Statistical estimation for common discrete diffusion processes

Salamata Ba[†] and Adja Deguene Gueye[‡]

Project advisor: Alderic Joulin

Abstract.

Diffusion processes are essential to finance for the modeling of stock market prices. They are the solution to a stochastic differential equation. Discretizing this solution by Euler scheme facilitates its numerical simulation for the prediction of future stock prices. Some stochastic differential equations found in the financial field include the well-known Black-Scholes model, the Cox-Ingersoll-Ross model, or the Ornstein-Uhlenbeck process. After solving a series of these stochastic differential equations when the solution exists, we would like to estimate each unknown parameter in the formula. However, the diffusion processes are not independent and identically distributed (i.i.d). In this paper, we will focus on the statistical estimation of non-i.i.d samples, which is the case of Markov chains.
Here we show that we can find these parameters through maximum likelihood estimation by writing the conditional distribution of the diffusion process for two successive periods. We then see that these estimated parameters can be expressed with a martingale. Therefore, the consistency and asymptotic normality of the maximum likelihood estimator are dependent on the martingale's behavior. In particular, the discretized version of the Ornstein-Uhlenbeck process gives the same solution as a homogeneous Markov Chain.
We also extend our analysis to continuous diffusion processes, without the upstream discretization work. We use more complex results like the Girsanov theorem and the Ito lemma and rethink our numerical simulation strategy.

1. Introduction.

1.1. Setting the context in Finance.

Continuous-time stochastic differential equations are used to solve various economic and financial issues such as asset pricing or portfolio selection. As pointed by Sundaresan (2000), diffusion processes have been leaving their mark on the financial and economic field throughout the last half-century. Their theorizing by Merton in the late 1960s established groundwork for option pricing in finance. Since the 1980s, these processes have been extended in order to allow parameter estimation and implementation using computational methods.

1.2. Presentation of the different processes and their use.

Continuous-time stochastic processes are often used to describe the time evolution of several models in finance. Given an issue, once the best model with good properties is chosen, finding the parameter's values which best fit the model on real market data is needed for future use such as making prediction. The results in this paper revolve around the parametric

[†]Student at the department of Applied Mathematics, National Institute of Applied Sciences(INSA), Toulouse, FR (sba@etud.insa-toulouse.fr,salamata68@gmail.com, https://www.linkedin.com/in/salamata-ba/).

[‡]Student at the department of Applied Mathematics, National Institute of Applied Sciences(INSA), Toulouse, FR (adgueye@etud.insa-toulouse.fr, adjadeguene@gmail.com,https://www.linkedin.com/in/adja-deguene-gueye/).

estimation of some diffusion processes in finance on the basis of their trajectory over a time interval. With this aim in mind, we will use the method of maximum likelihood estimation and we will draw a particular attention to the study of the estimators consistency.
In this paper, we have studied the Black-Scholes model, used to represent asset prices for a financial market containing derivative investment instruments. This model gives a theoretical method to calculate whether a European-style option has a fair value or not.
To describe the evolution of interest rates, the Ornstein-Uhlenbeck process is one of the most common approaches. It is particularly interesting on a statistical viewpoint because all the calculations are explicit (which is uncommon for a continuous-time stochastic process).

Furthermore, we have explored the Cox-Ingersoll-Ross model, an extension of the previous process which handles the problem of negative rates.

1.3. What did we do?.

Since, in practice, data used for modelling these processes are of a discrete nature, the Euler discretization scheme approach was employed here to approximate the continuous-time model. The maximum likelihood estimators are then calculated using this discrete-time processus. Otherwise, the use of Martingale theory, in particular the law of large numbers, has been very helpful to our study especially to determine the properties of these estimators.
Finally, to evaluate our statistical approach, simulations are done in order to illustrate the convergence of the theoretic estimators on some synthetic examples.

1.4. Definitions and theorems we are going to use.

`The Brownian motion`

While studying the movement of a fluid found in pollen grains in 1827, the botanist Robert Brown realized the liquid followed a continuous but very irregular trajectory. In fact, he theorized it in the 1918s as the random motion of a big particle immersed in a fluid and under the sole influence of the small neighboring molecules of the fluid. The Brownian motion will also be known as the Wiener process after the mathematician Norbert Wiener.
The Wiener process is often used in option pricing where the option price can be seen as a particle under constant bombing by smaller particles (illustrated by options exchanges or other local events). If we view the option price variations as the result of a big number of events considered globally independent and identically distributed, we can use a normal distribution to simulate them.

Definition 1.1. *(Brownian Motion)*

A Brownian Motion is a stochastic process $(B_t)_{t\geq 0}$ *such that:*
- $B_0 = 0$ *a.s*

- For any $0 \leq s \leq t$, the increments are stationary:

$$B_t - B_s \sim \mathcal{N}(0,\, t-s), B_{t-s} \text{ as well}$$

- For any $0 = t_0 < t_1 < \cdots < t_d$, the increments are independent i.e $(B_{t_i} - B_{t_{i-1}})_{1 \leq i \leq d}$ are independent.

(Properties)
A Brownian Motion is a Gaussian process.
It is characterized by its mean and its covariance:

$$\mathbb{E}[B_t] = 0,\; Cov(B_s, B_t) = min(s,t);\, \forall s, t \geq 0$$

The Brownian Motion is also a martingale. We can prove that its quadratic variation is:

$$[B, B]_t = t, \forall t > 0.$$

`The Ito lemma`

The Itô formula is one of the most important results of stochastic calculus.
It was introduced by japanese mathematician Kiyosi Itô in a series of papers published between 1942 and 1950.
The Itô lemma is defined as follows:

Lemma 1.2. *(The Ito lemma)*

Let M be a martingale and $f : \mathbb{R} \longrightarrow \mathbb{R}$ a function of class $\mathcal{C}^2$.
According to the Itô lemma, $\forall t \geq 0$,

$$f(M_t) - f(M_0) = \int_0^t f'(M_s)\mathrm{d}M_s + \frac{1}{2}\int_0^t f''(M_s)\mathrm{d}[M, M]_s$$

`The Euler scheme`

When the transition densities of the continuous system are not explicitly known, the maximum likelihood estimation starts with an approximation of the process using an Euler scheme. The Euler approximation corresponds to a simple discretization of the continuous time stochastic equation.

Let's assume that the process to approach is a solution of the equation:

$$X_t = x_0 + \int_0^t b(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\mathrm{d}B_s$$

or in a differential form:

$$\mathrm{d}X_t = b(t, X_t)dt + \sigma(t, X_t)\mathrm{d}Bt,\;\, X_0 = x_0$$

where B is a Brownian Motion. We consider the Euler scheme with time step $h = T/N$ i.e $t_i = ih$ are the discretization times in an interval $[0, T]$.

Definition 1.3. *(Euler scheme discretization)*

The approximation of the process, $(X_{t_i}^{(n)})_i$ *is defined as follow:*

$$X_{t_i}^{(n)} - X_{t_{i-1}}^{(n)} = (t_i - t_{i-1})\, b(t_{i-1}, X_{t_{i-1}}^{(n)}) + \sigma(t_{i-1}, X_{t_{i-1}}^{(n)})(B_{t_i} - B_{t_{i-1}}) \tag{1.1}$$

`Maximum Likelihood Estimator`

To estimate the parameters of the continous-time process, we consider the discrete model obtained by using the Euler scheme: Definition 1.3. Let's denote $(X_t)_t$ the process and assume that it is Markovian. The likelihood function $L_n(\theta)$, which is equal to the joint density of $(X_1, \ldots, X_n)$, can be derived from:

Definition 1.4.

$$L_n(\theta; x) = f^{\theta}_{X_1, X_2, \ldots, X_n}(x) \neq \prod_{i=1}^{n} f^{\theta}_{X_i}(x)$$

Since the values of $(X_n)_n$ are not independent, the joint density is not equal to the product of the marginal densities. However, given the Markovian nature of the process, we have:

Definition 1.5.

$$L_n(\theta; x) = \prod_{i=1}^{n} f^{\theta}_{X_i}(x_i / X_{i-1} = x_{i-1}) \tag{1.2}$$

where $f^{\theta}_{X_i}(x_i / X_{i-1} = x_{i-1})$ is the transition density function.
The logarithm of the likelihood function is given by:

$$log L_n(\theta; x) = \sum_{i=1}^{n} log f^{\theta}_{X_i}(x_i / X_{i-1} = x_{i-1})$$

The maximum likelihood estimator (MLE) is expressed as follow:

$$\hat{\theta}_{MV} = argmax_{\theta} \;\; log L_n(\theta, x) \tag{1.3}$$

`Law of large numbers for martingale`

Theorem 1.6. *(LLN martingale)*
We suppose that M is a square integrable martingale such that $[M,M]_\infty = +\infty$ a.s, it follows that:

$$\frac{M_n}{[M,M]_n} \xrightarrow[n \to +\infty]{} 0$$

`The Girsanov theorem`

In continuous time studies, we give the maximum likelihood using the process itself on a time interval of type $[0,T]$. In this case, every X_t is a component of $\mathcal{C}([0,T])$. Let $f_{\mathrm{x}}(x)$ be the density of X on $\mathcal{C}([0,T])$.

Theorem 1.7. *(Girsanov theorem)*

Given an SDE of form $\mathrm{d}X_t = b_\theta(X_t,t)\mathrm{d}t + \sigma(X_t,t)\mathrm{d}B_t$, this density is induced by the Girsanov theorem as:

$$f_{\mathrm{x}}(x) = exp\left(\int_0^T \frac{b_\theta(x_s;s)}{\sigma^2(x_s;s)}\,\mathrm{dx}_s - \frac{1}{2}\int_0^T \frac{b_\theta^2(x_s;s)}{\sigma^2(x_s;s)}\,\mathrm{ds} \right)$$

2. Maximum Likelihood Estimation for different diffusion processes.

In the previous sections, we have defined the likelihood function for an SDE of a general form with a prior discretization by Euler scheme, now we are going to do the calculation for several diffusion processes in finance: Black Scholes, Ornstein-Uhlenbeck and Cox-Ingersoll-Ross. Note that when the transition densities of the continuous-time processes are explicitly known, we do not need to employ Euler scheme for discretization. Thus, when it is possible, we will do both methods to find the MLE.

2.1. The Black-Scholes model.

Also called geometric brownian motion, the Black-Scholes model is given as followed:

$$\boxed{\mathrm{d}X_t = \theta X_t \mathrm{d}t + \sigma X_t \mathrm{d}B_t.}$$

This process is particularly easy to use because its probability distribution is clearly calculable. Therefore, its resolution can be done in two different ways: either by directly using its explicit representation, or by making an approximation by Euler scheme.
We assume that $\sigma > 0$ is known and $\theta > 0$ is unknown. We are going to estimate θ using the

maximum likelihood method.
$(B_t)_t$ is a brownian motion as defined in Definition 1.1.

2.1.1. Explicit representation of X_t.

According to the Ito lemma, with $f(X_t) = log(X_t)$, we have:

$$\mathrm{d}log(X_t) = \frac{\mathrm{d}X_t}{X_t} + \frac{1}{2}\Big(-\frac{1}{X_t^2}\Big)\sigma^2 X_t^2 \mathrm{d}_t$$

We replace $\mathrm{d}X_t$ by $\theta X_t \mathrm{d}t + \sigma X_t \mathrm{d}B_t$. We have:

$\mathrm{dlog(X}_t) = (\theta - \frac{\sigma^2}{2})\mathrm{d}_t + \sigma \mathrm{d}B_t$

$\rightarrow log(X_t) - log(X_0) = (\theta - \frac{\sigma^2}{2})t + \sigma B_t$

$\rightarrow X_t = X_0 \exp\Bigg((\theta - \frac{\sigma^2}{2})t + \sigma B_t\Bigg)$

Since B_t is a brownian motion, we can determine the distribution of $log(X_t)$:

$$log(X_t) \sim \mathcal{N}\Bigg(log(X_0) + (\theta - \frac{\sigma^2}{2})t,\, \sigma^2 t\Bigg)$$

In the Black-Scholes model, X_t follows a log-normal distribution.

2.1.2. Euler scheme discretization.

The Black-Scholes model can be rewritten as:

$$X_{t_i} - X_{t_{i-1}} = \theta \int_{t_{i-1}}^{t_i} X_r \,\mathrm{d}r + \sigma \int_{t_{i-1}}^{t_i} X_r \,\mathrm{d}B_r$$

.This can be approximated by the following expression using Definition 1.3:

$$\begin{aligned} X_{t_i}^{(n)} - X_{t_{i-1}}^{(n)} &\approx \theta(t_i - t_{i-1})X_{t_{i-1}}^{(n)} + \sigma X_{t_{i-1}}^{(n)}(B_{t_i} - B_{t_{i-1}}) \\ \Longrightarrow X_{t_i}^{(n)} &\approx X_{t_{i-1}}^{(n)}[1 + \theta(t_i - t_{i-1}) + \sigma(B_{t_i} - B_{t_{i-1}})] \end{aligned}$$

Given that $B_{t_i} - B_{t_{i-1}} \sim \mathcal{N}(0,\, t_i - t_{i-1})$ and if we set $X_{t_{i-1}}^{(n)} = x \in$, we then have:

$$X_{t_i}^{(n)} \overset{loi}{\approx} \mathcal{N}\Bigg((1+\theta(t_i - t_{i-1}))x,\, \sigma^2 x^2 (t_i - t_{i-1})\Bigg) \tag{2.1}$$

2.1.3. Calculating the Maximum Likelihood Estimator of θ.

The density of the Black-Scholes conditional distribution is:

$$f_{X_{t_i}}^{\theta}(x_i / X_{t_{i-1}}^{(n)} = x_{i-1}) = \frac{1}{\sqrt{x_{i-1}^2 \sigma^2 (t_i - t_{i-1}) 2\pi}} \exp\left(-\frac{(x_i - (1+\theta(t_i - t_{i-1}))x_{i-1})^2}{2\sigma^2 x_{i-1}^2 (t_i - t_{i-1})}\right)$$

Hence, it's likelihood function can be expressed as:

$$\mathrm{L}_n(\theta; x) = \prod_{i=1}^n \frac{1}{\sqrt{x_{i-1}^2 \sigma^2 (t_i - t_{i-1}) 2\pi}} \exp\left(-\frac{(x_i - (1+\theta(t_i - t_{i-1}))x_{i-1})^2}{2\sigma^2 x_{i-1}^2 (t_i - t_{i-1})}\right)$$

$$\Longrightarrow \log(L_n(\theta; x)) = \sum_{i=1}^n \log\left(\frac{1}{\sqrt{x_i \sigma^2 (t_i - t_{i-1}) 2\pi}}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^n \frac{(x_i - (1+\theta(t_i - t_{i-1}))x_{i-1})^2}{x_{i-1}^2 (t_i - t_{i-1})}$$

$$\Longrightarrow \frac{\partial}{\partial\theta} \log(L_n(\theta; x)) = \frac{1}{\sigma^2} \sum_{i=1}^n \frac{(x_i - (1+\theta(t_i - t_{i-1}))x_{i-1})}{x_{i-1}(t_i - t_{i-1})}$$

$$\Longrightarrow \frac{\partial}{\partial\theta} \log(L_n(\theta; x)) = 0 \Longrightarrow \theta = \frac{\sum_{i=1}^n (x_i - x_{i-1})}{\sum_{i=1}^n (t_i - t_{i-1}) x_{i-1}}$$

For the sake of simplification, we assume that we have the concavity of the log likelihood function and therefore the theta canceling its derivative is a maximum.
In this case, we find that the MLE of θ becomes:

$$\hat{\theta}_{MV} = \frac{\sum_{i=1}^n (X_{t_i} - X_{t_{i-1}})}{\sum_{i=1}^n (t_i - t_{i-1}) X_{t_{i-1}}}$$

2.1.4. Corresponding martingale and consistency .

To study the consistency of the estimator, we express $\hat{\theta}_{MV}$ according to θ.
We replace X_{t_i} by its approximation $X_{t_{i-1}}[1 + \theta h_n + \sigma(B_{t_i} - B_{t_{i-1}})]$ and we find:

$$\hat{\theta}_{MV} = \theta + \sigma \frac{\sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})}{\sum_{i=1}^n (t_i - t_{i-1})}$$

Let M_n be $\sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})$ and $M_0 = 0$

We can show that $(M_n)_n$ is an integrable square martingale and its quadratic variation is: $[M,M]_n = \sum_{i=1}^{n}(t_i - t_{i-1})$.
Thus, $\hat{\theta}_{MV}$ becomes:

$$\hat{\theta}_{MV} = \theta + \sigma \frac{M_n}{[M,M]_n}$$

From this new expression of the estimator, we can study its consistancy using the law of large numbers for martingales given at Theorem 1.6

Let h be the difference $t_i - t_{i-1}, \forall i \in 1..n$. We have :

$$[M,M]_n = \sum_{i=1}^{n}(t_i - t_{i-1}) = nh \xrightarrow[n \to +\infty]{} +\infty$$

Hence:

$$\frac{M_n}{[M,M]_n} \xrightarrow[n \to +\infty]{} 0$$

Thereby, $\hat{\theta}_{MV}$ is a consistant estimator, i.e.:

$$\hat{\theta}_{MV} = \theta + \sigma \frac{M_n}{[M,M]_n} \xrightarrow[n \to +\infty]{} \theta$$

2.2. The Ornstein-Uhlenbeck process.

The Ornstein-Uhlenbeck process is defined as followed:

$$\boxed{\mathrm{d}X_t = -\theta X_t \mathrm{d}t + \sigma \mathrm{d}B_t.}$$

As for the Black-Scholes model, this process can be solved using its explicit representation or by discretization with Euler scheme. However, this model does not translate very well on financial markets where the interest rate and the volatility are not always constants.

2.2.1. Explicit representation of X_t.

Using the Stochastic Differential Equation of the process and a partial integration we can derive that:

$$\begin{aligned} \mathrm{d}X_t &= -\theta X_t \mathrm{d}t + \sigma \mathrm{d}B_t \\ \Longrightarrow \mathrm{d}(e^{\theta t} X_t) &= \theta e^{\theta t} X_t \mathrm{d}t + e^{\theta t} \mathrm{d}X_t \end{aligned}$$

Then we use the equation to replace dX_t by $-\theta X_t \mathrm{d}t + \sigma \mathrm{d}B_t$:

$$\begin{aligned} \mathrm{d}(e^{\theta t} X_t) &= \theta e^{\theta t} X_t \mathrm{d}t + e^{\theta t}(-\theta X_t \mathrm{d}t + \sigma \mathrm{d}B_t) \\ \Longrightarrow \mathrm{d}(e^{\theta t} X_t) &= \sigma e^{\theta t} \mathrm{d}B_t \end{aligned}$$

Next, we integrate this new equation:

$$e^{\theta t_i} X_{t_i} - e^{\theta t_{i-1}} X_{t_{i-1}} = \sigma \int_{t_{i-1}}^{t_i} e^{\theta r} \, \mathrm{d}B_r$$

$$X_{t_i} = e^{-\theta(t_i - t_{i-1})} \left[X_{t_{i-1}} + \sigma \int_{t_{i-1}}^{t_i} e^{\theta(r - t_{i-1})} \, \mathrm{d}B_r \right]$$

If $X_{t_{i-1}} = x$,

$$X_{t_i} = e^{-\theta(t_i - t_{i-1})} \left[x + \sigma \int_{t_{i-1}}^{t_i} e^{\theta(r - t_{i-1})} \, \mathrm{d}B_r \right]$$

The Itô isometry helps us find that:

$$X_{t_i} \overset{law}{\sim} \mathcal{N}\left(e^{-\theta(t_i - t_{i-1})} x, \, \sigma^2 e^{-2\theta(t_i - t_{i-1})} \int_{t_{i-1}}^{t_i} e^{2\theta(r - t_{i-1})} \, \mathrm{d}r \right)$$

After some calculations, we obtain the transition density of the process and it follows that:

$$X_{t_i} \overset{law}{\sim} \mathcal{N}\left(e^{-\theta(t_i - t_{i-1})} x, \, \frac{\sigma^2}{2\theta} (1 - e^{-2\theta(t_i - t_{i-1})}) \right)$$

Considering $t = ih$ and $t_{i-1} = (i-1)h$ and if $X_{(i-1)h} = x$ we get finally:

$$X_{ih} \overset{loi}{\sim} \mathcal{N}(x e^{-\theta h}, \, \frac{\sigma^2}{2\theta}(1 - e^{-2\theta h})) \tag{2.2}$$

2.2.2. Euler scheme discretization.

Let's consider that $(X_{t_i}^{(n)})_{i=1,\ldots,n}$ is the Euler scheme associated to the process $(X_t)_{t\geq 0}$, according to Definition 1.3:

$$X_{t_i} - X_{t_{i-1}} = -\theta \int_{t_{i-1}}^{t_i} X_r \, \mathrm{d}r + \sigma (B_{t_i} - B_{t_{i-1}})$$

$$\Longrightarrow X_{t_i}^{(n)} - X_{t_{i-1}}^{(n)} \approx -\theta (t_i - t_{i-1}) X_{t_{i-1}}^{(n)} + \sigma (B_{t_i} - B_{t_{i-1}})$$

We set $V_i = B_{t_i} - B_{t_{i-1}}$, $V_i \sim \mathcal{N}(0, \, t_i - t_{i-1})$.
V_i and V_j are independent $\forall i, j$.

Hence, if we assume that $X_{t_{i-1}}^{(n)} = x \in \mathbb{R}$,

$$X_{t_i}^{(n)} \overset{(law)}{\approx} \mathcal{N}(x(1-\theta(t_i - t_{i-1})),\, \sigma^2(t_i - t_{i-1})) \tag{2.3}$$

Considering $t_i = ih$ and $t_{i-1} = (i-1)h$,

$$X_{ih}^{(n)} \overset{(law)}{\approx} \mathcal{N}(x(1-\theta h),\, \sigma^2 h) \tag{2.4}$$

Looking at the results, we notice that the formula obtained using the Euler scheme associated to $(X_t)_t$ is equivalent to the one obtained from the explicit representation of $(X_t)_t$ since we approximate $e^{-\theta h}$ by its Taylor expression at order 1:

$$e^{-\theta h} \approx 1 - \theta h$$

Nevertheless, the first method should be preferred because the latter remains an approximation.

2.2.3. Calculating the Maximum Likelihood Estimator of θ.

Once we have the conditional distribution of our process X_t, we can compute an estimator of θ with the maximum likelihood method.

The likelihood function is approximated by:

$$L_n(\theta; x) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2(t_i - t_{i-1})}} \exp\left(-\frac{[x_i - x_{i-1}(1-\theta(t_i - t_{i-1}))]^2}{2\sigma^2(t_i - t_{i-1})}\right)$$

$$\log(L_n(\theta; x)) = \frac{-1}{2}\sum_{i=1}^{n} \log(2\pi\sigma^2(t_i - t_{i-1})) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} \frac{[x_i - x_{i-1}(1-\theta(t_i - t_{i-1}))]^2}{(t_i - t_{i-1})}$$

To determine the MLE of θ, we are going to derivate this log-likelihood function.

$$\frac{\partial}{\partial\theta} log(L_n(\theta, x)) = \frac{-1}{\sigma^2}\sum_{i=1}^{n} \frac{x_{i-1}(t_i - t_{i-1})(x_i - x_{i-1} + \theta x_{i-1}(t_i - t_{i-1}))}{t_i - t_{i-1}}$$

As done previously, we assume the concavity of the function and we set the derivate to zero:

$$\frac{\partial}{\partial\theta} log(L_n(\theta, x)) = 0 \implies \sum_{i=1}^{n} x_{i-1}(x_i - x_{i-1} + \theta x_{i-1}(t_i - t_{i-1})) = 0$$

We therefore conclude that the MLE of θ is:

$$\hat{\theta}_{MV} = -\frac{\sum_{i=1}^{n} x_{i-1}(x_i - x_{i-1})}{\sum_{i=1}^{n} x_{i-1}^2(t_i - t_{i-1})} \tag{2.5}$$

Since $X_{t_i}^{(n)} - X_{t_{i-1}}^{(n)} \approx -\theta(t_i - t_{i-1})X_{t_{i-1}}^{(n)} + \sigma(B_{t_i} - B_{t_{i-1}})$ (using the Euler scheme), we can rewrite the estimator just found as a function of the parameter θ in order to facilitate the study of consistency:

$$\hat{\theta}_{MV} = \theta - \sigma \frac{\sum_{i=1}^{n} X_{t_{i-1}}^{(n)}(B_{t_i} - B_{t_{i-1}})}{\sum_{i=1}^{n} {X_{t_{i-1}}^{(n)}}^2(t_i - t_{i-1})} \tag{2.6}$$

2.2.4. Corresponding martingale and consistency .

We are going to use here the direct approach without any Euler approximation.

Let's $M_n = \sum_{i=1}^{n} X_{t_{i-1}}(B_{t_i} - B_{t_{i-1}})$ be a martingale. Thus, $\hat{\theta}_{MV}$ can be write in the following form:

$$\hat{\theta}_{MV} = \theta - \sigma \frac{M_n}{[M, M]_n}$$

Where

$$M_n = B_0 + \sum_{i=1}^{n} A_{t_i}(B_{t_i} - B_{t_{i-1}})$$

and $B_0 = 0, \quad A_{t_i} = X_{t_{i-1}}$.

Consider the σ-algebra $F_n = \sigma(B_{t_i}, i \leq n)$
.

From this we deduce that $(M_n)_n$ is the predictable transform of the martingale $(B_{t_n})_n$ by the predictable process $(A_{t_n})_n$ which is F_{n-1} -measurable.
$(M_n)_n$ is then, like $(B_t)_t$ an square integrable martingale. Its quadratic variation is:

$$[M, M]_n = \sum_{i=1}^{n} A_i^2(\,[B, B]_i - [B, B]_{i-1}\,)$$
$$[M, M]_n = \sum_{i=1}^{n} {X_{t_{i-1}}}^2(t_i - t_{i-1})$$

To study the consistency of $\hat{\theta}_{MV}$, we are going to use the law of large numbers introduced in Theorem 1.6. To apply the theorem, it is necessary to prove that $[M, M]_n \to +\infty$.

$$[M, M]_n \to +\infty \iff \sum_{i=1}^{n} {X_{t_{i-1}}}^2 \to +\infty$$

$(X_{t_i})_{i\geq 0}$ is a Markov chain with values in $\mathbb{R}$. We can prove that this Markov chain is irreducible, recurrent, aperiodic and admits a unique invariant probability that we will try to find. We know that:

$$Law(X_{t_i}/X_{t_{i-1}} = x) \sim \mathcal{N}\left(xe^{-\theta(t_i - t_{i-1})}, \frac{\sigma^2}{2\theta}(1 - e^{-2\theta(t_i - t_{i-1})})\right)$$

Assuming $X_0 = x$, we can show that:

$$X_{t_i} \sim \mathcal{N}\left(xe^{-\theta t_i}, \frac{\sigma^2}{2\theta}(1 - e^{-2\theta t_i})\right)$$

Because the chain is aperiodic, it converges to its unique invariant probability.
By extending $t_i = ih \to +\infty$, we have:

$$X_{t_i} \underset{n\to+\infty}{\longrightarrow} \mathcal{N}(0, \frac{\sigma^2}{2\theta}) = \pi \text{ which is the invariant probability.}$$

Now, we use the ergodic theorem for positive recurrent Markov chains which is defined as follows:

Theorem 2.1. *(Ergodic theorem)*

We assume that the process X is ergodic (positive recurrent markov chain).
For all function $f \in L^1(\pi)$, we have the following a.s convergence:

$$\frac{1}{n}\sum_{i=1}^{n} f(X_{t_{i-1}}) \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}} f(x)d\pi(x) = \mathbb{E}[f(X_\infty)]$$

In the case of Ornstein-Uhlenbeck, we set $f(x) = x^2$ to apply the theorem.

$$\frac{1}{n}\sum_{i=1}^{n} {X_{t_{i-1}}}^2 \xrightarrow{ps} \int_{\mathbb{R}} x^2 d\pi(x) = \mathbb{E}[X_\infty^2] = Var(X_\infty) + \mathbb{E}[X_\infty]^2 = \frac{\sigma^2}{2\theta}$$

Then we find:

$$\frac{1}{n}\sum_{i=1}^{n} {X_{t_{i-1}}}^2 \underset{n\to+\infty}{\approx} n\frac{\sigma^2}{2\theta} \xrightarrow[n\to+\infty]{} +\infty$$

Finally, we use the law of large numbers introduced in Theorem 1.6 to prove the consistency of the maximum likelihood estimator:

$$\hat{\theta}_{MV} = \theta - \sigma\frac{M_n}{[M,M]_n} \text{ and } [M,M]_n \to +\infty \iff \hat{\theta}_{MV} \xrightarrow[n\to+\infty]{} \theta$$

2.3. Cox-Ingersoll-Ross model.

The Cox-Ingersoll-Ross model is defined by the following stochastic differential equation:

$$\boxed{\mathrm{d}X_t = \mu(\nu - X_t)\mathrm{d}t + \sigma\sqrt{X_t}\mathrm{d}B_t.}$$

that can be rewritten as:

$$\boxed{\mathrm{d}X_t = (a - bX_t)\mathrm{d}t + \sigma\sqrt{X_t}\mathrm{d}B_t.}$$

We assume a et $\sigma > 0$ are known and we aim to estimate b.

2.3.1. Problem in discrete resolution.

Unlike the model seen earlier, the Cox-Ingersell-Ross model does not have any explicit representation. We are compelled to use the Euler scheme discretization directly. However, even using the Euler scheme like the other processes, we face some issues due to the square root in the equation. Indeed, $X_{t_i}^{(n)}$ can take negative values for a given i since this variable follows a normal distribution . This is why we have decide to study this model in continuous time.

2.3.2. Switching to continuous resolution.

We write the CIR model once more as:

$$\boxed{\mathrm{d}X_t = (a - \theta X_t)\mathrm{d}t + \sigma\sqrt{X_t}\mathrm{d}B_t.}$$

We are going to apply the Girsanov theorem to this equation where:
$b_\theta(X_t, t) = a - \theta X_t$ and $\sigma(X_t, t) = \sigma\sqrt{X_t}$.

The maximum likehood becomes:

$$\mathrm{L}_T(\theta; x) = f_{\mathrm{x}}(x) = \exp\left(\int_0^T \tfrac{(a-\theta x_s)}{\sigma^2 x_s}\,\mathrm{dx}_s - \tfrac{1}{2}\int_0^T \tfrac{(a-\theta x_s)^2}{\sigma^2 x_s}\,\mathrm{ds} \right)$$

$$\rightarrow log(L_T(\theta; x)) = \int_0^T \tfrac{(a-\theta x_s)}{\sigma^2 x_s}\,\mathrm{dx}_s - \tfrac{1}{2}\int_0^T \tfrac{(a-\theta x_s)^2}{\sigma^2 x_s}\,\mathrm{ds}$$

$$\rightarrow \tfrac{\partial}{\partial\theta} log(L_T(\theta; x)) = \tfrac{aT - x_T + x_0}{\sigma^2} - \theta\int_0^T \tfrac{x_s}{\sigma^2}\,\mathrm{ds}$$

$$\frac{\partial}{\partial\theta} log(L_T(\theta; x)) = 0 \Longrightarrow \theta = \frac{aT - x_T + x_0}{\int_0^T x_s\,\mathrm{ds}}$$

$\hat{\theta}_{MV}$ is expressed as followed:

$$\boxed{\hat{\theta}_{MV} = \frac{aT - X_T + X_0}{\int_0^T X_s \,\mathrm{ds}}}$$

$$\begin{aligned} X_T &= X_0 + \int_0^T (a - \theta X_t)\,\mathrm{dt} + \sigma \int_0^T \sqrt{X_t}\,\mathrm{d}B_t \\ \implies X_T &= X_0 + aT - \theta \int_0^T X_t\,\mathrm{dt} + \sigma \int_0^T \sqrt{X_t}\,\mathrm{d}B_t \end{aligned}$$

Hence,

$$\boxed{\hat{\theta}_{MV} = \theta - \sigma \frac{\int_0^T \sqrt{X_t}\,\mathrm{d}B_t}{\int_0^T X_t\,\mathrm{d}B_t}}$$

3. Validation studies & Discussion.

4. Conclusion. The paper is organized as follows. Our main results are in **??**, our new algorithm is in **??**, experimental results are in **??**, and the conclusions follow in **??**.

REFERENCES