}^{k-1} ce^{-x(h(k-j))}\sqrt{hV_j})$$

let

$$a_k = \sum_{j=0}^{k-1} ce^{-x(h(k-j))}(\lambda h(\theta - V_j))$$

and

$$b_k = \sum_{j=0}^{k-1} (e^{-x(h(k-j))}\sqrt{hV_j})^2$$

$$P(V_k|V_{k-1}) = \mathcal{N}(a_k \times c + V_0, c^2 b_k)$$

then we can write the likelihood function

$$L(v_0, ..., v_N) = \prod_{k=0}^{N} = \frac{1}{\sqrt{2\pi c^2 b_k^2}} e^{\frac{-1}{2c^2 b_k}(V_k - V_0 - a_k c)^2}$$

$$\mathcal{L}(v_0, ..., v_N) = \sum_{k=1}^{N} -\frac{1}{2} ln(2\pi b_k) - ln(c) - \frac{-1}{2b_k}(\frac{1}{c}(V_k - V_0) - a_k)^2$$

$$\frac{\partial \mathcal{L}}{\partial c} = \sum_{k=0}^{N} -\frac{1}{c} + \sum_{k=0}^{N} \frac{1}{b_k c^2}(V_k - V_0)(\frac{1}{c}(V_k - V_0) - a_k)$$

$$\frac{\partial \mathcal{L}}{\partial c} = 0 ====>$$

$$Nc^2 + c\sum_{k=1}^{N} \frac{a_k}{b_k}(V_k - V_0) - \sum_{k=1}^{N} \frac{(V_k - V_0)^2}{b_k} = 0$$

$$c^* = \frac{-\sum_{k=1}^{N} \frac{a_k}{b_k}(V_k - V_0) \pm \sqrt{(\sum_{k=1}^{N} \frac{a_k}{b_k}(V_k - V_0))^2 + 4 \times N \times \sum_{k=1}^{N} \frac{(V_k - V_0)^2}{b_k}}}{2N}$$

```
[1] "the real value of c is  0.5"
[1] "c_mv= 0.50467678457022"
[1] "the real value of x is  3"
[1] "x_optimize= 2.90288175859326"
```

Figure 9 : Result of estimation using the MLE for c and optimize for x

This method gives a right estimation for both parameters x and c

5.4 Third attempt : Estimation by writing the process as a diffusion

We have the following kernel

$$K(t) = ce^{-xt}$$

$$dV_t = [-x(V_t - V_0) + \lambda c(\theta - V_t)]dt + c\sqrt{V_t}dB_t$$
$$V_{t_{k+1}} - V_{t_k} = [\lambda c(\theta - V_k) - x(V_k - V_0)]dt + c\sqrt{V_{t_k}dt}\xi_k$$
with $\xi --> \mathcal{N}(0,1)$

$$V_{t_{k+1}} = V_k + [\lambda c(\theta - V_k) - x(V_k - V_0)]dt + c\sqrt{V_{t_k}dt}\xi_k$$

$$P(V_{t_k}|V_{t_{k-1}}) --> \mathcal{N}([\lambda c(\theta - V_k) - x(V_k - V_0)]dt, c^2 V_{t_k}dt)$$

we can then write the likelihood function

$$L(v_0, ..., v_N) = \prod_{k=1}^{N} P(v_{t_k}|v_{t_{k-1}}) * P(v_0)$$

$$L(v_0, ..., v_N) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi c^2 v_{t_{k-1}} dt}} e^{-\frac{1}{2c^2 v_{t_{k-1}} dt}(v_{t_k} - [\lambda c(\theta - v_k) - x(v_k - v_0)]dt^2}$$

and the loglikelihood function is

$$\mathcal{L}(v_0, ..., v_N) = \sum_{k=1}^{N} \frac{-}{2} ln(2\pi c^2 v_{t_{k-1}} dt) - \frac{-1}{2c^2 v_{t_{k-1}} dt} \times$$
$$(v_{t_k} - (v_{t_{k-1}} + [\lambda c(\theta - v_{t_{k-1}}) - x(v_{t_{k-1}} - v_0)]dt))^2$$
$$\frac{\partial \mathcal{L}}{\partial x} = \frac{-1}{c^2} \sum_{k=1}^{N} \frac{1}{v_{k-1}} (v_{t_{k-1}} - v_0)(v_{t_k} - (v_{t_{k-1}} + [\lambda c(\theta - v_{t_{k-1}}) - x(v_{t_{k-1}} - v_0)]dt))$$
$$\frac{\partial \mathcal{L}}{\partial x} = 0 ==>$$

$$x^* = \frac{\lambda c dt \sum_{k=1}^{N} \frac{1}{v_{t_{k-1}}} (\theta - v_{t_{k-1}})(v_{t_{k-1}} - v_0)}{dt \sum_{k=1}^{N} \frac{1}{v_{t_{k-1}}} (v_{t_{k-1}} - v_0)^2} -$$
$$\frac{\sum_{k=1}^{N} \frac{1}{v_{t_{k-1}}} (v_{t_{k-1}} - v_0)(v_{t_k} - v_{t_{k-1}})}{dt \sum_{k=1}^{N} \frac{1}{v_{t_{k-1}}} (v_{t_{k-1}} - v_0)^2}$$

$$\mathcal{L}(v_0,...,v_N) = \sum_{k=1}^{N} \frac{-1}{2} ln(2\pi c^2 v_{t_{k-1}} dt) - \frac{-1}{2c^2 v_{t_{k-1}} dt}(v_{t_k} - (v_{t_{k-1}} +$$

$$[\lambda c(\theta - v_{t_{k-1}}) - x(v_{t_{k-1}} - v_0)]dt))^2$$

$$\mathcal{L}(v_0,...,v_N) = \sum_{k=1}^{N} -ln(c) - \frac{1}{2} ln(2\pi v_{t_{k-1}} dt) -$$

$$\frac{-1}{2v_{t_{k-1}} dt}(\frac{1}{c}((v_{t_k} - v_{t_{k-1}}) + xdt(v_{t_{k-1}} - v_0)) - \lambda dt(\theta - v_{t_{k-1}}))^2$$

$$\frac{\partial \mathcal{L}}{\partial c} = \sum_{k=1}^{N} \frac{-1}{c} + \frac{1}{2dtv_{t_{k-1}}} \times 2 \times$$

$$\frac{1}{c^2}((v_{t_k} - v_{t_{k-1}}) + xdt(v_{t_{k-1}} - v_0))(\frac{1}{c}((v_{t_k} - v_{t_{k-1}}) + xdt(v_{t_{k-1}} - v_0)) - \lambda dt(\theta - v_{t_{k-1}}))$$

$$\frac{\partial \mathcal{L}}{\partial c} = 0 ==>$$

$$\frac{-N}{c} + \frac{1}{dtc^2}(\frac{1}{c}\sum_{k=1}^{N} \frac{1}{v_{t_{k-1}}}((v_{t_k} - v_{t_{k-1}}) + xdt(v_{t_{k-1}} - v_0)) -$$

$$\lambda dt \frac{1}{v_{t_{k-1}}}(\theta - v_{t_{k-1}})((v_{t_k} - v_{t_{k-1}}) + xdt(v_{t_{k-1}} - v_0))) = 0$$

let

$$a = (v_{t_k} - v_{t_{k-1}}) + xdt(v_{t_{k-1}} - v_0))$$

$$b = (\theta - v_{t_{k-1}})$$

$$e = \frac{1}{v_{t_{k-1}}}$$

then we have

$$-Nc + \frac{1}{dt}(\frac{1}{c}\sum_{k=1}^{N} e \times a^2 - \lambda dtc \sum_{k=1}^{N} e \times a \times b) = 0$$

$$Nc^2 + \lambda c \sum_{k=1}^{N} e \times a \times b - \frac{1}{dt}\sum_{k=1}^{N} e \times a^2$$

we have a second degree equation for c hence

$$c^* = \frac{(-\lambda \sum_{k=1}^{N} e \times a \times b) \pm \sqrt{(-\lambda \sum_{k=1}^{N} e \times a \times b)^2 + (4 \times N \times \frac{1}{dt} \sum_{k=1}^{N} e \times a^2)}}{2N}$$

```
[1] "the real value of x is  3"
[1] "x_mv= 2.84245217689268"
[1] "the real value of c is  0.5"
[1] "c_mv= 0.502978710969167"
```

Figure 10 : Result of estimation using the MLE after writing the process as a diffusion

As we've seen it in the first part, this method of estimation gives good estimate for both parameters x and c.

6 Study of Volatility with a kernel define by $K(t) = c_1 e^{-x_1 t} + c_2 e^{-x_2 t}$

6.1 The differential equation satisfied by the volatility V and its discretization

$$\begin{cases} V_t = V_0 + \int_0^t (c_1 e^{-x_1(t-s)} + c_2 e^{-x_2(t-s)}) dX_s \\ dX_s = \lambda(\theta - V_s)ds + \sqrt{V_s} dB_s \end{cases}$$

let

$$Y_t^{(i)} = \int_0^t exp(-x_i(t-s))dX_s$$

$$V_t = \sum_{i=1}^{M} c_i Y_t^{(i)}$$

hence , if we differentiate the previous equation we have

$$dY_t^{(i)} = [-x_i Y_t^{(i)} + \lambda(\theta - V_t)]\Delta t + \sigma\sqrt{V_t} dB_t$$

Then we discretize it and have the following recurrence formulae.

$$Y_{t_{k+1}}^{(i)} - Y_{t_k}^{(i)} = [-x_i Y_{t_{k+1}}^{(i)} + \lambda(\theta - V_{t_k})]\Delta t + \sigma\sqrt{V_{t_k}\Delta t}\xi_k$$

$$Y_{t_{k+1}}^{(i)} = \frac{1}{1 + x_i \Delta t}[Y_{t_k}^{(i)} + \lambda(\theta - V_{t_k})]\Delta t + \sigma\sqrt{V_{t_k}\Delta t}\xi_k]$$

$$Y_{t_k}^{(i)} = \sum_{j=0}^{k-1} \frac{1}{(1 + x_i \Delta t)^{(k-j)}} (\lambda(\theta - V_{t_j}))\Delta t + \sigma\sqrt{V_{t_j}\Delta t}\xi_j$$

$$V_{t_k} = \sum_{i=1}^{M} c_i Y_{t_k}^{(i)}$$

$$V_{t_k} = \sum_{i=1}^{M} c_i (\sum_{j=0}^{k-1} \frac{1}{(1 + x_i \Delta t)^{(k-j)}} (\lambda(\theta - V_{t_j}))\Delta t + \sigma\sqrt{V_{t_j}\Delta t}\xi_j)$$

Thanks to Fubini's theorem we can write

$$V_{t_k} = (\sum_{j=0}^{k-1} [\sum_{i=1}^{M} \frac{c_i}{(1 + x_i \Delta t)^{(k-j)}}](\lambda(\theta - V_{t_j}))\Delta t + \sigma\sqrt{V_{t_j}\Delta t}\xi_j)$$

$$\alpha_k = \sum_{j=0}^{k-1} \frac{1}{1 + x_1 \Delta t} (\lambda dt(\theta - V_j))$$

$$\beta_k = \sum_{j=0}^{k-1} \frac{1}{1 + x_2 \Delta t} (\lambda dt(\theta - V_j)$$

$$\gamma_{k,j} = \frac{1}{(1 + x_1 \Delta t)^{(k-j)}}$$

$$\delta_{k,j} = \frac{1}{(1 + x_2 \Delta t)^{k-j}}$$

$$V_k = V_0 + c_1 \alpha_k + c_2 \beta_k + \sum_{j=0}^{k-1} (c_1 \gamma_{k,j} + c_2 \delta_{k,j}) \sqrt{dt V_j} \xi_j$$

$$P(V_k | V_{j<k}) = \mathcal{N}(V_0 + c_1 \alpha_k + c_2 \beta_k, \sum_{j=0}^{k-1} (c_1 \gamma_{k,j} + c_2 \delta_{k,j})^2 dt V_j)$$

6.2 estimation of the kernel's parameters

Thanks to the previous formulae,we can write the likelihood function

$$L = \prod_{k=1}^{N} P(V_k | V_{j<k}) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi dt \sum_{j=0}^{k-1} (c_1 \gamma_{k,j} + c_2 \delta_{k,j})^2 V_j}} e^{\frac{-(V_k - V_0 - c_1 \alpha_k - c_2 \beta_k)^2}{2dt \sum_{j=0}^{k-1} (c_1 \gamma_{k,j} + c_2 \delta_{k,j})^2 V_j}}$$

and the log-likelihood function

$$\mathcal{L} = -\frac{-1}{2} [\sum_{k=1}^{N} ln(\sum_{j=0}^{k-1} ((c_1 \gamma_{k,j} + c_2 \delta_{k,j})^2 V_j) + ln(2\pi dt) + \frac{(V_k - V_0 - c_1 \alpha_k - c_2 \beta_k)^2}{dt \sum_{j=0}^{k-1} (c_1 \gamma_{k,j} + c_2 \delta_{k,j})^2 V_j}]$$

6.2.1 Estimation of x_1 and x_2

In order to estimate x_1 and x_2 we tried to optimize the log-likelihood function, however for unknown reasons that didn't work.

````
optim(c(3,15),lvx,method = "L-BFGS-B",lower = c(4,14),upper = c(7,26))
```

$par
[1]  4 14

$value
[1] -3486.893

$counts
function gradient
       2        2

$convergence
[1] 0

$message
[1] "CONVERGENCE: NORM OF PROJECTED GRADIENT <= PGTOL"
````

Figure 11 : Result of optim for the estimation of x_1 and x_2

That's why we've tried to find "manually" the maximum likelihood estimate Indeed when we plot the log-likelihood as a function of $x_1$and x_2 we see that there are multiple couples (x_1,x_2) for which the likelihood is maximal
````

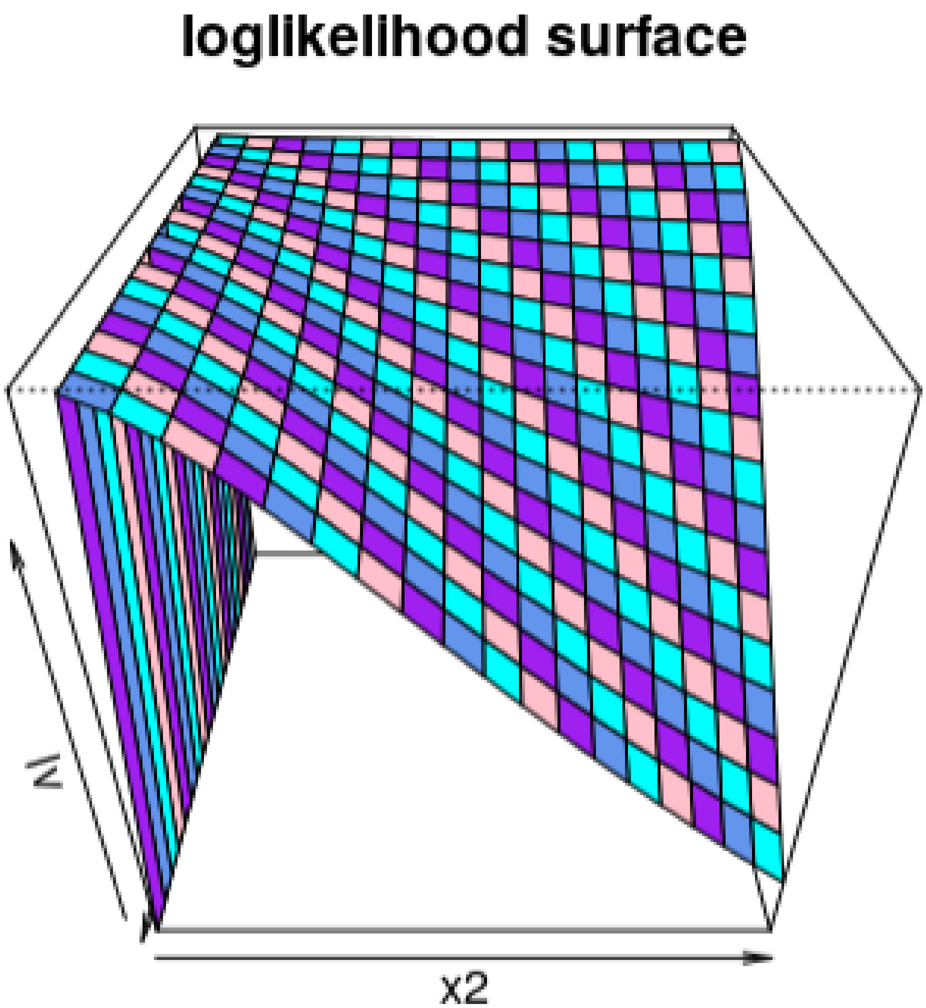

Figure 12 : Plot of log likelihood as a function of x_1 and x_2

If we calculate a random number of values for likelihood , the value of x_1 and x_2 that correspond to the maximum are not the same as the real value. However, when we plot the two kernels (the one with the real parameters and the one with the estimated parameters) we see that they are close

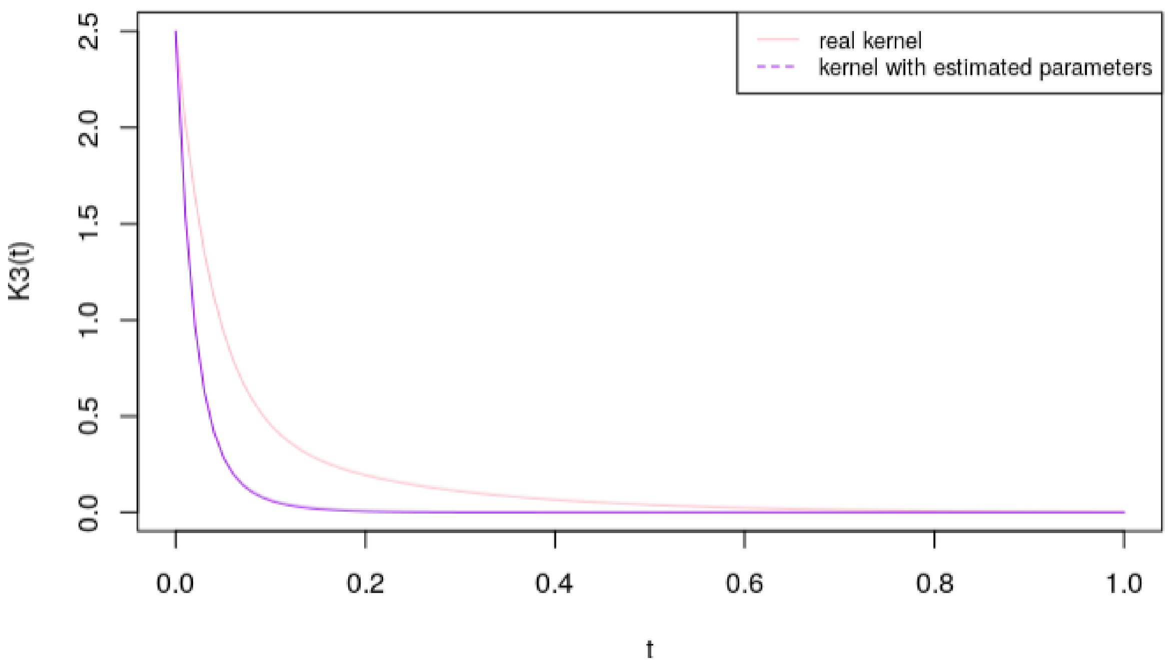

Figure 13 : Plot of the real and estimated Kernel

6.2.2 Estimation of c_1 and c_2

First we have plotted the log likelihood as a function of c_1 and c_2 to see if there was a maximum

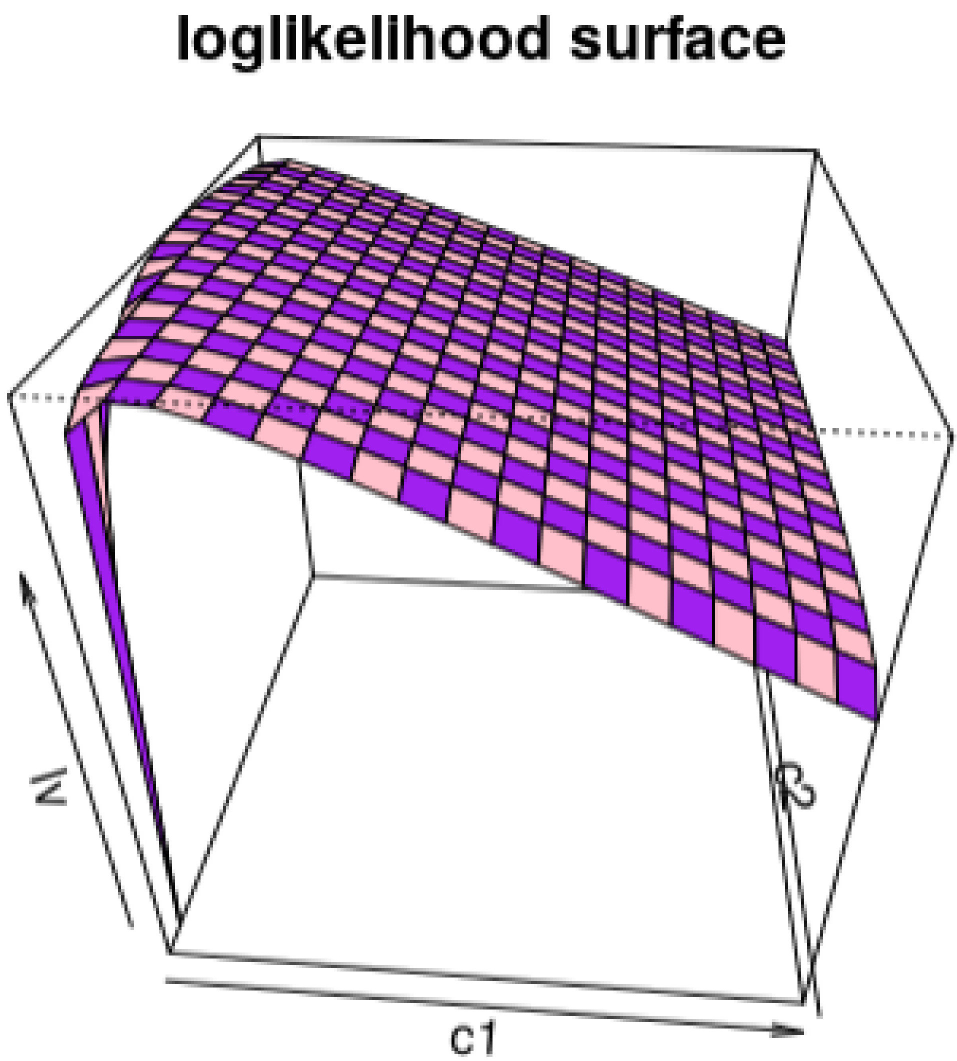

Figure 14 : Plot of the loglikelihood as a function of c_1 and c_2

In order to estimate c_1 and c_2 we tried to optimize the log-likelihood function, using the optim function and the method L-BFGS-B. It gave us a good approxiamtion of c_1 and c_2

```
optim(c(0.1,0.5),lv6,method = "L-BFGS-B",lower = c(0.05,0.1),upper = c(0.4,2))
```

```
[1] "the real value of c1 is 0.25 and the real value of c2 is 1"
$par
[1] 0.2324669 1.0646453

$value
[1] 49.04567

$counts
function gradient
      18       18

$convergence
[1] 0
```

Figure 15 :result of optim for c_1 and c_2

7 Study of Volatility with a kernel define by $K(t) = c_1 e^{-x_1 t} + c_2 e^{-x_2 t} + c_3 e^{-x_3 t}$

7.1 The differential equation satisfied by the volatility V and its discretization

$$\begin{cases} V_t = V_0 + \int_0^t (c_1 e^{-x_1(t-s)} + c_2 e^{-x_2(t-s)} + c_3 e^{-x_3 t}) dX_s \\ dX_s = \lambda(\theta - V_s)ds + \sqrt{V_s} dB_s \end{cases}$$

let

$$Y_t^{(i)} = \int_0^t exp(-x_i(t-s))dX_s$$

$$V_t = \sum_{i=1}^{M} c_i Y_t^{(i)}$$

We discretize theprevious equation and we obtain

$$dY_t^{(i)} = [-x_i Y_t^{(i)} + \lambda(\theta - V_t)]\Delta t + \sigma\sqrt{V_t} dB_t$$

Then we discretize it and have the following recurrence formulae.

$$Y_{t_{k+1}}^{(i)} - Y_{t_k}^{(i)} = [-x_i Y_{t_{k+1}}^{(i)} + \lambda(\theta - V_{t_k})]\Delta t + \sigma\sqrt{V_{t_k}\Delta t}\xi_k$$

$$Y^{(i)}_{t_{k+1}} = \frac{1}{1+x_i\Delta t}[Y^{(i)}_{t_k} + \lambda(\theta - V_{t_k})]\Delta t + \sigma\sqrt{V_{t_k}\Delta t}\xi_k]$$

$$Y^{(i)}_{t_k} = \sum_{j=0}^{k-1} \frac{1}{(1+x_i\Delta t)^{(k-j)}}(\lambda(\theta - V_{t_j}))\Delta t + \sigma\sqrt{V_{t_j}\Delta t}\xi_j$$

$$V_{t_k} = \sum_{i=1}^{M} c_i Y^{(i)}_{t_k}$$

$$V_{t_k} = \sum_{i=1}^{M} c_i(\sum_{j=0}^{k-1} \frac{1}{(1+x_i\Delta t)^{(k-j)}}(\lambda(\theta - V_{t_j}))\Delta t + \sigma\sqrt{V_{t_j}\Delta t}\xi_j)$$

Thanks to Fubini's theorem we can write

$$V_{t_k} = (\sum_{j=0}^{k-1}[\sum_{i=1}^{M} \frac{c_i}{(1+x_i\Delta t)^{(k-j)}}](\lambda(\theta - V_{t_j}))\Delta t + \sigma\sqrt{V_{t_j}\Delta t}\xi_j)$$

$$\alpha_k = \sum_{j=0}^{k-1} \frac{1}{(1+x_1\Delta t)^{k-j}}(\lambda dt(\theta - V_j))$$

$$\beta_k = \sum_{j=0}^{k-1} \frac{1}{(1+x_2\Delta t)^{k-j}}(\lambda dt(\theta - V_j)$$

$$\mu_k = \sum_{j=0}^{k-1} \frac{1}{(1+x_3\Delta t)^{k-j}}(\lambda dt(\theta - V_j)$$

$$\gamma_{k,j} = \frac{1}{(1+x_1\Delta t)^{(k-j)}}$$

$$\delta_{k,j} = \frac{1}{(1+x_2\Delta t)^{k-j}}$$

$$\nu_{k,j} = \frac{1}{(1+x_3\Delta t)^{k-j}}$$

$$V_k = V_0 + c_1\alpha_k + c_2\beta_k + c_3\mu_k \sum_{j=0}^{k-1}(c_1\gamma_{k,j} + c_2\delta_{k,j} + c_3\nu_{k,j})\sqrt{dtV_j}\xi_j$$

$$P(V_k|V_{j<k}) = \mathcal{N}(V_0 + c_1\alpha_k + c_2\beta_k + c_3\mu_k, \sum_{j=0}^{k-1}(c_1\gamma_{k,j} + c_2\delta_{k,j} + c_3\nu_{k,j})^2 dtV_j)$$

7.1.1 Estimation of c_1, c_2 and c_3

Thanks to the previous formulae,we can write the likelihood function

$$L = \prod_{k=1}^{N} P(V_k|V_{j<k}) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi dt \sum_{j=0}^{k-1}(c_1\gamma_{k,j} + c_2\delta_{k,j} + c_3\nu_{k,j})^2 V_j}} \times$$

$$\exp\left(\frac{-(V_k - V_0 - c_1\alpha_k - c_2\beta_k - c_3\mu_k)^2}{2dt \sum_{j=0}^{k-1}(c_1\gamma_{k,j} + c_2\delta_{k,j} + c_3\nu_{k,j})^2 V_j}\right)$$

and the log-likelihood function

$$\mathcal{L} = -\frac{-1}{2}[\sum_{k=1}^{N} ln(\sum_{j=0}^{k-1}((c_1\gamma_{k,j} + c_2\delta_{k,j} + c_3\nu_{k,j})^2 V_j) + ln(2\pi dt) +$$

$$\frac{(V_k - V_0 - c_1\alpha_k - c_2\beta_k - c_3\mu_k)^2}{dt \sum_{j=0}^{k-1}(c_1\gamma_{k,j} + c_2\delta_{k,j} + c_3\nu_{k,j})^2 V_j}]$$

we used optim in order to find the estimators of c_1, c_2 and c_3.

```
[1] "the real parameters are  0.333333333333333 2.66666666666667 9"
$par
[1]  0.2950429  3.8041339 15.4472687

$value
[1] 1451.957

$counts
function gradient
      21       21

$convergence
[1] 0

$message
[1] "CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH"
```

Figure 16 : Result of optim for the estimation of c_1, c_2 and c_3

Despite the fact that optim does not give the right values of the parameters we have found out that the plot of the kernel with the estimated parameters is the same as the plot with the real parameters.

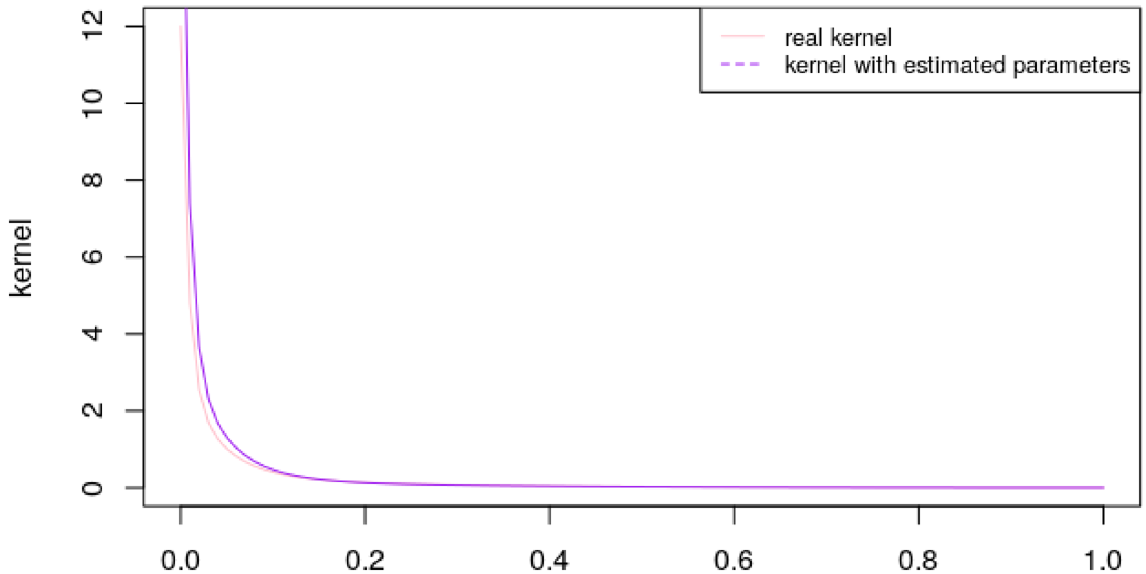

Figure 17 : Plot of the real and the estimated Kernel

As we can see it on the picture the two plots superimposed perfectly which can mean that even if we can not estimate the parameters one by one we can estimate the whole kernel. It can also mean that there is an identifiability problem.

8 conclusion

As a first year student at the national school of computer science for industry and business, I was asked to do an internship in order to put into practice the knowledge I have acquired during my year of formation. Since it is my ambition to pursue my studies toward a PhD, I welcomed the opportunity to work on the research subject Rough Volatility :Kernel estimation for a Volterra process. The simulation and the estimation of convolution kernels for rough volatility purposes was not an easy task. However we managed to find a way to estimate such kernels not by finding an estimator to each parameter but by estimating the whole function. After trying out many methods such as deconvolution, we've arrived to the conclusion that the proper way to such estimation is to use the decomposition othe volatility V in terms of Y_t.

9 Appendix 1 : Finding the resolvent of a convolution kernel

We have the following kernel

$$K(t) = c_1 e^{-\lambda_1 t} + c_2 e^{-\lambda_2 t}$$

we know that the resolvent will have this form

$$L(ds) = \alpha \delta_0(ds) + l(s)ds$$

We want to find α and $l(t)$ Knowing that the kernet K and the resolvent satisfy this equation

$$1 = \int_0^t K(t-s)L(ds)ds$$

hence

$$1 = \alpha K(t) + \int_0^t K(t-s)l(s)ds$$

We can find easy the value of alpha by evaluating the expresion $\alpha K(t) + \int_0^t K(t-s)l(s)ds$ at 0 $\boxed{\alpha = \frac{1}{K(0)}}$ Now let's find $l(t)$

$$1 = \alpha K(t) + \int_0^t K(t-s)l(s)ds$$

$$1 = \alpha K(t) + c_1 \int_0^t e^{-\lambda_1 (t-s)} l(s)ds + c_2 \int_0^t e^{-\lambda_2 (t-s)} l(s)ds$$

We derivatge it

$$0 = \alpha K'(t) - \lambda_1 c_1 \int_0^t e^{-\lambda_1 (t-s)} l(s)ds + c_1 l(t) - \lambda_2 c_2 \int_0^t e^{-\lambda_2 (t-s)} l(s)ds + c_2 l(t)$$

$$0 = (c_1 + c_2)l(t) + \alpha K'(t) - \lambda_1 c_1 \int_0^t e^{-\lambda_1 (t-s)} l(s)ds - \lambda_2 c_2 \int_0^t e^{-\lambda_2 (t-s)} l(s)ds$$

from that expression we can determine the initial value of l

$$l(0) = -\frac{\alpha K'(0)}{c_1 + c_2} = \frac{\lambda_1 c_1 + \lambda_2 c_2}{(c_1 + c_2)^2}$$

we replace

$$c_2 \int_0^t e^{-\lambda_2 (t-s)} l(s)ds$$

by

$$1 - \alpha K(t) - c_1 \int_0^t e^{-\lambda_1(t-s)} l(s) ds$$

$$0 = (c_1 + c_2)l(t) + \alpha K'(t) - \lambda_1 c_1 \int_0^t e^{-\lambda_1(t-s)} l(s) ds + c_1 l(t) - \lambda_2 (1 - \alpha K(t) - c_1 \int_0^t e^{-\lambda_1(t-s)} l(s) ds)$$

$$0 = (c_1 + c_2)l(t) + \alpha K'(t) - c_1(\lambda_1 + \lambda_2) \int_0^t e^{-\lambda_1(t-s)} l(s) ds - \lambda_2 + \lambda_2 \alpha K(t)$$

We derivate again

$$0 = (c_1 + c_2)l'(t) + \alpha K"(t) + \lambda_2 \alpha K'(t) + \lambda_1 c_1(\lambda_1 + \lambda_2) \int_0^t e^{-\lambda_1(t-s)} l(s) ds - c_1(\lambda_1 + \lambda_2) l(t)$$

We replace

$$c_1(\lambda_1 + \lambda_2) \int_0^t e^{-\lambda_1(t-s)} l(s) ds$$

by

$$(c_1 + c_2)l(t) + \alpha K'(t) - \lambda_2 + \lambda_2 \alpha K(t)$$

$$0 = (c_1 + c_2)l'(t) + \alpha K"(t) + \lambda_2 \alpha K'(t) + \lambda_1((c_1 + c_2)l(t) + \alpha K'(t) - \lambda_2 + \lambda_2 \alpha K(t)) - c_1(\lambda_1 + \lambda_2) l(t)$$

$$0 = (\lambda_1 c_2 + \lambda_2 c_1) l(t) + (c_1 + c_2) l'(t) + \alpha K"(t) + \alpha(\lambda_1 + \lambda_2) K'(t) + \alpha \lambda_1 \lambda_2 K(t) - \lambda_1 \lambda_2$$

$$l'(t) + \frac{(\lambda_1 c_2 + \lambda_2 c_1)}{c_1 + c_2} l(t) = -\frac{1}{c_1 + c_2} (\alpha K"(t) + \alpha(\lambda_1 + \lambda_2) K'(t) + \alpha \lambda_1 \lambda_2 K(t) - \lambda_1 \lambda_2)$$

$$K'(t) = -\lambda_1 c_1 e^{-\lambda_1 t} - \lambda_1 c_2 e^{-\lambda_2 t}$$

$$K"(t) = \lambda_1^2 c_1 e^{-\lambda_1 t} + \lambda_2^2 c_2 e^{-\lambda_2 t}$$

after replacing $K'(t)$ and $K"(t)$ by their expression we obtain the following differential equation satisfied by $l(t)$

$$l'(t) + \frac{(\lambda_1 c_2 + \lambda_2 c_1)}{c_1 + c_2} l(t) = \frac{\lambda_1 \lambda_2}{c_1 + c_2}$$

let

$$a = \frac{(\lambda_1 c_2 + \lambda_2 c_1)}{c_1 + c_2}$$

and

$$b = \frac{\lambda_1 \lambda_2}{c_1 + c_2}$$

then

$$l'(t) + a l(t) = b$$

hence $\boxed{l(t) = \frac{b}{a} + k_1 e^{-at}}$ with

$$k_1 = l(0) - \frac{b}{a}$$

Now that we've determine α and $l(t)$ we have the expression of $L(ds)$ hence

$$\boxed{L(ds) = \frac{1}{c_1+c_2}\delta_0(ds) + (\frac{b}{a} + k_1 e^{-as})ds}$$

10 Appendix 2 : sustainable development

10.1 Strategy and governance

Appointment of a Chargé de Mission DD RS in February 2019
Appointment of a gender equality referent in December 2018
Adoption of a strategy document and school SD RS charter in May 2019

10.2 Education and training

Addition of an appendix DD RS in the reports of internship of initial trainings in 2018-2019
Three Environmental Awareness Conferences for the first and second years, and a course on the eco-design of digital services in the first year (in 2018-2019) Creation of the Charter of the eco-citizen of the school for students early 2019
Awareness of eco-gestures by the association Ecologiie as part of the student week of sustainable development in April 2019
Training in digital eco-gestures for staff in April 2019 by the Point de MIR association.
Exhibitions on Sustainable Development : "10 clichés (disassembled) on computing and the environment" of the LIMSI CNRS / Paris Saclay University in November 2018, and "Change the system, not the climate" by the Maison des Citoyens du Monde Nantes in May 2019

10.3 Environmental Management

REFEDD carbon footprint training by Avenir Climatique in May 2019, in association with IMT Business School , for students from both schools
Establishment of selective sorting of glass in May 2019

10.4 Social policy and territorial anchoring

Events for the International Women's Rights Day of March 8, 2019 : double photo exhibition, lunch for women
Internal communication on the non-stereotypical communication guide of the High Council for Equality between Women and Men at the beginning of 2019 and adaptation for ENSIIE
Work on the communication documents of ENSIIE for a non-stereotyped communication : website directory, job offers of the website in early 2019
Launch of the Computer Science Project at the end of 2018 in collaboration between students and

management
Participation of the students of the IT project at the Feminine at the event " They innovate for digital " in Ile-de-France in April 2019
Participation in the gender equality barometer of the CGE in May 2019
Provision of documents concerning the rights and duties of maternity, paternity, parental leave, part-time work and violence and harassment for school staff in April 2019
Parity of external members of the Board since September 2018
Integration projects led by Colombbus to give a training of web developer or integrator to people remote from employment in 2018-2019
Junior Digital Academy to introduce children to digital-based creative practices in November 2018
Declick introductory workshops on programming by students in 4 junior high schools of Évry in 2018-2019

Références

[1] Eduardo Abi Jaber and Omar El Euch. Multifactor approximation of rough volatility models. *SIAM Journal on Financial Mathematics*, 10(2) :309–349, 2019.

[2] Mohamed Ben Alaya and Ahmed Kebaier. Parameter estimation for the square-root diffusions : ergodic and nonergodic cases. *Stochastic Models*, 28(4) :609–634, 2012.

[3] Sidi Mohamed Ould Aly. Asymptotic behavior for the joint distribution of the cir process and its integral. Technical report, 2012.

[4] Fabienne Comte and Eric Renault. Long memory in continuous-time stochastic volatility models. *Mathematical finance*, 8(4) :291–323, 1998.

[5] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6) :933–949, 2018.

[6] Eduardo Abi Jaber, Martin Larsson, and Sergio Pulido. Affine volterra processes. *arXiv preprint arXiv :1708.08796*, 2017.