Introduction to Uncertainty Quantification Coursework 2023/24

The deadline for this coursework is Friday 8th December at 15:00 GMT. A copy must be submitted to Blackboard by this time (instructions on how to do this will be found on Blackboard in the coursework folder). To avoid any connectivity issues, please make sure to submit the coursework well before the deadline. The coursework should take on average no more than 15 hours to complete.

Marks will be awarded for presentation of numerical results, as well as for the correctness and interpretation of the results. Clearly state what parameter values you use in your numerical experiments (chosen discretisation parameters, number of samples, etc). Please provide all of the code that you have created in an appendix at the end of your report, and ensure that the code is commented and as readable as possible. Prof Cotter will be available for help with any coding issues during the lab sessions in weeks 9 and 11. Please note that you must work on this project alone, and that there are serious penalties for plagiarism. This coursework is worth 20% of your final mark.

---

1. **Monte Carlo** In cell biology, there are many examples of oscillatory chemical reaction networks which help to regulate important functions. For example, the Circadian clock in many organisms is a system of reactions which oscillates with a period of approximately 24 hours, helping to regulate processes according to the time of day. Another important oscillatory system within our cells is that of the cell cycle, which regulates how frequently our cells divide. The cell cycle is very well studied, and there are many models which approximate the dynamics that is observed in real cells. The following is a simple model by Ferrell et al[1], which involves 3 chemical species with concentrations x, y, z respectively. Under certain parameter regimes this system exhibits oscillations which replicate those found in real cells.

$$\begin{aligned} \frac{dx}{dt} &= \alpha_1 - \beta_1 x \frac{z^n}{K^n + z^n} \\ \frac{dy}{dt} &= \alpha_2(1-y)\frac{x^n}{K^n + x^n} - \beta_2 y, \\ \frac{dz}{dt} &= \alpha_3(1-z)\frac{y^n}{K^n + y^n} - \beta_3 z. \end{aligned} \tag{1}$$

   (a) Derive the forward Euler approximation of (1), and implement it on a computer. Use this implementation to approximate the solution of this ODE with $\alpha_1 = 0.1$, $\alpha_2 = 3$, $\alpha_3 = 3$, $\beta_1 = 3$, $\beta_2 = 1$, $\beta_3 = 1$, $K = 0.5$, $n = 8$, up to time $T = 10$, with initial condition $[x(0), y(0), z(0)]^\top = [0.1, 3, 3]^\top$, plotting the approximation with $\Delta t = 10^{-5}$. Plot the absolute value of the discrepancy between this approximation and the approximations obtained when using $\Delta t = 10^{-1}$, $\Delta t = 10^{-2}$, $\Delta t = 10^{-3}$, and $\Delta t = 10^{-4}$. Comment on the decay of the discrepancy as the timestep decreases. Hint: Consider using the log scale.

[1] Ferrell, James E., Tony Yu-Chen Tsai, and Qiong Yang. "Modeling the cell cycle: why do certain circuits oscillate?." Cell 144, no. 6 (2011): 874-885.

[5 marks]

(b) Suppose that the initial condition is itself uncertain, and can be modelled by the following random variable:

$$\begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix} \sim \{\omega \sim \mathcal{N}([0.1; 3; 3]^\top, \sigma^2 I) | \omega_1, \omega_2, \omega_3 \geq 0\} \tag{2}$$

where $I \in \mathbb{R}^{3\times 3}$ is the identity matrix, and $\sigma = 0.1$. In other words, the initial condition is a normally distributed random variable conditioned on all components being non-negative. This distribution can very simply be sampled from using rejection sampling; we sample from the normal distribution indicated but only use samples whose components are all non-negative.

We now wish to ascertain the distribution of the approximation of the ODE at time $T = 10$, subject to this uncertainty in the initial condition. For simplicity we focus only on the distributions of x and y. Implement a Monte Carlo algorithm in order to estimate $\mathbb{E}(x(10))$ and $\mathbb{E}(y(10))$ using the Euler approximation from part (a). Make appropriate plots to visualise the joint and marginal distributions of the random variables $x(10)$, $y(10)$ and estimated using the Monte Carlo samples, using the timestep $\Delta t = 10^{-2}$.

You may use the m-file "hist2d.m" to create your visualisations of the joint distribution, which can be found in the coursework folder on Blackboard[2].

Using your numerical results, state estimates for $\mathbb{E}(x(10))$, $\mathbb{E}(y(10))$ and $\mathbb{E}(z(10))$. How many samples would you need to have a standard error less than 10^{-3} for estimators of $\mathbb{E}(x(10))$, $\mathbb{E}(y(10))$ and $\mathbb{E}(z(10))$ when $\Delta t = 10^{-2}$? Justify your answers.

[5 marks]

2. **Numerical approximation of SDEs.** In reality cells do not behave like an ODE as the small scale of the reactions means that small random fluctuations can have a big impact on the dynamics. We consider a stochastic version of the cell cycle model, given by:

$$\begin{aligned} dX &= \left(\alpha_1 - \beta_1 X \frac{Z^n}{K^n + Z^n}\right) dt + \gamma X dB_t^1, \\ dY &= \left(\alpha_2 (1 - Y) \frac{X^n}{K^n + X^n} - \beta_2 Y\right) dt + \gamma Y dB_t^2, \\ dZ &= \left(\alpha_3 (1 - Z) \frac{Y^n}{K^n + Y^n} - \beta_3 Z\right) dt + \gamma Z dB_t^3. \end{aligned} \tag{3}$$

where B_t^i for $i \in \{1, 2, 3\}$ are independent standard Brownian motions, and where the parameters are the same as the ODE (1), with the addition of $\gamma = 0.1$.

(a) Derive the Euler-Maruyama method for the approximation of realisations of this SDE. Implement this method in any computer language you wish. Simulate and

[2] Usage: hist2d(X,Y,N,'pdf'), where X and Y are arrays containing your samples, and N is the number of bins in each direction (I recommend 100, but make sure you have enough samples).

plot 5 approximate realisations of this SDE, with initial condition $[X, Y, Z]^\top = [0.3, 0.1, 0.1]^\top$, up to time $T = 10$ with $\Delta t = 10^{-2}$. Comment on the variation that you observe in the different realisations, in comparison with the ODE trajectories from Q1.

[5 marks]

(b) Implement a Monte Carlo sampler to estimate the values of $\mathbb{E}(X)$, $\mathbb{E}(Y)$ and $\mathbb{E}(Z)$ at $T = 10$, using the Euler-Maruyama approximation with $\Delta t = 10^{-2}$. Make appropriate plots to visualise the marginal and joint distributions for the random variables X_T, Y_T and Z_T at $T = 10$. Using your numerical results, state estimates for $\mathbb{E}(X_T)$, $\mathbb{E}(Y_T)$ and $\mathbb{E}(Z_T)$. How many samples would you need in order for all of the estimators to have a standard error less than 10^{-3}? Justify your answer.

[5 marks]

[TOTAL: 20 marks]