swissQuant
Mastering Complexity

Mathematical Challenge October 2019

Variance Reduction Techniques in Monte Carlo Methods

References

[1] Paul Glasserman. *Monte Carlo methods in financial engineering*. Springer, New York, 2004.

[2] Paul Glasserman, Philip Heidelberger, and Perwez Shahabuddin. *Importance sampling and stratification for value-at-risk*. IBM Thomas J. Watson Research Division, 1999.

Description

Motivation

Monte-Carlo (MC) methods are ubiquitous in Financial Mathematics (see [1] for a reference), especially for problems that involve a high number of underlyings/factors (e.g. pricing of structured products on multiple underlyings, LIBOR Market Model applications, etc.). The latter comes from the fact that MC methods show a convergence of order $\sqrt{(1/N)}$, where N denotes the number of simulations, independently of the underlyings dimension. Depending on the problem complexity, the computational costs of evaluating samples can be significantly high and variance reduction techniques are often used to speed-up the convergence rate, therefore reducing the number of simulations needed to achieve a certain accuracy.

Technical Details

In a nutshell, MC methods allow to evaluate numerically the expectation of a generic random variable $Y = \phi(X)$ by generating samples from the known distribution of X as:

$$\mathbb{E}[Y] \approx \frac{\sum_{i=1}^{N} \phi(X_i)}{N} = \hat{\mu}_Y \tag{1}$$

where X_i denotes a sample of X and N is the total number of samples. If $\phi(X)$ is square-integrable with variance σ^2 then, applying the Central Limit Theorem,

$$\hat{\mu}_Y - \mathbb{E}[Y] \sim \mathcal{N}\left(0, \frac{\sigma^2}{N}\right). \tag{2}$$

The last equation shows that the convergence rate of the standard deviation of the error decreases with the $\sqrt{N}$. To further improve the convergence, the following variance reduction techniques are often used:

- **Antithetic Variables:** this method reduces the variance by introducing negative dependence between pairs of samples. It is based on the observation that if U is uniformly distributed on [0, 1], then $\bar{U} = 1 - U$ is too (the latter extends to a generic random variable X applying the inverse transform of the distribution function, i.e. $X = F^{-1}(U)$). Hence for each realization U_i, the symmetric one $\bar{U}_i$ is also considered. The variables U_i and $\bar{U}_i$ form an antithetic pair in the sense that a large value of one is accompanied by a small value of the other. Intuitively this may result, depending on $\phi(\cdot)$, in a reduced variance. The antithetic variable MC estimate reads:

$$\hat{\mu}_Y^{AV} = \frac{\sum_{i=1}^{N} Y_i + Y_i}{2N} \tag{3}$$

- **Control Variables:** this method exploits information about errors in estimates of known quantities to reduce the error in an estimate of an unknown quantity. Suppose that from the i-th sample we are able to compute, jointly with Y_i, a realization Z_i of Z and that $\mathbb{E}[Z]$ is known analytically. Then for a fixed b we can compute the control variable MC estimate as:

$$\hat{\mu}_Y^{CV} = \frac{\sum_{i=1}^{N} Y_i - b(Z_i - \mathbb{E}[Z])}{N} \tag{4}$$

  The derivation of b^* , i.e. the optimal b, is left as an exercise to the reader (see Q2), but in practice the latter quantity is in most of the cases unknown and it has to be estimated upfront with a smaller MC estimate.

- **Importance Sampling:** this latter method tries to reduce the variance by changing the probability measure from which the paths are generated. Changing measure is a standard tool in Financial Mathematics. In Importance Sampling, the measure is changed to give more weight to "important" outcomes thereby increasing sampling efficiency. To make this idea concrete, rewrite the target expectation as

$$\mathbb{E}[\phi(X)] = \int \phi(x) f(x)\, \mathrm{d}x = \int \phi(x) \frac{f(x)}{g(x)} g(x)\, \mathrm{d}x = \mathbb{E}^g\left[\phi(X)\frac{f(X)}{g(X)}\right] \tag{5}$$

  where $f(\cdot)$ and $g(\cdot)$ are the density in the original and sampling probability measures respectively. The Importance Sampling MC estimator therefore reads:

$$\hat{\mu}_Y^{IS} = \frac{1}{N}\sum_{i=1}^{N} Y_i \frac{f(X_i)}{g(X_i)} \tag{6}$$

  where Y_i are now sampled accordingly to $g(\cdot)$. It can be shown that, in order to minimize the variance, $g(\cdot)$ should be selected as close as possible to $\phi(\cdot)f(\cdot)$: intuitively the optimal sampling probability density takes into account both the "true" probability and the outcome of the function ϕ (for instance considering a vanilla option, all samples which terminate Out-of-The-Money are “useless” in the sense that they do not contribute to $\mathbb{E}[Y]$).

Questions

- **Q1:** *Referring to the Antithetic Variables technique and assuming that the computational effort required to generate a pair* $(Y_i, \bar{Y}_i) = (\phi(X_i), \phi(\bar{X}_i))$ *is approximately twice the effort to generate* Y_i*, prove that:*

$$\mathrm{Var}(\hat{\mu}_Y^{AV}) < \mathrm{Var}\left(\frac{1}{2N}\sum_{i=1}^{2N} Y_i\right) \quad \text{iff} \quad \mathrm{Cov}(Y_i, \bar{Y}_i) < 0 \tag{7}$$

Which condition on the payoff function $\phi(\cdot)$ *guarantees the above inequality? Mention one derivative product/structure that would not benefit from Antithetic Sampling.*

- **Q2:** *Referring to the Control Variables technique prove that:*

$$b^* = \frac{\mathrm{Cov}(Y, Z)}{\mathrm{Var}(Z)} \tag{8}$$

and

$$\frac{\mathrm{Var}(\hat{\mu}_Y^{CV})}{\mathrm{Var}(\hat{\mu}_Y)} = 1 - \rho_{Y,Z}^2 \tag{9}$$

Considering a call option and using its underlying as control variable, how would you expect the above ratio to behave as a function of the moneyness? Think of cases where a good control candidate is available for unknown pricing functions.

- **Q3:** *Regarding the Importance Sampling technique*:
  - *given the Black-Scholes model, how can you change the sampling algorithm to price a deep Out-of-The-Money more efficiently?*
  - *consider a portfolio which contains derivative products. The latter cannot be priced with analytical formula, but Deltas and Gammas are available. We are interested in computing the Value-at-Risk of the portfolio. Relying on a nested-MC would be too computationally expensive. How can you speed-up the estimation of the Value-at-Risk of the portfolio using Importance Sampling? (see [2])*