Homework 1
Numerical methods for random partial differential equations: hierarchical approximation and machine learning approaches

April 14, 2023

IMPORTANT: Please refer to the course syllabus and the course outline for what an acceptable level on a handed in homework is.

1 MC under exact sampling

1.1 Objective

Hands-on experience on Monte Carlo sampling methods under the idealised assumptions of exact sampling. Tasks include the numerical versification and computable confidence regions. The idea is to connect theory and practical aspects and experiment with the methods.

1.2 Problem

We are interested in calculating the following quantity

$$g = \int_{[0,1]^N} f(\mathbf{x})\, d\mathbf{x}$$

for some given function f. We try different cases, which mainly differ by the degree of regularity of the function f. In this exercise f is known and the problem is simply integrating the function in a high dimension N.

1.3 Model problems

We look at different examples of f for some real constants $\{c_n, w_n\}_{n=1}^N$ taken from

1. Oscillatory: $f(\mathbf{x}) = \cos\left(2\pi w_1 + \sum_{n=1}^N c_n x_n\right)$, with $c_n = 9/N$, $w_1 = \frac{1}{2}$.
   Exact solution:
   $$\int_{[0,1]^N} f(\mathbf{x})\, d\mathbf{x} = \Re\left(e^{i2\pi w_1} \prod_{n=1}^N \frac{1}{ic_n}(e^{ic_n} - 1)\right)$$
   (here i denote the imaginary unit and $\Re(z)$ the real part of $z \in \mathbb{C}$)

2. Product peak: $f(\mathbf{x}) = \prod_{n=1}^N \left(c_n^{-2} + (x_n - w_n)^2\right)^{-1}$, with $c_n = 7.25/N$ and $w_n = \frac{1}{2}$.
   Exact solution:
   $$\int_{[0,1]^N} f(\mathbf{x})\, d\mathbf{x} = \prod_{n=1}^N c_n\left(\arctan(c_n(1 - w_n)) + \arctan(c_n w_n)\right)$$

3. Corner peak: $f(\mathbf{x}) = \left(1 + \sum_{n=1}^N c_n x_n\right)^{(-N+1)}$
   with $c_n = 1.85/N$.

4. Gaussian: $f(\mathbf{x}) = \exp\left(-\sum_{n=1}^N c_n^2 (x_n - w_n)^2\right)$, with $c_n = 7.03/N$ and $w_n = \frac{1}{2}$.
   Exact solution:
   $$\int_{[0,1]^N} f(\mathbf{x})\, d\mathbf{x} = \prod_{n=1}^N \frac{\sqrt{\pi}}{2c_n}\left(\operatorname{erf}(c_n(1 - w_n)) + \operatorname{erf}(c_n w_n)\right)$$

5. Continuous: $f(\mathbf{x}) = \exp\left(-\sum_{n=1}^N c_n |x_n - w_n|\right)$,with $c_n = 2.04/N$ and $w_i = \frac{1}{2}$.
   Exact solution:
   $$\int_{[0,1]^N} f(\mathbf{x})\, d\mathbf{x} = \prod_{n=1}^N \frac{1}{c_n}\left(2 - e^{-c_n w_n} - e^{-c_n(1 - w_n)}\right)$$

6. Discontinuous: $f(\mathbf{x}) = \begin{cases} 0 & \text{if } x_1 > w_1 \text{ or } x_2 > w_2 \\ \exp\left(\sum_{n=1}^N c_n x_n\right) & \text{otherwise} \end{cases}$, with $c_i = 4.3/N$, $w_1 = \frac{\pi}{4}$ and $w_2 = \frac{\pi}{5}$.
   Exact solution:
   $$\int_{[0,1]^N} f(\mathbf{x})\, d\mathbf{x} = \frac{1}{\prod_{n=1}^N c_n}(e^{c_1 w_1} - 1)(e^{c_2 w_2} - 1)\prod_{n=3}^N (e^{c_n} - 1)$$

1.4 Numerical tasks

1. Write a Monte Carlo estimator to numerically estimate (some of) these integrals. Use `MATLAB`'s function `rand(N,M)` to generate M random points in the hypercube $[0,1]^N$.

2. Estimate the error using the Central Limit Theorem. Plot the exact error and error estimates versus the number of samples used. Estimate the convergence rate. Report the square of the coefficient of variation for all cases.

   Do your computations at least for $N = 2$ and $N = 20$.

3. Repeat the previous problem using Berry-Esseen theorem.

4. Consider the approximation of the following integral, for values of $K = 3$ and $K = 6$
$$\frac{1}{2\pi}\int_{\mathbb{R}^2} \max\left(e^{x_1} + e^{x_2} - K, 0\right) e^{\frac{-(x_1^2+x_2^2)}{2}} dx_1 dx_2$$
   Use importance sampling based on the shift-dilation technique, and compare it to the standard Monte Carlo estimator. Use Central Limit Theorem to control the error.

2 MC for PDEs using approximate sampling

2.1 Objective

This exercise gives hands-on experience on Monte Carlo sampling for approximate (discretized) sampling. It also makes the student work on sampling Gaussian random fields and thus experimenting with Karhunen-Loève and Fourier expansions.

2.2 Problem

Next we look at the solution of a one dimensional boundary value problem
$$-(a(x,\omega)u'(x,\omega))' = \underbrace{4\pi^2\cos(2\pi x)}_{=:f(x)}, \text{ for } x \in (0,1)$$
with $u(0,\cdot) = u(1,\cdot) = 0$. In this case, we are interested in computing the expected value of the following QoI
$$Q(u(\omega)) = \int_0^1 u(x,\omega)dx.$$
We use an $I+1$ uniform grid $0 = x_0 < x_1 < \ldots < x_I = 1$ on $[0,1]$ with uniform spacing $h = x_i - x_{i-1} = 1/I$, $i = 1,\ldots,I$. Using this grid we can build a piecewise linear FEM approximation,
$$u_h(x,\omega) = \sum_{i=1}^{I-1} \mathbf{u}_{h,i}(\omega)\varphi_i(x),$$

yielding a tridiagonal linear system for the nodal values, $A(\omega)\mathbf{u}_h(\omega) = F$, with

$$\begin{aligned} A_{i,i-1}(\omega) &= -\frac{a(x_{i-1/2},\omega)}{h^2} \\ A_{i,i}(\omega) &= \frac{a(x_{i-1/2},\omega) + a(x_{i+1/2},\omega)}{h^2} \\ A_{i,i+1}(\omega) &= -\frac{a(x_{i+1/2},\omega)}{h^2} \end{aligned}$$

and

$$F_i = f(x_i).$$

Here we used the notation $x_{i+1/2} = \frac{x_i+x_{i+1}}{2}$ and $x_{i-1/2} = \frac{x_i+x_{i-1}}{2}$.

The integral in $Q(u_h)$ can be then computed exactly by a trapezoidal method, yielding

$$Q(u_h) = h\sum_{i=1}^{I-1} \mathbf{u}_{h,i}.$$

2.3 Models for the random coefficient a.

We look at two different models for $a(x,\omega)$, namely

2.3.1 Model 1: Piecewise constant coefficient

$$a(x,\omega) = 1 + \sigma \sum_{n=1}^{N} Y_n(\omega)\mathbb{I}_{[\hat{x}_{n-1},\hat{x}_n]}(x),$$

with equispaced nodes $\hat{x}_n = \frac{n}{N}$ for $0 \leq n \leq N$ and i.i.d. uniform random variables $Y_n \sim U([-\sqrt{3},\sqrt{3}])$. Consider different uniform mesh refinements, i.e. values of $I = N\,2^\ell$, $\ell \geq 0$, and different number of input random variables $N = 10$, $N = 20$ and $N = 40$. Remember to choose σ ensuring coercivity, namely $1 - \sigma\sqrt{3} > 0$.

2.3.2 Model 2: Log-Normal

$$a(x,\omega) = \exp(\kappa(x,\omega)),$$

where $\kappa(x,\omega)$ is a stationary Gaussian random field with mean zero and the Matérn covariance function

$$C(x,y) = \sigma^2 \frac{1}{\Gamma(\nu)2^{\nu-1}} \left(\sqrt{2\nu}\frac{|x-y|}{\rho}\right)^{\nu} K_\nu \left(\sqrt{2\nu}\frac{|x-y|}{\rho}\right),$$

where Γ is the gamma function and K_ν is the modified Bessel function of the second kind. We look at the following special cases of C

$$\begin{aligned} \nu = 0.5, \quad C(x,y) &= \sigma^2 \exp\left(-\tfrac{|x-y|}{\rho}\right) \\ \nu = 1.5, \quad C(x,y) &= \sigma^2 \left(1 + \tfrac{\sqrt{3}|x-y|}{\rho}\right) \exp\left(-\tfrac{\sqrt{3}|x-y|}{\rho}\right) \\ \nu = 2.5, \quad C(x,y) &= \sigma^2 \left(1 + \tfrac{\sqrt{5}|x-y|}{\rho} + \tfrac{\sqrt{3}|x-y|^2}{\rho^2}\right) \exp\left(-\tfrac{\sqrt{5}|x-y|}{\rho}\right) \\ \nu \to \infty, \quad C(x,y) &= \sigma^2 \exp\left(-\tfrac{|x-y|^2}{2\rho^2}\right). \end{aligned}$$

Moreover we choose $\rho = 0.1$ and $\sigma^2 = 2$.

2.4 Possible numerical approximations of Log-Normal random fields

1. Given a grid, consider multivariate Gaussian random vector X_h with mean μ_h covariance matrix Σ_h consisting of the Gaussian random field's covariance function evaluated on the grid. Let A such that $\Sigma_h = AA^T$ (e.g., via SVD or Cholesky or spectral decomposition). Sampling of X_h then via the representation $X_h = \mu_h + A\,Z$, where $Z \sim \mathcal{N}(0, I)$.

2. We expand $\kappa(x,\omega)$ using truncated Karhunen-Loève expansion with N terms

$$\kappa(x,\omega) \approx \sum_{n=1}^{N} \sqrt{\lambda_n} Y_n(\omega) e_n(x)$$

Where $\{Y_k\}$ is a set of i.i.d. standard Gaussian. To find the eigenfunctions and eigenvalues of C we solve the eigenvalue problem

$$\int_0^1 C(x,y) e_n(y)\, \mathrm{d}y = \lambda_n e_n(x) \tag{1}$$

We do this by discretizing C as a matrix by evaluating the function $C(x_{i+\frac{1}{2}}, x_{j+\frac{1}{2}})$ over the grid $\{x_{i+\frac{1}{2}}\}_{i=0}^{I-1} \times \{x_{i+\frac{1}{2}}\}_{i=0}^{I-1}$ with $N \leq I$. Then use `MATLAB`'s function `eig()`. This discretization corresponds to a piecewise constant FEM approximation to the eigenvalue problem (1). Make sure that the computed eigenfunctions have norm 1 with respect to the continuous $L^2([0,1])$ norm.

3. Now, with the same covariance field as before, use a Fourier expansion of a periodic extension of it, namely

$$Cov_\kappa^{\#}(x-y) = \sum_{n=0}^{\infty} \kappa_n^2 \cos\left(\frac{n\pi(x-y)}{L_p}\right).$$

Take $L_p = 2$ for instance. As a result, the stationary random field κ admits the exact representation in $[0, L_p]$

$$\kappa(\omega, x) = \sum_{n=0}^{\infty} \kappa_n \left(y_n(\omega) \cos\left(\frac{n\pi x}{L_p}\right) + z_n(\omega) \sin\left(\frac{n\pi x}{L_p}\right)\right)$$

and

- $\mathbb{E}[y_n] = \mathbb{E}[z_n] = 0$
- $Var[y_n] = Var[z_n] = 1$
- $\{y_n, z_n\}_n$ uncorrelated.

Here take iid y_n, z_n with standard Normal distribution.

4. **Task:** Plot the spectral contribution of C of Model 2 (Log-Normal) to the Karhunen-Loève expansion, i.e. the values of $\sqrt{\lambda_n}\|e_n\|_\infty$, and the size of the Fourier terms κ_n for the different values of ν given above.

2.5 Numerical approximation of $\mathbb{E}[Q(u)]$.

For the following questions (1) and (2), use Model 1 or Model 2 for $a(x, \omega)$. For question (3), use Model 2 for $a(x, \omega)$. When working with Model 2, use the previous Task to deduce a relevant choice of N. Moreover, pick two different values of the parameter ν.

1. Repeat Exercise 1 for the approximation of $\mathbb{E}[Q(u_h)]$. To this end, compute it using plain Monte Carlo sampling. Estimate the statistical error using the Central Limit Theorem. Estimate the bias error using Richardson extrapolation. Choose the number of samples and the mesh size h such that the relative error is less than 10%.

2. Use a coarser discretization of the problem (say with double the gridsize h) with the same realization of $\{Y_k\}$ as a control variate for the approximation of $\mathbb{E}[Q(u_h)]$ and check the variance reduction in Monte Carlo sampling. Does it pay to use it?

3. Approximate the probability $P(Q(u) < K)$ using importance sampling. Compute with different values of K, namely $K = -5$, $K = -10$, and $K = -20$. Estimate the statistical error with the Central Limit Theorem choosing the number of samples so that the relative error is less than 10%. Use the same mesh size as in question 1.