Project for the course "Numerical Integration of Stochastic Differential Equations"

Parameter Inference for SDEs

Teacher: *Fabio Nobile*

Academic Year 2024/2025

We consider the problem of inferring the parameters of a stochastic differential equation (SDE) given discrete-time observations of a single trajectory. Parameter estimation for SDEs can be challenging for several reasons; when a single trajectory is observed, one has to rely on time invariance and ergodicity properties of the SDE to be able to infer the values of the parameters. Moreover, often the likelihood of the observation is not available in closed form and sophisticated simulation based approaches (e.g. Monte Carlo methods, approximate Bayesian computation, etc.) may be needed, which can be computationally intensive and time-consuming. Therefore, finding reliable and easy-to-compute estimators is of fundamental importance.
Let $T > 0$ be a final time and $X = (X(t), 0 \leq t \leq T)$ be the Ornstein-Uhlenbeck process which solves the Itô SDE

$$\mathrm{d}X(t) = -\alpha X(t)\mathrm{d}t + \sqrt{2\sigma}\mathrm{d}W(t), \quad X(0) = X_0 \in \mathbb{R}, \tag{1}$$

where $\alpha > 0$ is the drift coefficient and $\sigma > 0$ is the diffusion coefficient. The Ornstein-Uhlenbeck process is extensively used in physics and biology (see e.g. [2] and [1]), therefore a well-understanding of its stochastic properties is absolutely beneficial.

- **(Q1)**

  1. Derive the exact solution $X(t)$ of (1) when $X_0 \in \mathbb{R}$ is given. Then, derive its distribution μ_t and write its probability density function ρ_t.
  2. Is the solution of (1) a martingale? Justify your answer.
  3. The solution $X(t)$ of (1) has the property of being ergodic, i.e., its distribution μ_t tends for $t \to \infty$ to an invariant measure, which we denote by μ_∞ and which admits a probability density function ρ_∞. The function ρ_∞ is the unique solution of the stationary Fokker-Planck equation, a partial differential equation (PDE) which reads

     $$\begin{aligned} &\mathcal{L}^*\rho = 0, \text{ on } \mathbb{R}, \\ &\int_{\mathbb{R}} \rho(x)\mathrm{d}x = 1, \end{aligned} \tag{2}$$

     where the normalization condition is taken to ensure the uniqueness of the solution, and that ρ_∞ is indeed a probability density function. The differential operator $\mathcal{L}^*$ is the L^2-adjoint of the generator $\mathcal{L}$ of (1), which is defined as

     $$\mathcal{L}\varphi(x) = -\alpha x\varphi'(x) + \sigma\varphi''(x) \tag{4}$$

     for all sufficiently smooth functions $\varphi : \mathbb{R} \to \mathbb{R}$. In particular, $\mathcal{L}^*$ is defined by the relation

     $$\int_{\mathbb{R}} v(x)\mathcal{L}u(x)\mathrm{d}x = \int_{\mathbb{R}} u(x)\mathcal{L}^*v(x)\mathrm{d}x$$

     where $u, v : \mathbb{R} \to \mathbb{R}$ are smooth functions with compact support.
     Compute the operator $\mathcal{L}^*$ and write the stationary Fokker-Planck equation (2) explicitly for the SDE (1). Derive the invariant measure μ_∞ and verify that its probability density function ρ_∞ satisfies the stationary Fokker-Planck equation (2).
  4. Prove that the operator $\mathcal{L}$ is a non positive operator on

     $$H^2(\mathbb{R}, \rho_\infty) := \{f \in L^2(\mathbb{R}, \rho_\infty) \ : \ f', f'' \in L^2(\mathbb{R}, \rho_\infty)\},$$

     i.e., $\langle f, \mathcal{L}f\rangle_{L^2(\mathbb{R},\rho_\infty)} \leq 0$ for $f \in H^2(\mathbb{R}, \rho_\infty)$ and its null space is given by constant functions.
     Hint: You can assume that $C_0^\infty(\mathbb{R})$ is dense in $H^2(\mathbb{R}, \rho_\infty)$.

5. Sometimes, a positive constant term is added in the drift of (1), i.e.

$$\mathrm{d}X(t) = \alpha(p - X(t))\mathrm{d}t + \sqrt{2\sigma}\mathrm{d}W(t), \quad X(0) = X_0, \tag{3}$$

for $p > 0$. Equation (3) is a mean-reverting process. Explain the meaning of "mean-reverting". If $X(0) = p$, what can we say about $\mathbb{E}[X(t)]$? And if $X(0) \neq p$?

(Q2)

1. Compute an approximate solution of (1) with final time $T = 10^3$ employing the Euler-Maruyama method with a discretization step $h = 2^{-9}$ for $M = 10^4$ different realizations of the Brownian motion. Set the drift coefficient $\alpha = 1$ and the diffusion coefficient $\sigma = 1$. Verify numerically that the solution $X(T)$ at the final time is approximately distributed accordingly to the invariant measure μ_∞ by comparing the histogram of $\left\{X^{(m)}(T)\right\}_{m=1}^{M}$ with the density ρ_∞.

2. Which is the strong order of convergence of the Euler-Maruyama in this case? And the weak order?

- **(Q3)** Show that the covariance function of the process $X(t)$ at stationarity, i.e., when both $t, s \to \infty$, is given by

$$\mathcal{C}(t, s) = \frac{\sigma}{\alpha} e^{-\alpha|t-s|}. \tag{4}$$

- **(Q4)** Let $\Delta > 0$ be a sampling rate and assume that we are provided with discrete data in the form $\left\{\widetilde{X}_n\right\}_{n=0}^{N}$ where $N = T/\Delta$ and $\widetilde{X}_n = X_{n\Delta}$ for $n = 0, \dots, N$, i.e., equispaced observations from a single realization of the solution of (1) until time T. Assume furthermore that the coefficients α and σ are unknown and we aim to estimate them employing the data. In this context, approaches based on the classic estimators fail. In particular, the coefficient σ could be estimated approximating the quadratic variation of the path X with the available data, i.e., defining the estimator

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

Moreover, for the coefficient α, one could derive an approximate maximum likelihood estimator in the following form

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

However, these estimators do not converge to the exact coefficients in the limit of infinite data (and fixed Δ).

1. Show that the estimators $\widehat{\sigma}_N^\Delta$ and $\widehat{\alpha}_N^\Delta$ are asymptotically biased, by computing the almost sure limits

$$\sigma_\infty^\Delta = \lim_{N\to\infty} \widehat{\sigma}_N^\Delta \quad \text{and} \quad \alpha_\infty^\Delta = \lim_{N\to\infty} \widehat{\alpha}_N^\Delta$$

and verify that $\sigma_\infty^\Delta \neq \sigma$ and $\alpha_\infty^\Delta \neq \alpha$.
Hint. Since the solution $X(t)$ of (1) is ergodic, the data satisfy the following ergodic theorems for all functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^2 \to \mathbb{R}$ smooth enough

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f\left(\widetilde{X}_n\right) &= \mathbb{E}^{\mu_\infty}\left[f(X_0)\right], \quad \text{a.s.} \\ \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} g\left(\widetilde{X}_n, \widetilde{X}_{n+1}\right) &= \mathbb{E}^{\mu_\infty}\left[g(X_0, X_\Delta)\right], \quad \text{a.s.} \end{aligned} \tag{5}$$

where the superscript μ_∞ denotes the fact that X_0 and X_Δ are at stationarity, i.e., distributed according to the invariant measure μ_∞. These results yield an equality between time averages (on the left-hand side) and ensemble averages (on the right-hand side).

2. Verify that

$$\lim_{\Delta \to 0} \sigma_\infty^\Delta = \sigma \quad \text{and} \quad \lim_{\Delta \to 0} \alpha_\infty^\Delta = \alpha$$

which imply that these estimators provide good approximations of the true unknown coefficients when the sampling rate Δ is sufficiently small.

- **(Q5)**

1. Generate an "exact" trajectory of the SDE (1) with parameters $\alpha = 1$, $\sigma = 1$, and final time $T = 10^3$ (in practice, you could use the the Euler-Maruyama method with a discretization step $h = 2^{-10}$). Consider now different values of the sampling rate $\Delta = 2^{-i}$ for $i = 0, 1, \dots, 7$ and generate synthetic observations $\left\{ \widetilde{X}_n \right\}_{n=0}^N$ by evaluating the trajectory at time $n\Delta$. For each value of Δ compute now the estimators $\widehat{\alpha}_N^\Delta$ and $\widehat{\sigma}_N^\Delta$ and plot the results varying Δ together with the exact values of the coefficients α and σ.

2. Repeat the experiment using this time the Milstein Method. Comment the results that you obtain.

- **(Q6)** In concrete applications one is usually not allowed to choose the sampling rate Δ because the data are given, and therefore we cannot rely on the the previous estimators if Δ is too large. Let us for now focus only on the drift coefficient α and assume the diffusion coefficient σ to be known. A different approach consists in constructing estimating functions based on the eigenvalues and the eigenfunctions of the operator $-\mathcal{L}_a$ given in (4) and where the exact drift coefficient α is replaced by the parameter a. It can be shown that the operator $-\mathcal{L}_a$ has a countable set of distinct nonnegative eigenvalues $\{\lambda_j(a)\}_{j=0}^\infty$ which satisfy $0 \leq \lambda_0(a) < \lambda_1(a) < \cdots < \lambda_j(a) \nearrow +\infty$ and whose corresponding eigenfunctions $\{\phi_j(\cdot; a)\}_{j=0}^\infty$ form an orthonormal basis of $L^2(\mathbb{R}, \rho_\infty(\cdot\,; a))$ where $\rho_\infty(\cdot\,; a)$ is the invariant density distribution found in **(Q1)** where α is replaced by a.

1. State the eigenvalue problem $-\mathcal{L}_a \phi(x; a) = \lambda(a)\phi(x; a)$ in this context.

2. Verify that the eigenvalues are given by

$$\lambda_j(a) = ja, \quad j \in \mathbb{N}$$

and the corresponding eigenfunctions satisfy the following recurrence relation

$$\begin{aligned} &\phi_0(x; a) = 1 \\ &\phi_1(x; a) = x \\ &\phi_j(x; a) = x\phi_{j-1}(x; a) - \frac{\sigma}{a}(j-1)\phi_{j-2}(x; a), \quad j \geq 2 \end{aligned} \tag{6}$$

Hint. Prove and use the fact that the functions defined by the recurrence relation (6) satisfy

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

3. Let J be a positive integer and consider the first eigenpairs $\{(\lambda_j(a), \phi_j(a))\}_{j=1}^J$ and a set $\{\psi\}_{j=1}^J$ of smooth functions $\psi_j : \mathbb{R} \to \mathbb{R}$. Define the estimating function

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

and let the estimator $\widetilde{\alpha}_N^\Delta$ be the solution of the nonlinear equation $G(a) = 0$.

Set $J = 1$ and $\psi_1(x) = x$. Compute the almost sure limit

$$\mathcal{G}(a) = \lim_{N \to \infty} G(a)$$

and verify that $\mathcal{G}(a) = 0$ if and only if $a = \alpha$.

4. Give the analytical expression of the estimator $\widetilde{\alpha}_N^\Delta$ in the case $J = 1$ and $\psi_1(x) = x$ and show that it is strongly consistent, i.e., prove that

$$\lim_{N\to\infty} \widetilde{\alpha}_N^\Delta = \alpha, \qquad \text{a.s.},$$

independently of the sampling rate Δ.

- **(Q7)**

  1. Generate an "exact" trajectory of the SDE (1) with parameters $\alpha = 1$, $\sigma = 1$, and final time $T = 10^3$ (in practice, you could use the Euler-Maruyama method with a discretization step $h = 2^{-10}$). Consider now different values of the sampling rate $\Delta = 2^{-i}$ for $i = 0, 1, \dots, 7$ and generate synthetic observations $\left\{ \widetilde{X}_n \right\}_{n=0}^N$ by evaluating the trajectory at time $n\Delta$. For each value of Δ compute now the estimators found in point **(Q6)** and plot the results varying Δ together with the exact value of the drift coefficient α.

  2. Repeat the procedure of point **(Q7)-1** for $M = 10^4$ times, each time generating a new independent trajectory and synthetic observations with sampling rate $\Delta = 1$. For each repetition $m = 1, \dots, M$ of the experiment, compute the estimator $\widetilde{\alpha}_N^{\Delta,(m)}$. Verify numerically that the estimator satisfies a central limit theorem, i.e., that $\sqrt{N}\left(\widetilde{\alpha}_N^\Delta - \alpha\right)$ is approximately distributed as $\widetilde{\mu} = \mathcal{N}(0, \Sigma)$ where

  $$\Sigma = \frac{e^{2\alpha\Delta} - 1}{\Delta^2}$$

  by comparing the histogram of $\left\{ \sqrt{N}\left(\widetilde{\alpha}_N^{\Delta,(m)} - \alpha\right) \right\}_{m=1}^M$ and the density $\widetilde{\mu}$ (you may use a larger time step of $h = 2^{-9}$ to reduce the overall computational cost).

- **(Q8)** Let us now assume that also the diffusion coefficient σ is unknown. In this case we replace σ in the generator by a parameter s and therefore also the eigenvalues and the eigenfunctions can depend on both a and s. Moreover, we choose 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

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

and the couple of estimators $\left(\widetilde{\alpha}_N^\Delta, \widetilde{\sigma}_N^\Delta\right)$ is the solution of the two-dimensional nonlinear system $\mathbf{G}(a, s) = \mathbf{0}$.

  1. Set $J = 2$ and $\Psi_1(x) = \Psi_2(x) = \begin{pmatrix} x^2 & x \end{pmatrix}^\top$. Write explicitly the nonlinear system $\mathbf{G}(a, s) = \mathbf{0}$ in this case.

  2. Generate an "exact" trajectory of the SDE (1) with parameters $\alpha = 1$, $\sigma = 1$, and final time $T = 10^3$ (in practice, you could use the the Euler-Maruyama method with a discretization step $h = 2^{-10}$ or $h = 2^{-9}$ to lower the computational cost). Generate synthetic observations $\left\{ \widetilde{X}_n \right\}_{n=0}^N$ with sampling rate $\Delta = 1$ by evaluating the trajectory at time $n\Delta$ and compute the pair of estimators $\left(\widetilde{\alpha}_N^\Delta, \widetilde{\sigma}_N^\Delta \right)$ by solving the nonlinear system found in point **(Q8)**-1. Plot the evolution of the estimators $\widetilde{\alpha}_n^\Delta$ and $\widetilde{\sigma}_n^\Delta$ varying the number of available observations $n = 2, 3, \dots, N$, together with the exact values of the coefficients α and σ.
  Hint. In order to solve the nonlinear system you can use the function fsolve in Matlab or scipy.optimize.fsolve in PYTHON with initial value $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix}^\top$.

References

[1] Odd O Aalen and Håkon K Gjessing. "Survival models based on the Ornstein-Uhlenbeck process". In: *Lifetime data analysis* 10 (2004), pp. 407–423.

[2] Wolfgang Paul and Jörg Baschnagel. "Stochastic processes". In: *From Physics to Finance, Springer, Berlin* (1999).