Computational Stochastic Processes
Problem Sheet 2

U. Vaes

with problems from G. Pavliotis, A. Duncan and S. Gomes

March 2020

You are free to return a selection of your work to me for marking. This is entirely optional and the mark will not count for assessment.

1 Stochastic differential equations

Problem 1 (Weak error)**.** Let us consider the weak Euler–Maruyama update defined by

$$X_{n+1}^{\Delta t} = X_n^{\Delta t} + b(X_n^{\Delta t})\,\Delta t + \sigma(X_n^{\Delta t})\,\sqrt{\Delta t}\,\xi_n,$$

where $\{\xi_n\}_{n=0}^{N-1}$ are i.i.d. discrete-valued random variables taking values 1 and -1 with equal probability. Show that the weak error, for geometric Brownian motion and for the observables $f(x) = x$ and $f(x) = x^2$, scales as Δt, i.e. show that

$$\left|\mathbb{E}\left[f(X_{N\Delta t}) - f(X_N^{\Delta t})\right]\right| \leq C\Delta t,$$

for a constant C independent of Δt.

Problem 2 (Variance reduction)**.** Consider the *overdamped Langevin* equation

$$\mathrm{d}X_t = -V'(X_t)\,\mathrm{d}t + \sqrt{2\beta^{-1}}\,\mathrm{d}W_t, \qquad X_0 = -1, \tag{1}$$

where $V(\cdot)$ is the double well potential:

$$V(x) = \frac{x^4}{4} - \frac{x^2}{2}.$$

1. By using a Monte Carlo simulation wth the Euler–Maruyama method, estimate the probability P defined by
$$P := \mathbb{P}[X_T > 0], \qquad T = 1.$$
2. By using importance sampling, implement an estimator for P with a lower variance.

Problem 3 (Maximum Likelihood estimator)**.** Consider the SDE

$$\mathrm{d}X_t = (\alpha X_t - \beta X_t^3)\,\mathrm{d}t + \mathrm{d}W_t.$$

Our objective is to derive maximum likelihood estimators for α and β for a given observation of the path X_t, $t \in [0, T]$.

1. Show that the log of the likelihood function is

$$\log L = \alpha B_1 - \beta B_3 - \frac{1}{2}\alpha^2 M_2 - \frac{1}{2}\beta^2 M_6 + \alpha\beta M_4,$$

   where

$$M_n\left(\{X_t\}_{t\in[0,T]}\right) = \int_0^T X_t^n \,\mathrm{d}t \quad \text{and} \quad B_n(\{X_t\}_{t\in[0,T]}) := \int_0^T X_t^n \,\mathrm{d}X_t.$$

2. Consequently show that the MLE for α and β are given by

$$\hat{\alpha} = \frac{B_1 M_6 - B_3 M_4}{M_2 M_6 - M_4^2} \quad \text{and} \quad \hat{\beta} = \frac{B_1 M_4 - B_3 M_2}{M_2 M_6 - M_4^2}.$$

Problem 4 (Nonlinear SDEs in population dynamics)**.** The following SDE appears in population dynamics:

$$\mathrm{d}X_t = -\mu X_t(1 - X_t)\,\mathrm{d}t - \sigma X_t(1 - X_t)\,\mathrm{d}W_t \tag{2}$$

1. Show that $X_t = 1$ is a fixed point for (2) and that linearizing about this fixed point we obtain the SDE for geometric Brownian motion.

2. Solve (2) numerically using the explicit Euler scheme for $\mu = -1$, $X_0 = 1.1$ and for $\sigma = .5, .6, .7, .8, .9$. Calculate numerically $\mathbb{E}|X_t - 1|^2$ and comment on the mean square stability of the explicit Euler scheme for the nonlinear SDE (2).

3. Solve (2) using the θ-Euler scheme with $\theta = \frac{1}{2}$. Investigate the mean square stability of this numerical scheme when applied to (2).

2 Markov chain Monte Carlo

Problem 5. Read Section 3.3 in the lecture notes, and show that π_{st} and π_{pt} are reversible for the Markov chains generated by the *simulated tempering* and *parallel tempering* algorithms, respectively. For the case of *parallel tempering*, consider for simplicity the case where $N = 2$. In both cases, assume that the MCMC schemes employed with probability α_0, in the notations of the lecture notes, are such that the associated transition distributions $p_i(x, y)$ satisfy detailed balance:

$$\pi_i(x)\,p_i(x, y) = \pi_i(y)\,p_i(y, x), \qquad \pi_i \propto \exp\left(-\frac{H(x)}{T_i}\right). \tag{3}$$

Here T_i denote positive constants, called *temperatures* by analogy with physical systems, and $H(x)$ denotes a smooth confining potential – a function such that $\lim_{|x|\to\infty} H(x) = +\infty$ and $\mathrm{e}^{-H(x)/T} \in L^1(X)$ for all $T > 0$. (This second condition guarantees that $\mathrm{e}^{-H(x)/T}$ defines a probability measure, up to the normalization constant.)

Problem 6 (Metropolis–Hastings)**.** In this question we explore the Metropolis–Hastings algorithm in a discrete state space.

1. Suppose we wish to sample from the binomial distribution

$$p_k = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}, \quad k \in \{0, 1, \ldots, n\},$$

   with parameters $n \in \mathbb{N}$ and $p \in (0, 1)$. Derive an independence sampler using a uniform distribution on $0, \ldots, n$ as proposal distribution.

2. The geometric probability distribution is

$$p_k = p(1-p)^{k-1}, \quad k \in \{1, 2, 3, \ldots\},$$

with parameter $p \in (0,1)$. Derive a simple symmetric random walk Metropolis Hastings algorithm to sample from this distribution.

In both cases implement the samplers in a programming language of your choice (with your chosen values of p and n), and confirm that they work by comparing the estimated means and variances with the known theoretical means and variances of these distributions.

Problem 7 (Metropolis-Hastings using deterministic transformations)**.** Suppose we wish to sample from a distribution $\pi(x)$. We consider sampling from this distribution using a Metropolis-Hastings algorithm in which the proposal distribution $q(y\,|x)$ is an equal mixture of two uniform distributions, as follows:

$$q(\cdot\,|\,x) = \frac{1}{2}\mathcal{U}((a-\varepsilon)x, (a+\varepsilon)x) + \frac{1}{2}\mathcal{U}(x/(a+\varepsilon), x/(a-\varepsilon),$$

where a is a constant greater than one, and $0 < \varepsilon < a-1$, for $x \geq 0$, and analogously (i.e. with bounds flipped) for $x < 0$.

1. Derive an expression for the MH acceptance probability for this proposal distribution.

2. Consider the limit of this algorithm as $\varepsilon \to 0$. Describe the resulting algorithm in this limit.

3. For the particular case where π is the following distribution

$$\pi(x) = \frac{2}{\pi}\frac{1}{(1+x^2)^2},$$

would either of the schemes proposed be efficient for sampling from π?

Problem 8 (Alternative acceptance probabilities)**.** While the Metropolis-Hastings acceptance probability is by far the most widely used acceptance probability, there are several other choices. One alternative rule is the *Barker rule*:

$$\alpha(x,y) = \left(1 + \frac{\pi(x)q(y|x)}{q(x|y)\pi(y)}\right)^{-1}.$$

1. Show that the Metropolis Hastings scheme using this acceptance rule is reversible with respect to π, in the case of a continuous state space.

2. Using a proposal $q(\cdot|x) \sim \mathcal{N}(x, \delta^2)$, implement the Barker-rule based scheme, as well as a standard RWMH for a standard Gaussian target distribution π. Plotting the acceptance rate averaged over time, how do they compare?

3. Compare the performance in terms of effective sample size.

Problem 9. Consider the independence sampler, i.e. of the Metropolis–Hastings algorithm with proposal $q(\cdot|x) = g(\cdot)$. Show that, if $\pi(x) \leq M\,g(x)$ for some constant M, then the probability of an acceptance from state x is bounded from below by $\frac{1}{M}$.

Problem 10 (MALA algorithm and preconditioning)**.** Let $\pi(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$ be a probability density function and suppose that we want to calculate the expectation

$$\mathbb{E}_\pi f = \int_{\mathbb{R}^d} f(\mathbf{x})\pi(\mathbf{x})\, \mathrm{d}\mathbf{x}, \tag{4}$$

where $f(\mathbf{x})$ is an arbitrary function such that $\mathbb{E}_\pi f < +\infty$.

1. Explain how you can use a diffusion process of the form

$$d\mathbf{X}_t = \nabla \log \pi(\mathbf{X}_t)\, dt + \sqrt{2}\, d\mathbf{W}_t, \qquad \mathbf{X}_0 \sim \rho_0, \tag{5}$$

   where $\mathbf{W}_t$ denotes standard Brownian motion on $\mathbb{R}^d$ in order to calculate $\mathbb{E}_\pi f$.

2. Let $\pi(x)$ be a bivariate normal distribution $\pi \sim \mathcal{N}(\mu, \Sigma)$, where

$$\mu = \begin{bmatrix} 3 \\ 6 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 2 & 0.5 \\ 0.5 & 1 \end{bmatrix}.$$

   (a) Write down $\pi(x)$, $\log \pi(x)$ and $\nabla \log \pi(x)$.

   (b) Use the above calculations to sample from π using the MALA distribution.

   (c) Compute an estimator for $I = \mathbb{E}(f(\mathbf{X}))$, where $\mathbf{X} = (X, Y) \sim \pi$ and $f(\mathbf{x}) = x^3 + y^2$.

   (d) Track the acceptance rate for different time steps δ.