Department of Applied Physics, Chalmers, Göteborg
Göran Wahnström, 2020-11-19

E3 Monte Carlo integration

To get familiar with the Monte Carlo method you are here asked to first consider a one-dimensional integral. You will implement importance sampling using the Transformation method. You will then implement the Metropolis algorithm and use that to solve a three dimensional integral. You will also consider the evaluation of the statistical inefficiency s, which can be used to determine proper error bars.

Task

1. Consider the one-dimensional integral

$$I = \int_0^1 f(x)\,\mathrm{d}x = \int_0^1 x\,(1-x)\,\mathrm{d}x$$

Generate N points x_i uniformly on the interval [0,1] and evaluate the integral with the Monte Carlo method. Perform the calculations for $N = 10^1$, 10^2, 10^3, and 10^4, and estimate the error. Compare your results with the exact value. (0.5 p)

2. The efficiency of the method can be improved using importance sampling. Use a sine function

$$p(x) \propto \sin(\pi x)$$

as weight function. Generate N points x_i according to the weight function using the Transformation method. Convience yourself that the method is working correctly by plotting a histogram of the generated points. Perform the calculations for $N = 10^1$, 10^2, 10^3, and 10^4 and estimate the error. Compare the results with the previous method. Is the present method more efficient? (0.5 p)

3. Consider now the three-dimensional integral

$$I = \pi^{-3/2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x^2+x^2y^2+x^2y^2z^2) \exp\left[-(x^2+y^2+z^2)\right] dxdydz$$

Use

$$w(xyz) = \pi^{-3/2} \exp\left[-(x^2+y^2+z^2)\right]$$

as weightfunction, where

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(xyz)dxdydz = 1$$

and the Metropolis algorithm to generate points according to the above weightfunction. A good choice for the trial changes $\tau_{m\to n}$ are

$$x_n = x_m + \Delta(r - 0.5)$$

and similarly for y and z, and where r is a uniform random number on the interval [0,1] and Δ is a parameter. A "rule of thumb" is to choose Δ such that about half or slightly less of the trials becomes accepted. Compute the integral. You do not need to give error bars. (1 p)

4. In large scale Monte Carlo simulations it is important to determine the statistical inefficiency s to be able to properly estimate the errors. On the course homepage you find the text file `MC.txt` containing $N = 10^6$ values f_i from a Monte Carlo simulation. Use this data set and determine s using both the method based on evaluation of the correlation function as well as on block averaging.

   Determine the auto-correlation function

$$\Phi_k = \frac{\langle f_{i+k} f_i \rangle - \langle f_i \rangle^2}{\langle {f_i}^2 \rangle - \langle f_i \rangle^2}, \quad \Phi_{-k} = \Phi_k$$

   and estimate s from

$$\Phi_{k=s} = \exp(-2) \approx 0.135$$

   This is based on the assumption that the correlation function Φ_k decays exponentially.

   Consider also block-averaging. In that method new blocked variables

$$F_j = \frac{1}{B} \sum_{i=1}^{B} f_{i+(j-1)B}$$

   have to be determined for various block sizes. If the block size becomes sufficiently large ($B > s$) the blocked variables F_j become statistically independent and

$$s = \frac{B \ \mathrm{Var}[\, F \,]}{\mathrm{Var}[\, f \,]}$$

   Do you get consistent results for s using the two different methods? Notice, if the mean value of f is large compared with the typical deviations from the mean value the above formula for the correlation can introduce numerical inaccuracies. It is then better that first to determine the mean value and then calculate the correlation for the deviation from the mean value. (2 p)