Parameter Inference for SDE

Federico Betti
MATH-450, Numerical Integration of SDEs, EPFL Lausanne, Switzerland

June 5, 2022

Abstract

In this report the results obtained in the Mini Project of the course are presented. The notation is the same used in the project descriptions and more in general in the course, unless stated otherwise.

Introduction

We start by presenting the Ornstein-Uhlenbeck process and some important asymptotic results concerning the sample paths of such a stochastic process. In particular, by ergodicity the latter admits a limiting invariant measure with a density function which we can explicitly calculate and characterize as the solution to the Fokker-Planck equation. We then focus on the problem of estimating the drift and the diffusion coefficients, when they are unknown, given some discrete observations of a sample path. We first derive classical estimators, and we show that they need the sampling rate to be sufficiently small to provide a good estimate of the true coefficients. As this condition is not necessarily satisfied in the applications, where observations are given and cannot tune the sampling rate, we take a different approach and construct estimators, based on the properties of the limiting distribution of the process, which are asymptotically unbiased independently of the sampling rate. Finally, we show numerically that as we increase the number of observations at disposal, these estimators do converge towards the true value in the limit of infinite data, and they satisfy a central limit theorem.

We start by quickly recalling the general setting and some important definitions. Let $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space; we define a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ as a standard one-dimensional Gaussian process $\{W(t)\}_{t\geq 0}$ with $\mathbb{E}\left[W(s)W(t)\right] = \min\{s,t\}$. For some final time $T > 0$, we then consider the following Itô stochastic differential equation

$$dX(t) = -\alpha X(t)dt + \sqrt{2\sigma}dW(t), X(0) = X_0, \quad 0 \leq t \leq T, \tag{1}$$

where $\alpha > 0$ is some drift coefficient and the stochastic term represent a diffusion governed by some parameter $\sigma > 0$. We call a sample path of the solution to (1) a realization of the Ornstein-Uhlenbeck process.

Question 1

Let $u : [0,T] \times \mathbb{R} \to \mathbb{R}$ defined by $(t,x) \mapsto xe^{\alpha t}$. Since $u(\cdot,\cdot)$ is smooth enough, applying the Itô formula on the latter we can easily compute the solution to (1). Indeed since $\frac{\partial^2 u}{\partial x^2}(t, X(t)) = 0$ we have that

$$d(u(t,X(t))) = \frac{\partial u}{\partial X}dX(t) + \frac{\partial u}{\partial t}dt \overset{(1)}{=} e^{\alpha t}\sqrt{2\sigma}dW(t).$$

Integrating on both sides now gives

$$u(t,X(t)) = u(0,X(0)) + \sqrt{2\sigma}\int_0^t e^{\alpha s}dW(s), \tag{2}$$

or equivalently

$$X(t) = e^{-\alpha t}X_0 + \sqrt{2\sigma}\int_0^t e^{-\alpha(t-s)}dW(s). \tag{3}$$

Since $g(s) = e^{\alpha s}$ is continuous and hence in $L^2(0,T)$, we know that by the Itô isometry

$$\int_0^t e^{-\alpha(t-s)}dW(s) \sim \mathcal{N}\left(0, \int_0^t e^{-2\alpha(t-s)}ds\right) = \mathcal{N}\left(0, \frac{1}{2\alpha}\left(1 - e^{-2\alpha t}\right)\right). \tag{4}$$

Finally, the distribution μ_t of $X(t)$ is

$$\mu_t \sim \mathcal{N}\left(e^{-\alpha t}X_0, \frac{\sigma}{\alpha}\left(1-e^{-2\alpha t}\right)\right), \tag{5}$$

where we assume that X_0 is fixed, and its probability density function is

$$\rho_t(x) = \sqrt{\frac{\alpha}{2\pi\sigma\left(1-e^{-2\alpha t}\right)}} \exp\left(-\frac{\alpha\left(x-e^{-\alpha t}X_0\right)^2}{2\sigma\left(1-e^{-2\alpha t}\right)}\right). \tag{6}$$

Question 2

We know by ergodicity that the solution to (1) admits a limiting invariant measure. This means that the distribution μ_t tends as $t \to +\infty$ towards a limiting distribution μ_∞. Moreover, it is a well known fact that if we are given a sequence of Gaussian random variables $\{X_n \sim \mathcal{N}\left(\mu_n, \sigma_n^2\right)\}_{n\in\mathbb{N}}$ with $\mu_n \to \mu$ and $\sigma_n^2 \to \sigma^2$ then the sequence $\{X_n\}_{n\in\mathbb{N}}$ converges in distribution to $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$. Then, using the fact that $\lim_{t\to+\infty} e^{-\alpha t}X_0 = 0$ and $\lim_{t\to+\infty} \frac{\sigma}{\alpha}\left(1-e^{-2\alpha t}\right) = \frac{\sigma}{\alpha}$, from (5) we conclude that asymptotically every sample path of the Ornstein-Uhlenbeck process will be distributed according to the limiting measure $\mu_\infty \sim \mathcal{N}\left(0, \frac{\sigma}{\alpha}\right)$. The interpretation of this result is that at stationarity the ratio between the diffusion coefficient and the drift controls the deviation of the process from 0.

Consider now the probability distribution function associated to μ_∞, that is

$$\rho_\infty(x) = \sqrt{\frac{\alpha}{2\pi\sigma}} \exp\left(\frac{-\alpha x^2}{2\sigma}\right). \tag{7}$$

Moreover, let us consider the differential operator $\mathcal{L}$, defined by

$$\mathcal{L}u(x) = -\alpha x u'(x) + \sigma u''(x). \tag{8}$$

We now show that (7) solves the Fokker-Planck equation

$$\mathcal{L}^*\rho = 0, \quad \int_{\mathbb{R}} \rho = 1. \tag{9}$$

where $\mathcal{L}^*$ is the L_2-adjoint of $\mathcal{L}$, namely the operator satisfying

$$\int_{\mathbb{R}} \phi(x)\mathcal{L}u(x)dx = \int_{\mathbb{R}} u(x)\mathcal{L}^*\phi(x)dx \quad \forall \phi \in \mathcal{C}_0^\infty(\mathbb{R}). \tag{10}$$

It is straightforward to derive an expression for $\mathcal{L}^*$. Let $u, \phi : \mathbb{R} \to \mathbb{R}$ be smooth and compactly supported functions. Then, from (10), using the definition of $\mathcal{L}$ and integrating by parts:

$$\begin{aligned}
&\int_{\mathbb{R}} \phi(x)\mathcal{L}u(x)dx = -\alpha \int_{\mathbb{R}} \phi(x)xu'(x)dx + \sigma \int_{\mathbb{R}} \phi(x)u''(x)dx \\
&= -\alpha\left(\Big[xu(x)\phi(x)\Big]_{-\infty}^{+\infty} - \int_{\mathbb{R}} (x\phi(x))'u(x)dx\right) + \sigma \int_{\mathbb{R}} \phi''(x)u(x)dx \\
&= \alpha \int_{\mathbb{R}} (x\phi(x))'\, u(x)dx + \sigma \int_{\mathbb{R}} \phi''(x)u(x)dx = \int_{\mathbb{R}} \left(\alpha\, (x\phi(x))' + \sigma\phi''(x)\right) u(x)dx,
\end{aligned}$$

where we used the fact that u, ϕ have compact support together with all their derivatives (it is a well known fact that $\operatorname{supp}(\psi^{(n)}) \subseteq \operatorname{supp}(\psi)$ for all $\psi \in \mathcal{C}_0^\infty(\mathbb{R})$ and for any $n \in \mathbb{N}$). By the fundamental lemma of calculus of variations we derive that

$$\mathcal{L}^*u(x) = \alpha(xu(x))' + \sigma u''(x) = \alpha u(x) + \alpha x u'(x) + \sigma u''(x). \tag{11}$$

Thus, we have

$$\begin{aligned}
\rho_\infty'(x) &= -\frac{\alpha}{\sigma} x \rho_\infty(x), \\
\rho_\infty''(x) &= -\frac{\alpha}{\sigma}\rho_\infty(x) + \left(\frac{\alpha}{\sigma}\right)^2 x^2 \rho_\infty(x),
\end{aligned}$$

and, substituting (7) in (11), the claim follows:

$$\begin{aligned}
&\mathcal{L}^*\rho_\infty(x) = \alpha\rho_\infty(x) - \frac{\alpha^2}{\sigma}x^2\rho_\infty(x) + \sigma\left(\frac{\alpha}{\sigma}\right)^2 x^2\rho_\infty(x) - \sigma\frac{\alpha}{\sigma}\rho_\infty(x) = 0, \\
&\int_{\mathbb{R}} \rho_\infty = 1.
\end{aligned}$$

Question 3

We now verify numerically that for T large enough the numerical solution $X(T)$ is approximately distributed according to the invariant measure μ_∞. To solve numerically (1), we consider a step size Δt yielding a uniform partition $\{0 = t_0 < t_1 < t_2 < \ldots < t_N = T\}$ of $[0, T]$, with $N = \frac{T}{\Delta t}$. Then, the Euler-Maruyama method applied to (1) reads

$$X_{n+1} = X_n - \alpha X_n \Delta t + \sqrt{2\sigma}\left(W(t_{n+1}) - W(t_n)\right), \tag{12}$$

for $n = 0, \ldots, N-1$ and given a realization of the Wiener process $\{W(t)\}_{t\geq 0}$. Figure 1 shows the histogram of $X(T)$ for $T = 10^3$, obtained with $M = 10^4$ different realizations of a standard Brownian motion. We see that for T large enough the solution $X(T)$ is approximately distributed according to the invariant measure μ_∞ as the histogram approaches the probability density function ρ_∞.

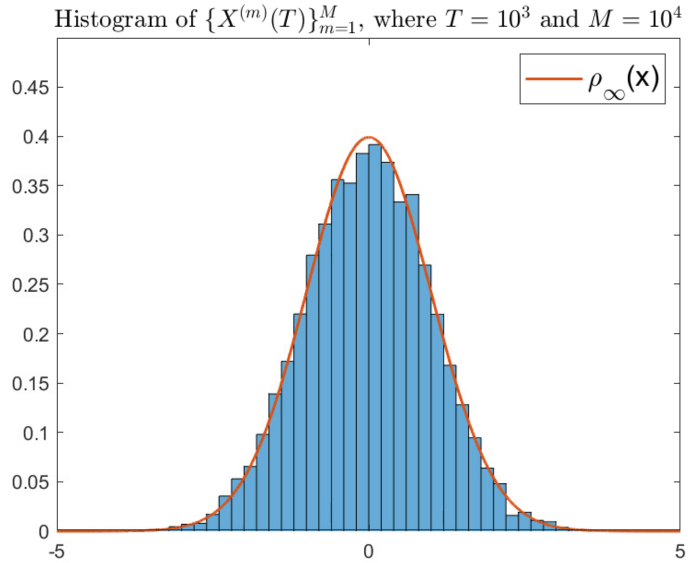

Figure 1: Histogram of the numerical solution $\{X^{(m)}(T)\}_{m=1}^M$ of (1) obtained from (12) with step size $\Delta t = 0.01$ and final time $T = 10^3$ for $M = 10^4$ independent realizations of a Wiener process: the distribution of $X(T)$ approaches the invariant measure μ_∞.

Question 4

We aim at computing the covariance function of the Ornstein–Uhlenbeck process at stationarity, which we define as

$$\mathcal{C}(t,s) = \mathrm{Cov}(X(t), X(s)), \tag{13}$$

for s and t large enough. We first state and prove the following identity: for any $f, g \in L^2(0,T)$

$$\mathbb{E}\left[\int_0^t f(r)dW(r) \int_0^t g(r)dW(r)\right] = \int_0^t f(r)g(r)dr, \quad 0 \leq t \leq T. \tag{14}$$

Indeed, since both terms on the left hand side of (14) are zero mean Gaussian random variables, we can write that

$$\begin{aligned}
&\mathbb{E}\left[\int_0^t f(r)dW(r) \int_0^t g(r)dW(r)\right] = \\
&\mathrm{Cov}\left(\int_0^t f(r)dW(r), \int_0^t g(r)dW(r)\right) = \\
&\frac{1}{2}\left(\mathrm{Var}\left[\int_0^t f(r)dW(r) + \int_0^t g(r)dW(r)\right] - \mathrm{Var}\left[\int_0^t f(r)dW(r)\right] - \mathrm{Var}\left[\int_0^t g(r)dW(r)\right]\right) = \\
&\frac{1}{2}\left(\mathrm{Var}\left[\int_0^t (f(r)+g(r))\, dW(r)\right] - \mathrm{Var}\left[\int_0^t f(r)dW(r)\right] - \mathrm{Var}\left[\int_0^t g(r)dW(r)\right]\right),
\end{aligned}$$

where in the last equality we use linearity of the stochastic integral. Finally, since all the terms above have zero mean and using the Itô isometry

$$\mathbb{E}\left[\int_0^t f(r)dW(r) \int_0^t g(r)dW(r)\right] = \frac{1}{2}\left(\mathbb{E}\left[\int_0^t (f(r)+g(r))^2 dr\right] - \mathbb{E}\left[\int_0^t f(r)^2 dr\right] - \mathbb{E}\left[\int_0^t g(r)^2 dr\right]\right). \tag{15}$$

The claim follows since f, g are deterministic and developing the square in the first term.

Thus we have that, from (5), for all $t, s \in \mathbb{R}$

$$\mathcal{C}(t,s) = \mathbb{E}\Bigg[\bigg(X(t) - \mathbb{E}[X(t)]\bigg)\bigg(X(s) - \mathbb{E}[X(s)]\bigg)\Bigg] = \mathbb{E}\left[\sqrt{2\sigma}\int_0^t e^{-\alpha(t-r)}dW(r)\sqrt{2\sigma}\int_0^s e^{-\alpha(s-r)}dW(r)\right] =$$
$$2\sigma e^{-\alpha(t+s)}\mathbb{E}\left[\int_0^t e^{\alpha r}dW(r)\int_0^s e^{\alpha r}dW(r)\right] \overset{(*)}{=} \frac{\sigma}{\alpha}e^{-\alpha(t+s)}\left(e^{2\alpha\min\{t,s\}} - 1\right) =$$
$$\frac{\sigma}{\alpha}\left(e^{-\alpha|t-s|} - e^{-\alpha(t+s)}\right).$$

where at $(*)$ we apply (14) to the functions $f(r) = e^{\alpha r}\mathbb{1}_{\{r<t\}}$, $g(r) = e^{\alpha r}\mathbb{1}_{\{r<s\}} \in L^2(0,T)$, namely we use the identity

$$\mathbb{E}\left[\int_0^t e^{\alpha r}dW(r)\int_0^s e^{\alpha r}dW(r)\right] = \mathbb{E}\left[\int_0^T e^{\alpha r}\mathbb{1}_{\{r<t\}}e^{\alpha r}\mathbb{1}_{\{r<s\}}dr\right] = \int_0^{\min\{t,s\}} e^{2\alpha r}dr. \tag{16}$$

Finally, at stationarity the second term vanishes faster than the first one (i.e. the first term is the dominant one) leading to

$$\mathcal{C}(t,s) = \frac{\sigma}{\alpha}e^{-\alpha|t-s|}. \tag{17}$$

Question 5

Let $\Delta > 0$ a sampling rate and assume that some discrete observations $\{\tilde{X}_n\}_{n=0}^N$ are given, where $N = T/\Delta$ and $\widetilde{X}_n$ is the value of a sample path of the solution to (1) at time $n\Delta$. We want to construct good estimators of the drift coefficient α and of the diffusion coefficient σ when they are not known. To this aim, we define the following quantities:

$$\hat{\alpha}_N^\Delta = -\frac{\sum_{n=0}^{N-1}\widetilde{X}_n\left(\widetilde{X}_{n+1} - \widetilde{X}_n\right)}{\Delta\sum_{n=0}^{N-1}\widetilde{X}_n^2}, \tag{18}$$

$$\hat{\sigma}_N^\Delta = \frac{1}{2\Delta N}\sum_{n=0}^{N-1}\left(\widetilde{X}_{n+1} - \widetilde{X}_n\right)^2. \tag{19}$$

First, notice that we can rewrite (18) as

$$\hat{\alpha}_N^\Delta = \frac{1}{\Delta}\left(1 - \frac{\sum_{n=0}^{N-1}\widetilde{X}_n\widetilde{X}_{n+1}}{\sum_{n=0}^{N-1}\widetilde{X_n}^2}\right). \tag{20}$$

To compute the almost sure limits of (18) in the limit of infinite data, i.e. as $N \to \infty$, we can apply the ergodic theorem stated in the Hint of the question to the functions

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = x^2, \qquad g : \mathbb{R}^2 \to \mathbb{R}, \quad g(x,y) = xy. \tag{21}$$

which are smooth enough so that the ergodic theorem applies as $f \in C^\infty(\mathbb{R},\mathbb{R})$ and $g \in C^\infty(\mathbb{R}^2,\mathbb{R})$. Thus we can compute

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\widetilde{X}_n^2 = \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}f\left(\widetilde{X}_n\right) = \mathbb{E}^{\mu_\infty}\left[X_t^2\right] = \mathcal{C}(t,t) \overset{(17)}{=} \frac{\sigma}{\alpha},$$
$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\widetilde{X}_n\widetilde{X}_{n+1} = \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}g\left(\widetilde{X}_n,\widetilde{X}_{n+1}\right) = \mathbb{E}^{\mu_\infty}\left[X_tX_{t+\Delta}\right] = \mathcal{C}(t,t+\Delta) \overset{(17)}{=} \frac{\sigma}{\alpha}e^{-\alpha\Delta}.$$

where we assume t large enough so that at stationarity $\mathbb{E}^{\mu_\infty}[X_t] = 0$. Thus

$$\alpha_\infty^\Delta := \lim_{N\to\infty}\hat{\alpha}_N^\Delta = \frac{1}{\Delta}\left(1 - e^{-\alpha\Delta}\right) \neq \alpha. \tag{22}$$

Similarly, for $\hat{\sigma}_N^\Delta$ we apply the ergodic theorem to the smooth function $h : \mathbb{R}^2 \to \mathbb{R}, h(x,y) = (x-y)^2$, which satisfies $h \in C^\infty(\mathbb{R}^2,\mathbb{R})$, to conclude that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\left(\widetilde{X}_{n+1} - \widetilde{X}_n\right)^2 = \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}h\left(\widetilde{X}_n,\widetilde{X}_{n+1}\right) = \tag{23}$$

$$\mathbb{E}^{\mu_\infty}\left[(X_t - X_{t+\Delta})^2\right] = \mathbb{E}^{\mu_\infty}\left[X_t^2\right] + \mathbb{E}^{\mu_\infty}\left[X_{t+\Delta}^2\right] - 2\mathbb{E}^{\mu_\infty}\left[X_tX_{t+\Delta}\right]. \tag{24}$$

Taking again t large enough so that $\mathbb{E}^{\mu_\infty}[X_t] = 0$ we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \left(\widetilde{X}_{n+1} - \widetilde{X}_n\right)^2 = \mathcal{C}(t,t) + \mathcal{C}(t+\Delta, t+\Delta) - 2\mathcal{C}(t, t+\Delta) \overset{(17)}{=} \frac{\sigma}{\alpha} + \frac{\sigma}{\alpha} - 2\frac{\sigma}{\alpha} e^{-\alpha\Delta} \tag{25}$$

which directly leads to

$$\sigma_\infty^\Delta = \lim_{N\to\infty} \hat{\sigma}_N^\Delta = \frac{1}{2\Delta}\left(\frac{\sigma}{\alpha} + \frac{\sigma}{\alpha} - 2\frac{\sigma}{\alpha}e^{-\alpha\Delta}\right) = \frac{\sigma}{\alpha\Delta}\left(1 - e^{-\alpha\Delta}\right) \neq \sigma. \tag{26}$$

In the above, we claim that $\alpha_\infty^\Delta \neq \alpha$ and similarly $\sigma_\infty^\Delta \neq \sigma$ because we always assume $\alpha, \Delta > 0$. Namely, we can conclude that the two estimators (18) and (19) are asymptotically biased.

Question 6

On the other hand, if we take the sampling rate Δ sufficiently small, then the asymptotic bias of (18) and (19) vanishes. Indeed, using the fact that $\lim_{\Delta\to 0}\,(e^\Delta - 1)\,/\,\Delta = 1$, it is immediate from (22) and (26) to see that

$$\lim_{\Delta\to 0} \alpha_\infty^\Delta = \alpha, \quad \lim_{\Delta\to 0} \sigma_\infty^\Delta = \sigma. \tag{27}$$

Question 7

Figure 2 shows the convergence of the estimators $\hat{\alpha}_N^\Delta$ and $\hat{\sigma}_N^\Delta$ as a function of the sampling rate Δ for $\Delta = 2^{-i}, i = 0, \dots, 7$. Coherently with the previous results the quality of the estimate of the true unknown coefficients increases as $\Delta \to 0$; perhaps one can observe that the estimate of the diffusion coefficient σ seems to be more precise. The solution to (1) is again approximated by employing the Euler-Maruyama method with a step size $\Delta t = 2^{-10}$.

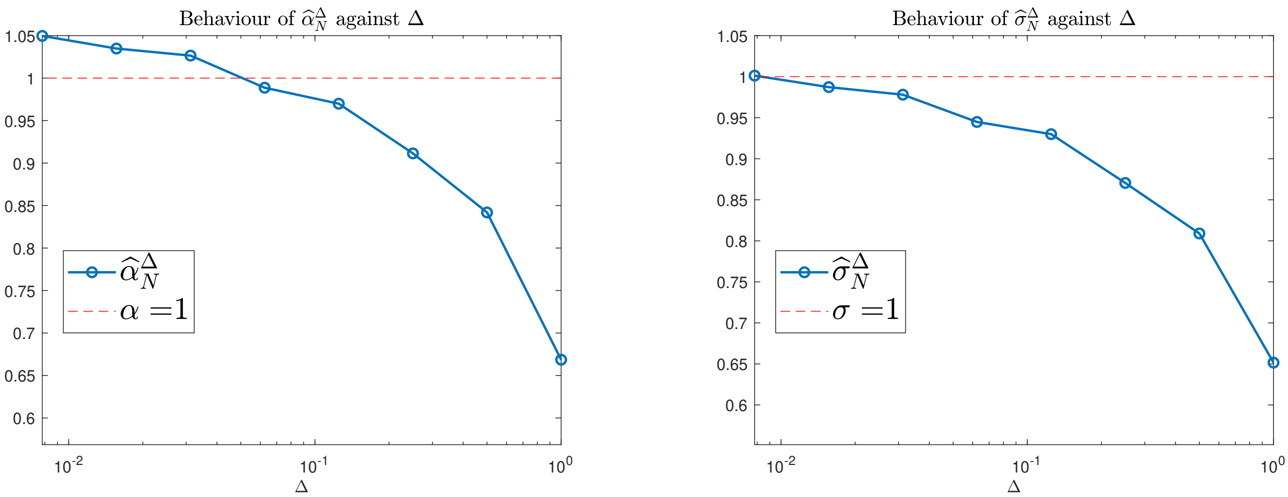

Figure 2: Estimators (18) and (19) against the sampling rate Δ: the red line shows the true coefficients α and σ.

Question 8

As highlighted by the previous results at (22) and (26), the drawback of the estimators defined in (18) and (19) is that they are not asymptotically unbiased independently of the sampling rate Δ and this is sub-optimal for real world applications where one is not allowed, in general, to choose the latter, but observations are already given. We thus take a different approach: let $-\mathcal{L}_a$ be the differential operator defined in (8) where we replace the true drift coefficient α by some $a \in \mathbb{R}$. We know that $-\mathcal{L}_a$ admits a countable set of eigenvalues $\{\lambda_j(a)\}_{j=1}^\infty$ with corresponding eigenfunctions $\{\phi_j(\cdot, a)\}_{j=1}^\infty$ which represent an orthonormal basis of the weighted space $L^2(\mathbb{R}, \rho_\infty(\cdot, a))$, where $\rho_\infty(\cdot, a)$ is the probability density function of a Gaussian with zero mean and variance a.

The eigenvalue problem in this context reads: find $\phi_j(\cdot, a) : \mathbb{R} \to \mathbb{R}$ is sufficiently smooth with $\phi_j(\cdot, a) \neq 0$ and $\lambda_j(a) \in \mathbb{R}$ such that

$$-\mathcal{L}_a \phi_j(x, a) = \lambda_j(a)\phi_j(x, a), \tag{28}$$

that is

$$a x \phi_j'(x, a) - \sigma \phi_j''(x, a) = \lambda_j(a)\phi_j(x, a). \tag{29}$$

Question 9

Let us now characterize the eigenfunctions $\{\phi_j(\cdot, a)\}_{j=1}^{\infty}$ by the second-order recurrence relation

$$\begin{aligned}
\phi_0(x, a) &= 1, \\
\phi_1(x, a) &= x, \\
\phi_j(x, a) &= x\phi_{j-1}(x, a) - \frac{\sigma}{a}(j-1)\phi_{j-2}(x, a), \quad j \geq 2.
\end{aligned} \tag{30}$$

Then we can show by strong induction on j that the following holds:

$$\phi_j'(x, a) = j\phi_{j-1}(x, a), \quad j \geq 1. \tag{31}$$

The base cases $j = 0$ and $j = 1$ are trivial since $\phi_0'(x, a) = 0$ and $\phi_1'(x, a) = 1$. Assume now that (31) holds for j and $j-1$ for some $j \geq 2$, then using Leibniz's Rule and (30) we have that

$$\begin{aligned}
&\phi_{j+1}'(x, a) = \phi_j(x, a) + x\phi_j'(x, a) - \frac{\sigma}{\alpha} j\phi_{j-1}'(x, a) \overset{\text{ind. hp.}}{=} \\
&\phi_j(x, a) + xj\phi_{j-1}(x, a) - \frac{\sigma}{\alpha} j(j-1)\phi_{j-2}(x, a) = \phi_j(x, a) + j\phi_j(x, a) = (j+1)\phi_j(x, a).
\end{aligned}$$

Finally, we use (29) to compute the eigenvalues $\{\lambda_j(a)\}_{j=1}^{\infty}$. We claim that $\lambda_j(a) = ja, j \in \mathbb{N}$. The base cases $j = 0$ and $j = 1$ follow trivially from (29), because we have $\lambda_0(a)\phi_0(\cdot, a) = 0$ with $\phi_0(\cdot, a) \neq 0$ and $ax = \lambda_1(a)x$ for all $x \in \mathbb{R}$, respectively. For $j \geq 2$, by (31)

$$\begin{aligned}
ax\phi_j'(x, a) - \sigma\phi_j''(x, a) = axj\phi_{j-1}(x, a) - \sigma j(j-1)\phi_{j-2}(x, a) &= ja\left[x\phi_{j-1}(x, a) - \frac{\sigma}{\alpha}(j-1)\phi_{j-2}(x, a)\right] \\
&\overset{(30)}{=} ja\phi_j(x, a) \overset{(29)}{=} \lambda_j(a)\phi_j(x, a),
\end{aligned}$$

and the claim follows.

Question 10

We use now the previous results to construct a more efficient estimator for the unknown coefficient α. Let us consider $J \in \mathbb{N}$ and the first J eigenpairs $\{(\lambda_j(a), \phi_j(\cdot, a))\}_{j=1}^{J}$ together with some smooth functions $\psi_j : \mathbb{R} \to \mathbb{R}$ for $j = 1, \dots, J$. We then define the estimating function $G(a)$ as

$$G(a) = \frac{1}{N} \sum_{j=1}^{J} \sum_{n=0}^{N-1} \psi_j(\tilde{X}_n) \left(\phi_j(\tilde{X}_{n+1}; a) - e^{-\lambda_j(a)\Delta}\phi_j(\tilde{X}_n; a)\right), \tag{32}$$

from which we construct the estimator $\widetilde{\alpha}_N^{\Delta}$ of the drift coefficient as the solution of the non-linear equation $G(a) = 0$. Consider the simple case $J = 1$ with $\psi_1(x) = x$. Then (32) reduces to

$$G(a) = \frac{1}{N} \sum_{n=0}^{N-1} \widetilde{X}_n \left(\widetilde{X}_{n+1} - e^{-a\Delta}\widetilde{X}_n\right). \tag{33}$$

By ergodicity of the Ornstein–Uhlenbeck process we can again apply the ergodic theorem to the function $g_{a,\Delta} : \mathbb{R}^2 \to \mathbb{R}$, $g_{a,\Delta}(x, y) = x(y - e^{-a\Delta}x)$, since $g_{a,\Delta} \in C^{\infty}(\mathbb{R}^2, \mathbb{R})$. Then, the almost sure limit of (33) is given by

$$\mathcal{G}(a) := \lim_{N\to\infty} G(a) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} g_{a,\Delta}\left(\widetilde{X}_n, \widetilde{X}_{n+1}\right) = \mathbb{E}^{\mu_\infty}\left[X_t\left(X_{t+\Delta} - e^{-a\Delta}X_t\right)\right] = \mathcal{C}(t, t+\Delta) - e^{-a\Delta}C(t, t),$$

for t large enough so that at stationarity $\mathbb{E}^{\mu_\infty}[X_t] = 0$. Finally, by (17) we have

$$\mathcal{G}(a) = \frac{\sigma}{\alpha}e^{-\alpha\Delta} - \frac{\sigma}{\alpha}e^{-a\Delta}, \tag{34}$$

which is equal to 0 if and only if $a = \alpha$.

Question 11

We consider again the simple case $J = 1$ with the same $\psi_1(x) = x$ of *Question 10* to derive a closed form expression for the estimator $\tilde{\alpha}_N^\Delta$. Setting (33) equal to zero yields

$$\sum_{n=0}^{N-1} \tilde{X}_n \tilde{X}_{n+1} = e^{-a\Delta} \sum_{n=0}^{N-1} \tilde{X}_n^2. \tag{35}$$

We have that clearly $\sum_{n=0}^{N-1} \tilde{X}_n^2 > 0$ almost surely. On the other hand, we showed in *Question 5* that almost surely

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \tilde{X}_n \tilde{X}_{n+1} = \frac{\sigma}{\alpha} e^{-\alpha\Delta} > 0. \tag{36}$$

Therefore for N sufficiently large we can apply the logarithm on both sides and isolating the unknown a in (35) gives

$$\widetilde{\alpha}_N^\Delta := a = \frac{1}{\Delta} \log \left(\frac{\sum_{n=0}^{N-1} \tilde{X}_n^2}{\sum_{n=0}^{N-1} \tilde{X}_n \tilde{X}_{n+1}} \right). \tag{37}$$

To compute the almost sure limit of (37) as $N \to \infty$, by continuity of the logarithm we can pass the limit inside so that

$$\lim_{N\to\infty} \tilde{\alpha}_N^\Delta = \frac{1}{\Delta} \log \left(\lim_{N\to\infty} \frac{\sum_{n=0}^{N-1} \tilde{X}_n^2}{\sum_{n=0}^{N-1} \tilde{X}_n \tilde{X}_{n+1}} \right) = \log \left(\frac{\frac{\sigma}{\alpha}}{\frac{\sigma}{\alpha} e^{-\alpha\Delta}} \right) = \frac{1}{\Delta} \log \left(e^{\alpha\Delta} \right) = \alpha, \tag{38}$$

where once again we applied the ergodic theorem to the smooth functions defined at (21). We conclude that (37) is an asymptotically unbiased estimator of α independently of the sampling rate Δ.

Question 12

Figure 3 shows the behaviour of the estimator (37) against Δ for $\Delta = 2^{-i}, i = 0, \ldots, 7$. In this case the quality of the estimate of the drift coefficient α does not increase for a vanishing sampling rate Δ, as we have shown in *Question 11* that the latter is independent of the former. As in *Question 7*, (1) is solved numerically with the Euler-Maruyama method (12) using a step-size $\Delta t = 2^{-10}$.

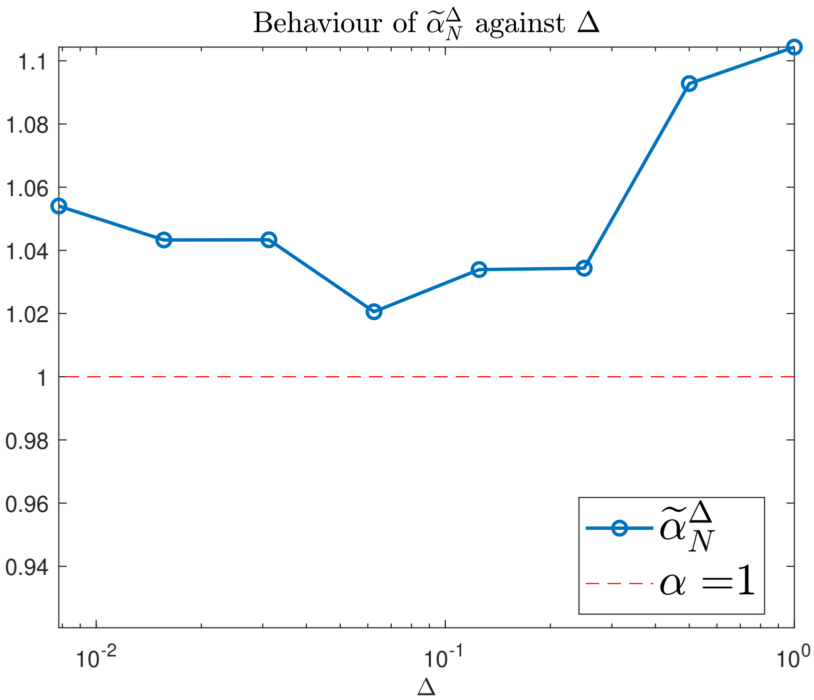

Figure 3: Evolution of the estimator (37) as a function of Δ: the red line shows the true coefficient α.

Question 13

We now set the sampling rate $\Delta = 1$ and we assume to be given some observations $\{\widetilde{X}_n^{(m)}\}_{n=0}^N$. By computing for each of $M = 10^4$ realizations of a Brownian motion the estimator (37), we observe numerically, as one can see in Figure 4, that the latter satisfies a central limit theorem (because it is unbiased, which was not holding true before), namely the quantity $\sqrt{N}\left(\tilde{\alpha}_N^\Delta - \alpha\right)$ is approximately distributed according to $\mu_{\alpha,\Delta} \sim \mathcal{N}(0, \Sigma)$, where $\Sigma = \left(e^{2\alpha\Delta} - 1\right)/\Delta^2$. Equivalently, the histogram of $\sqrt{N}\left(\{\tilde{\alpha}_N^{\Delta,(m)}\}_{m=1}^M - \alpha\right)$ approaches the density $\rho_{\mu_{\alpha,\Delta}}$ of the distribution $\mu_{\alpha,\Delta}$.

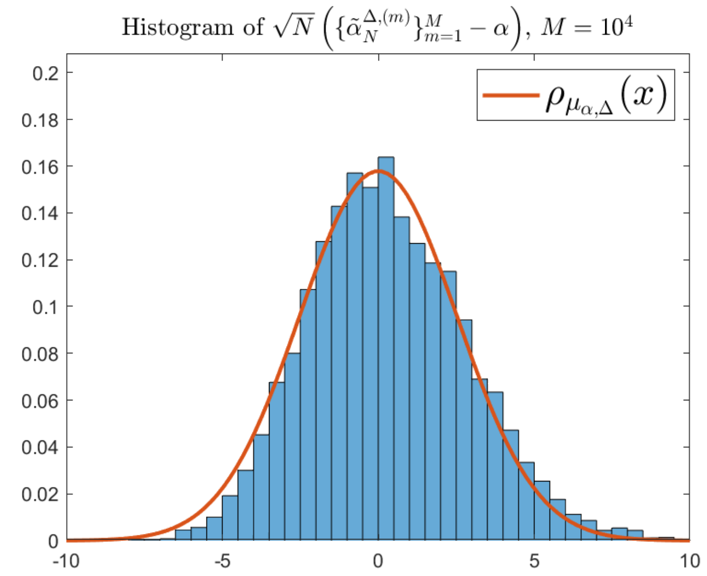

Figure 4: Histogram of $\sqrt{N}\left(\{\tilde{\alpha}_N^{\Delta,(m)}\}_{m=1}^M - \alpha\right)$.

Question 14

Let us now assume that also the diffusion coefficient σ is unknown. Note that, in the calculations in *Question 9*, the fact that the diffusion coefficient σ was known was never really exploited: then we can generalise the previous reasoning, substituting it with some parameter s. That is, starting from the recurrence relation

$$\begin{aligned} &\phi_0(x; a, s) = 1, \\ &\phi_1(x; a, s) = x, \\ &\phi_j(x; a, s) = x\phi_{j-1}(x, a) - \frac{s}{a}(j-1)\phi_{j-2}(x, a), \quad j \geq 2, \end{aligned} \tag{39}$$

one can show also in this case that $\phi_j'(x; a, s) = j\phi_{j-1}(x; a, s), j \geq 1$. For space reasons, we skip the details, but one can repeat exactly the calculations carried out in *Question 9*. Similarly, by defining a new differential operator $\mathcal{L}_{a,s}$ which is equal to $\mathcal{L}_a$ up to the substitution of σ with s, namely

$$\mathcal{L}_{a,s}\phi(x) = -ax\phi'(x) + s\phi''(x), \tag{40}$$

then the eigenvalue problem in this case reads $-\mathcal{L}_{a,s}\phi_j(x; a, s) = \lambda_j(a, s)\phi_j(x; a, s)$ where $\lambda_j(a, s) = ja, j \in \mathbb{N}$.

Consider this time a set $\{\Psi_j\}_{j=1}^J$ of vector-valued smooth functions $\Psi_j : \mathbb{R} \to \mathbb{R}^2$. Then, the estimating function reads

$$\mathbf{G}(a, s) = \frac{1}{N}\sum_{n=0}^{N-1}\sum_{j=1}^{J} \Psi_j(\widetilde{X}_n)\left(\phi_j(\widetilde{X}_{n+1}; a, s) - e^{-\lambda_j(a,s)\Delta}\phi_j(\widetilde{X}_n; a, s)\right), \tag{41}$$

and we construct an estimate for both the drift and the diffusion coefficient $(\widetilde{\alpha}_N^\Delta, \widetilde{\sigma}_N^\Delta)$ as the solution of the nonlinear two-dimensional system $\mathbf{G}(a, s) = 0$.

Let $J = 2$ and $\Psi_1(x) = \Psi_2(x) = \begin{pmatrix} x^2 & x \end{pmatrix}^T$. Then by (39) we have that $\phi_2(x; a, s) = x^2 - \frac{s}{a}$ and we can write explicitly the equation $\mathbf{G}(a,s)=0$:

$$\mathbf{G}(a,s)=\frac{1}{N}\sum_{n=0}^{N-1}\left(\widetilde{X}_{n+1}^2-\frac{s}{a}-e^{-2a\Delta}\left(\widetilde{X}_n^2-\frac{s}{a}\right)+\widetilde{X}_{n+1}-e^{-a\Delta}\widetilde{X}_n\right)\begin{bmatrix}\widetilde{X}_n^2\\ \widetilde{X}_n\end{bmatrix}= \tag{42}$$

$$\frac{1}{N}\sum_{n=0}^{N-1}\left(\widetilde{X}_{n+1}^2+\widetilde{X}_{n+1}-e^{-2a\Delta}\widetilde{X}_n^2-e^{-a\Delta}\widetilde{X}_n-\frac{s}{a}\left(1-e^{-2a\Delta}\right)\right)\begin{bmatrix}\widetilde{X}_n^2\\ \widetilde{X}_n\end{bmatrix}=0. \tag{43}$$

Question 15

Figure 5 and Figure 6 show respectively the estimators $(\widetilde{\alpha}_n^\Delta, \widetilde{\sigma}_n^\Delta)$ for α and σ plotted against the number of observations available n, for $n=2,\dots,N$ with $N=5*10^3$. The solution to (1) is again approximated numerically with the Euler-Maruyama method using step-size $\Delta t=0.01$ with final time $T=5*10^3$ (hence we set the sampling rate $\Delta=1$). As the limit of infinite data is approached the oscillations of the estimates become more and more negligible and the estimators converge to the true coefficients α and σ. Thus, in real world applications where we cannot tune the sampling rate, we can alternatively solve the non-linear system (42) to get a good estimation of the drift and diffusion coefficients.

Conclusions

We introduced the Ornstein-Uhlenbeck process as an ergodic process and the problem of doing inference on the drift and diffusion coefficients from discrete-time samples. We showed initially that exploiting the asymptotic behaviour of the sample paths of the process itself, we can easily construct estimators to approximate both coefficients. However, these classical estimators fail to give accurate estimates because their bias does not vanish asymptotically if the sampling rate is not taken sufficiently small. Hence for the applications, where one cannot choose the latter but observations are given, it is necessary to estimate the drift and diffusion coefficient by taking a different approach. In this direction we proposed a different estimator based on the eigenvalue problem associated with a modification of the differential operator $\mathcal{L}$ whose adjoint represents the Fokker-Plank equation solved by the invariant density of the process, where we replace the true (unknown) coefficients by parameters to be estimated. For the case in which only the drift α is unknown we derive an asymptotically unbiased estimator of α, independently of the sampling rate. Numerical results have been carried out to support the claim and showed that such an estimator satisfies a central limit theorem, thus converges to the true value of α in the limit of infinite data. Finally, we assessed the case in which also the diffusion coefficient σ is unknown; we showed numerically that by introducing another unknown parameter we can reduce the problem of estimating α and σ to solving a relatively easy non-linear system and this leads to converging estimators also for a non-vanishing fixed sampling rate.

Implementation aspects

The code part of the project is developed in MATLAB. To reproduce the obtained results, please run the script *RunAll.m* provided in the zip attached to the submission. All the implementations for the single questions are equally provided in the same folder.

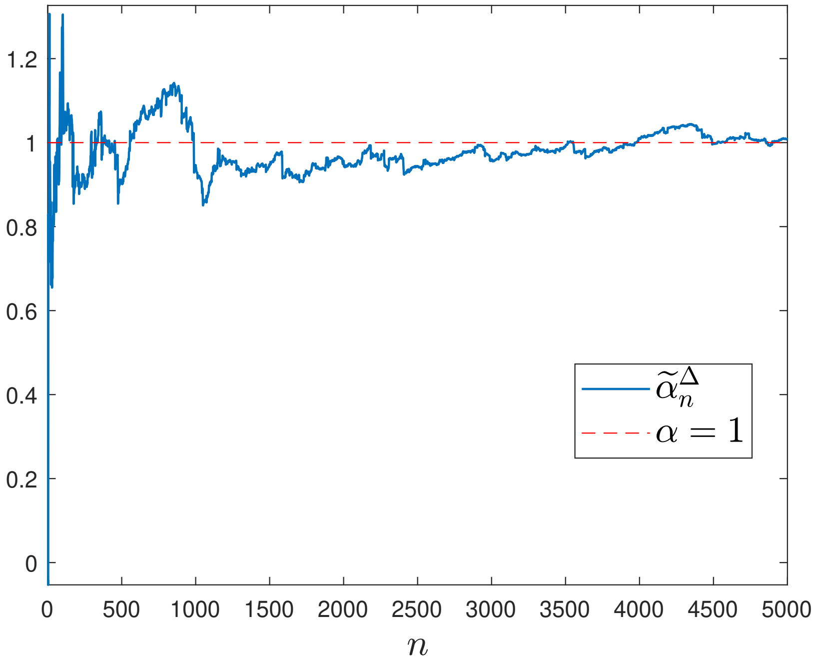

Figure 5: Evolution of the estimate of α given by (42) as a function of the number of observations n: the red line shows the true coefficient α.

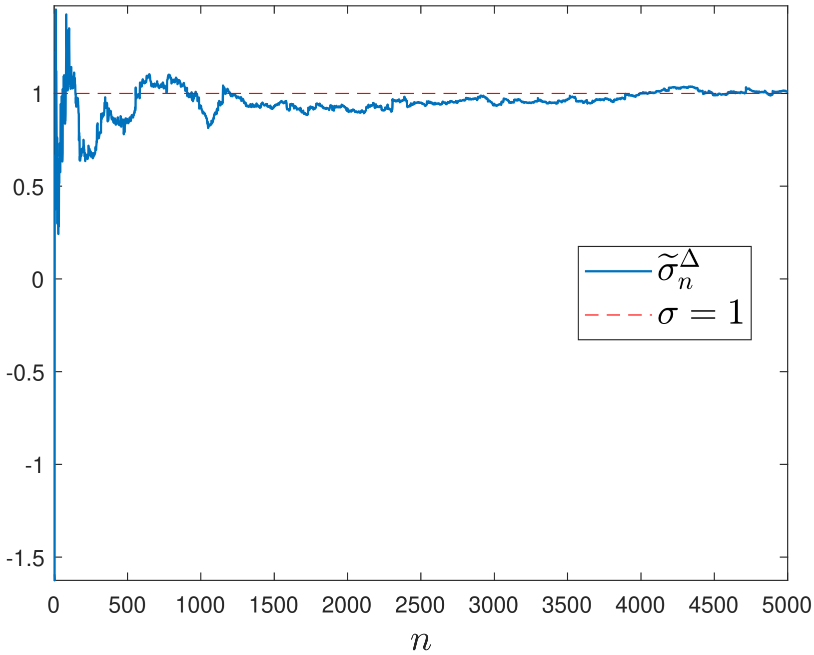

Figure 6: Evolution of the estimate of σ given by (42) as a function of the number of observations n: the red line shows the true coefficient σ.