Mini Project – Parameter Inference for SDEs

We consider the problem of inferring the parameters of a stochastic differential equation (SDE) given discrete-time observations of its solution.

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

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

where $\alpha > 0$ is the drift coefficient and $\sigma > 0$ is the diffusion coefficient.

(Q1) State 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.

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

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

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

$$\mathcal{L}\varphi(x) = -\alpha x \varphi'(x) + \sigma \varphi''(x), \tag{3}$$

for all sufficiently smooth functions $\varphi \colon \mathbb{R} \to \mathbb{R}$. 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 $u, v \colon \mathbb{R} \to \mathbb{R}$ are smooth functions with compact support.

(Q2) 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).

(Q3) Solve equation (1) with final time $T = 10^3$ employing the Euler–Maruyama method with a discretization step $h = 10^{-2}$ 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 $\{X^{(m)}(T)\}_{m=1}^M$ and the density ρ_∞.

(Q4) Show that the covariance function at stationarity, i.e., when both $t, s \to \infty$, is given by

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

Let $\Delta > 0$ be a sampling rate and assume that we are provided with discrete data in the form $\{\widetilde{X}_n\}_{n=0}^N$ 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

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

Moreover, for the coefficient α, one could discretize the *maximum likelihood estimator* and define

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

However, these estimators do not converge to the exact coefficients in the limit of infinite data.

(Q5) Show that the estimators $\widehat{\sigma}_N^\Delta$ and $\widehat{\alpha}_N^\Delta$ are asymptotically biased, i.e., compute the almost sure limits

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

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 $f\colon \mathbb{R} \to \mathbb{R}$ and $g\colon \mathbb{R}^2 \to \mathbb{R}$ smooth enough

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

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

(Q6) Verify that

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

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

(Q7) Solve equation (1) with final time $T = 10^3$ employing the Euler–Maruyama method with a discretization step $h = 10^{-3}$. Set the drift coefficient $\alpha = 1$ and the diffusion coefficient $\sigma = 1$. Assume to know discrete observations $\{\widetilde{X}_n\}_{n=0}^N$ for different values of the sampling rate $\Delta = 2^{-i}$ for $i = 0, 1, \dots, 7$. For each value of Δ compute 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.

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. 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 (3) 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 $0 \leq \lambda_0(a) < \lambda_1(a) < \cdots < \lambda_j(a) \uparrow +\infty$ and whose corresponding eigenfunctions $\{\phi_j(\cdot; a)\}_{j=0}^{\infty}$ form an orthonormal basis of the L^2 space weighted by the probability density function $\rho_\infty(\cdot; a)$ found in (Q2) and where α is replaced by a.

(Q8) State the eigenvalue problem $-\mathcal{L}_a \phi(x; a) = \lambda(a)\phi(x; a)$ in this context.

(Q9) Verify that the eigenvalues are given by

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

and the corresponding eigenfunctions satisfy the following recurrence relation

$$\begin{aligned}
&\phi_0(x; a) = 1, \\
&\phi_1(x; a) = x, \\
&\phi_j(x; a) = x\phi_{j-1}(x; a) - \frac{\sigma}{a}(j-1)\phi_{j-2}(x; a), \qquad j \geq 2.
\end{aligned} \tag{4}$$

Hint. Prove and use the fact that the functions defined by the recurrence relation (4) satisfy

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

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 $\psi_j \colon \mathbb{R} \to \mathbb{R}$. Define the estimating function

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

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

(Q10) 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 $a = \alpha$.

(Q11) 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 asymptotically unbiased, i.e., prove that

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

independently of the sampling rate Δ.

(Q12) Solve equation (1) with final time $T = 10^3$ employing the Euler–Maruyama method with a discretization step $h = 10^{-3}$. Set the drift coefficient $\alpha = 1$ and the diffusion coefficient $\sigma = 1$. Assume to know discrete observations $\{\widetilde{X}_n\}_{n=0}^N$ for different values of the sampling rate $\Delta = 2^{-i}$ for $i = 0, 1, \dots, 7$. For each value of Δ compute the estimator $\widetilde{\alpha}_N^\Delta$ found in point (Q11) and plot the results varying Δ together with the exact value of the drift coefficient α.

(Q13) Solve equation (1) with final time $T = 10^3$ employing the Euler–Maruyama method with a discretization step $h = 10^{-2}$ for $M = 10^4$ different realizations of the Brownian motion. Set the drift coefficient $\alpha = 1$ and the diffusion coefficient $\sigma = 1$. Assume to know discrete observations $\{\widetilde{X}_n^{(m)}\}_{n=0}^N$ with sampling rate $\Delta = 1$ and compute the estimator $\widetilde{\alpha}_N^{\Delta,(m)}$ for each realization of the Brownian motion. Verify numerically that the estimator satisfies a central limit theorem, i.e., that $\sqrt{N}(\widetilde{\alpha}_N^\Delta - \alpha)$ 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 $\{\sqrt{N}(\widetilde{\alpha}_N^{\Delta,(m)} - \alpha)\}_{m=1}^M$ and the density $\widetilde{\mu}$.

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 $\Psi_j \colon \mathbb{R} \to \mathbb{R}^2$. Then, the estimating function reads

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

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

(Q14) 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.

(Q15) Solve equation (1) with final time $T = 5 \cdot 10^3$ employing the Euler–Maruyama method with a discretization step $h = 10^{-2}$. Set the drift coefficient $\alpha = 1$ and the diffusion coefficient $\sigma = 1$. Assume to know discrete observations $\{\widetilde{X}_n\}_{n=0}^N$ with sampling rate $\Delta = 1$ and compute the couple of estimators $(\widetilde{\alpha}_N^\Delta, \widetilde{\sigma}_N^\Delta)$ by solving the nonlinear system found in point (Q14). 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 functions `fsolve` in MATLAB or `scipy.optimize.fsolve` in PYTHON with initial value $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix}^\top$.

Rules

The rules for the submission of your project are the following:

(*1*) Your report should address all the previous points, with clear references to the correspondence between the questions (Qx) and your answers.

(*2*) Submit your solution via email to `andrea.zanoni@epfl.ch` in an archive folder named `familyname.zip` (e.g., `zanoni.zip`) which should contain your report and a subfolder with your implementation. The deadline for submitting your solution is **Sunday 5 June 2022 at 23:59**.

(*3*) Your report must not exceed the length of **10 pages** (minimum font size 10pt, figures and references included), and should be typeset in LaTeX. The setting and results of your numerical experiments have to be included, and all questions above have to be addressed in your report.

(*4*) Your implementation should be clear and a set of easy-to-run numerical tests should be provided. The programming language is of your choice, but a `Matlab`, `C++/C` or `Python` implementation would be appreciated.

(*5*) Whenever you exploit results from existing literature, please cite your source accordingly in the bibliography.

(*6*) The project **is optional**. In case you submit a solution, your final grade F for the course will be computed as

$$F = \max\{0.8W + 0.2P, W\},$$

where W is the grade of the written exam and P is the grade of the project.