TU Chemnitz – Professur Numerische Mathematik 18. Mai 2016
Prof. Dr. Oliver Ernst
Dipl.-Math. Björn Sprungk

Mathematische Methoden der Unsicherheitsquantifizierung

Sommersemester 2016

Exercise Sheet 4: Monte Carlo methods with control variates

Another method of variance reduction for Monte Carlo simulations is the use of *control variates*:

$$\widetilde{Y}_b := Y - b(Z - \mathbf{E}Z), \qquad b \in \mathbb{R} \tag{1}$$

leading to the following Monte Carlo estimator

$$\widetilde{\mu}_{N,b} := \frac{1}{N}\sum_{n=1}^{N} \widetilde{Y}_{b,n}$$

where the $\widetilde{Y}_{b,n}$ are stochastically independent copies of $\widetilde{Y}_b$.

Exercise 1

The lecture proves an equation for the optimal value b^*. Find an estimator $B_N = g(Y_1, Z_1, \ldots, Y_N, Z_N)$ of b^* and check the consistency of

$$\widetilde{\mu}_N := \frac{1}{N}\sum_{n=1}^{N} \left[Y_n - B_N(Z_n - \mathbf{E}Z)\right],$$

or in other words show

$$\lim_{N\to\infty} \widetilde{\mu}_N = \mathbf{E}Y \qquad \mathbb{P}\text{-a.s..}$$

Exercise 2

Instead of estimating the optimal value b^*, the problem is often simplified and $b = 1$ is chosen. Furthermore, in many applications the random variable Y is given as $Y = f(X)$ for some $f : \mathbb{R}^d \to \mathbb{R}$ and a d dimensional real-valued random variable X. It is now possible to use $Z = \widetilde{f}(X)$ as a control variate with $\widetilde{f} : D \to \mathbb{R}$ being an approximation of f.

(a) Deduce an optimal approximation $\widetilde{f}$ of f – optimal in the sense of minimal estimation error – in case of $b = 1$. (*Hint:* It turns out to be a best approximation in a certain Hilbert space.)

(b) We will now consider the class of affine approximations $\widetilde{f}(X) = aX + c$. By which factor is the sample variance of $\widetilde{\mu}_N$ reduced when choosing $Z = aX + c$ and $b = 1$ as control variables compared to the sample variance of the standard Monte Carlo estimator $\widehat{\mu}_N$?
Furthermore, show that the best L^2 approximation of f among all affine functions is also the best control variate in this class with respect to the sample variance.

(c) Let $X \sim \mathrm{U}[0,1]$, $f(x) = \exp(x)$ and also $f(x) = x(1-x)$. By which factor is the sample variance reduced when choosing $\widetilde{f}(X)$ and $b = 1$, where $\widetilde{f}$ is the linear interpolation of f at points $x_0 = 0$ and $x_1 = 1$, as control variates?

Exercise 3

We want to estimate π with the following random experiment. We simulate the random variable

$$Y := 4 \cdot \mathbf{1}_{[0,1]}(U_1^2 + U_2^2),$$

where $U_1, U_2 \sim \mathrm{U}[0,1]$ i.i.d. There holds $\mathbf{E}[Y] = \pi$. (Why?)
Further, we want to employ the following control variates

$$Z_c := \mathbf{1}_{[0,c]}(U_1 + U_2), \qquad c \in (0,2).$$

(a) Illustrate the meaning of Y and Z_c in a grafical way. Find an explicit expression for $\mathbf{E}[Z_c]$.

(b) Assume that we will always use the optimal choice b_c^* for each Z_c. What is then the best choice for c^* in the sense of minimal variance of the resulting estimator (or greatest improvement of that compared to standard Monte Carlo, respectively)? Determine the value c^* numerically.

(c) Conduct some Monte Carlo simulations with the control variate Z_{c^*} and the optimal $b_{c^*}^*$ and compare the resulting error $|\widetilde{\mu}_N - \pi|$ for estimating π to the error $|\widehat{\mu}_N - \pi|$ of a standard Monte Carlo simulation with the same number of samples.

(d) Repreat the comparison for a non-optimal c and a non-optimal b. Try to find very bad choices of b and c.

(e) Now, illustrate the distribution of the errors $|\widetilde{\mu}_N - \pi|$ and $|\widehat{\mu}_N - \pi|$ for a moderate N, e.g., $N = 1000$ and for the same choices of c und b as in the previous two tasks.