Stochastic Simulation

Autumn Semester 2024

Prof. Fabio Nobile *Assistant: Matteo Raviola*

Project - 7

Submission deadline: 16 January 2025

Multi-level Monte Carlo methods for option pricing

1 Introduction and Background

Consider the goal of computing $\mu = \mathbb{E}[Y]$ where Y is the output of some stochastic model. In many applications, the stochastic model involves some differential operator and, as such, it can not be simulated exactly and a "discretization step" is necessary, characterized by a discretization parameter h, as, for instance, the characteristic mesh size in a finite difference approximation. It follows that we can only simulate an approximate output quantity Y_h. A Monte Carlo estimator for μ will look like

$$\hat{\mu}_h^{\mathrm{MC}} = \frac{1}{N}\sum_{i=1}^{N} Y_h^{(i)}, \quad \text{with } Y_h^{(i)} \overset{iid}{\sim} Y_h, \tag{1}$$

and will be a biased estimator, since, due to the discretization step, $\mathbb{E}[\hat{\mu}_h^{MC}] = \mathbb{E}[Y_h] \neq \mathbb{E}[Y]$.

As an example of the above framework, we consider in this project the problem of computing the expectation of a quantity $Y = f(S_T)$, which involves the final time value of the solution of a stochastic differential equation (SDE)

$$dS(t) = a(S(t),t)dt + b(S(t),t)dW_t, \quad t \in (0,T], \quad S(0) = S_0, \tag{2}$$

where W_t is a standard Brownian motion. The exact solution of equation (2) is in general not known, except for special forms of a and b, but it can be approximated using, for instance, the Euler-Maruyama scheme. Let $t_m = mh, \quad m = 1,\dots,M = T/h$, and S^m be the approximation of $S(t_m)$ given by

$$S^{m+1} = a(S^m,t_m)h + b(S^m,t_m)\Delta W_m, \quad \text{with } \Delta W_m \overset{iid}{\sim} N(0,h), \quad m = 0,\dots,M-1$$

and $S^0 = S_0$. Under certain regularity conditions on a, b, and f, it can be shown that

$$|\mathbb{E}[Y_h] - \mathbb{E}[Y]| = \mathcal{O}(h),$$

so that the mean squared error (MSE) of the Monte Carlo estimator satisfies

$$\mathrm{MSE}(\hat{\mu}_h^{\mathrm{MC}}) := \mathbb{E}\left[(\hat{\mu}_h^{MC} - \mu)^2\right] = \frac{\mathrm{Var}[Y_h]}{N} + \mathcal{O}(h^2).$$

The convergence rate $\alpha = 1$ of the bias term $|\mathbb{E}[Y_h] - \mathbb{E}[Y]| = \mathcal{O}(h^\alpha)$ is refereed to as the *weak converge rate* (convergence on the expectation), as opposed to the *strong convergence rate*, which is the convergence rate of the error $\mathbb{E}[|Y_h - Y|^p]^{1/p} = \mathcal{O}(h^\eta)$, for some $p \geq 1$, usually with $\eta = 1/2$. If we assume that the cost per simulation is proportional to the number of steps, hence $\mathcal{O}(h)$ and $\text{Var}[Y_h] = \mathcal{O}(1)$, then, the choice $N_\epsilon = O(\epsilon^{-2})$ and $h_\epsilon = O(\epsilon)$ guarantees a MSE of order ϵ^2 and the corresponding total cost of computing $\hat{\mu}_h^{MC}$ is $C_\epsilon = O(Nh^{-1}) = O(\epsilon^{-3})$. The aim of Multi-level Monte Carlo is to improve this order. A brief description is given in the following sections, however, you are encouraged to read the first chapter of [1] for more details.

1.1 Multi-level Monte Carlo: a two level approach

MLMC is, essentially, a control variate technique for which we use the quantity of interest obtained with a coarser (i.e, less accurate) discretization as a control variable. We begin by discussing the two-level MLMC. Let $h_1 < h_0$ correspond to two discretization parameters, such that $Y_{h_1} := Y_1$ is a more accurate approximation of Y than $Y_{h_0} := Y_0$, yet being also more expensive to simulate. Notice that we can write

$$\mathbb{E}[Y_1] = E[Y_0] + \mathbb{E}[Y_1 - Y_0],$$

and as such, we can use the following 2-level estimator,

$$\hat{\mu}^{\text{2-level}} = N_0^{-1} \sum_{n=1}^{N_0} Y_0^{(n,0)} + N_1^{-1} \sum_{n=1}^{N_1} (Y_1^{(n,1)} - Y_0^{(n,1)}), \tag{3}$$

where $Y^{(n,0)}$ and $Y_1^{(n,1)} - Y_0^{(n,1)}$ are simulated independently, whereas $Y_1^{(n,1)}$, and $Y_0^{(n,1)}$ are simulated using the same underlying noise. If we let C_0 denote the cost of simulating Y_0 , C_1 the cost of simulating $Y_1 - Y_0$, and introduce the notation $V_0 = \text{Var}[Y_0]$ and $V_1 = \text{Var}[Y_1 - Y_0]$, then an optimal allocation of N_0, N_1 (when treating N_0, N_1 as real numbers) that minimizes the variance of the estimator $\hat{\mu}^{\text{2-level}}$ at a fixed cost satisfies:

$$\frac{N_1}{N_0} = \frac{\sqrt{V_1/C_1}}{\sqrt{V_0/C_0}}. \tag{4}$$

1.2 Multi-level Monte Carlo: multiple levels

The idea of the previous subsection can be easily generalized to a higher number of levels. Consider a sequence of levels $\ell = 0, 1, \ldots, L$, such that $Y_0, Y_1, \ldots, Y_L$ approximate Y with increasing accuracy and cost. As before, we can write $\mathbb{E}[Y_L]$ as

$$\mathbb{E}[Y_L] = \mathbb{E}[Y_0] + \sum_{\ell=1}^{L} \mathbb{E}[Y_\ell - Y_{\ell-1}] = \sum_{\ell=0}^{L} \mathbb{E}[Y_\ell - Y_{\ell-1}],$$

where in the last equality, we have used the convention that $Y_{-1} = 0$, and define the multi-level estimator

$$\hat{\mu}_L^{\text{MLMC}} \approx \sum_{\ell=0}^{L} \left[N_\ell^{-1} \sum_{n=1}^{N_\ell} (Y_\ell^{(n,\ell)} - Y_{\ell-1}^{(n,\ell)}) \right]$$

where, again, $Y_\ell^{(n,\ell)}$ and $Y_{\ell-1}^{(n,\ell)}$ are simulated using the same underlying noise, whereas $Y_\ell^{(n,\ell)} - Y_{\ell-1}^{(n,\ell)}$ and $Y_k^{(m,k)} - Y_{k-1}^{(m,k)}$ are generated independently for $k \neq \ell$ or $m \neq n$. Proceeding as before, we denote by C_ℓ the cost of evaluating $Y_\ell - Y_{\ell-1}$ and by V_ℓ its variance.

It can be shown (see [1]) that a quasi-optimal allocation of $N_0, N_1, \dots N_L$ that minimizes the total cost of the estimator $\hat{\mu}_L^{\mathrm{MLMC}}$ for a total variance $V_{\mathrm{MLMC}} \leq \epsilon^2$, is:

$$N_\ell = \left\lceil \epsilon^{-2} \left(\sum_{\ell=0}^{L} \sqrt{V_\ell C_\ell} \right) \sqrt{\frac{V_\ell}{C_\ell}} \right\rceil, \tag{5}$$

where $\lceil \cdot \rceil$ denotes the ceiling function (approximation to the nearest larger natural number). Notice that, in practice, $V_\ell \leq V_{\ell-1}$, as the discretization gets more accurate as ℓ increases, and $C_\ell \geq C_{\ell-1}$. Thus we have that $N_\ell < N_{\ell-1}$. The idea of the method is then to use large sample sizes at the lower accuracy-cost discretizations, and correct with smaller and smaller sample sizes as we move up on the levels ℓ. The following theorem, whose proof can be found in [1], gives a bound on the total cost of the MLMC estimator to achieve MSE$\leq \epsilon^2$. In particular, the theorem shows that, asymptotically, as $\epsilon \to 0$, the cost of the (optimally tuned) MLMC estimator is always smaller than the cost of the single-level Monte Carlo estimator (1), provided some assumptions on the decay of the weak and strong errors and increase of the cost with respect to ℓ are satisfied:

Theorem 1: *Let Y denote a random variable and let Y_ℓ denote the approximation of Y at level ℓ. Let $\hat{\mu}_\ell$ be a Monte Carlo estimator of $\mathbb{E}[Y_\ell - Y_{\ell-1}]$ based on N_ℓ independent replicas of $Y_\ell - Y_{\ell-1}$, each with cost C_ℓ and variance V_ℓ. If there exist positive constants $\alpha, \beta, \gamma, c_1, c_2, c_3$ such that $\alpha \geq \frac{1}{2}\min(\beta, \gamma)$, and*

i. $|\mathbb{E}[Y_\ell - Y]| \leq c_1 2^{-\alpha\ell}$

ii. $\mathbb{E}[\hat{\mu}_\ell] = \mathbb{E}[Y_\ell - Y_{\ell-1}], \quad Y_{-1} = 0,$

iii. $V_\ell \leq c_2 2^{-\beta\ell}$,

iv. $C_\ell \leq c_3 2^{\gamma\ell}$,

then there exists a positive constant c_4 such that for any $\epsilon < e^{-1}$ there are values L and N_ℓ for which the multi-level estimator

$$\hat{\mu}_L^{MLMC} = \sum_{\ell=0}^{L} \hat{\mu}_\ell$$

has a mean squared error smaller than ϵ^2, with computational cost C_{MLMC} given by

$$\mathbb{E}[C_{MLMC}] \leq \begin{cases} c_4 \epsilon^{-2}, & \text{if } \beta > \gamma, \\ c_4 \epsilon^{-2} |\log(\epsilon)|^2, & \text{if } \beta = \gamma, \\ c_4 \epsilon^{-2-(\gamma-\beta)/\alpha}, & \text{if } \beta < \gamma. \end{cases}$$

2 Goals of the project

We would like to apply the MLMC algorithm for option pricing. We consider the underlying SDE to follow a simple geometric Brownian motion described by

$$dS(t) = rS(t)dt + \sigma S(t)dW_t, \quad 0 < t < 1, \tag{6}$$

with $S(0) = 1, r = 0.05$ and $\sigma = 0.2$ and W_t a standard Brownian motion. The idea then is to investigate the expected payoff $\mu = \mathbb{E}[Y]$ for different types of options, namely, we will investigate:

1) Asian option: $Y^{(1)} = \exp(-r)\max(0, \bar{S} - K)$, with

$$\bar{S} = \int_0^1 S(t)dt.$$

2) Barrier call option: $Y^{(2)} = \exp(-r)\max(0, S(1) - K)\mathbb{1}_{\{\max_{t\in[0,1]} S(t)<S_{max}\}}$.

for $K = 1$. Please address the following points:

(a) Give a sketch of the proof of the optimality of formulas (4) and (5) when N_ℓ are treated as real numbers.

(b) Implement a standard Monte Carlo estimator for payoff 1. To do so, approximate $\bar{S}$ as $\bar{S} \approx h\sum_{m=1}^{M} \frac{S^m+S^{m-1}}{2}$ where $S^m = S(mh)$ (or an approximation of it) and $M = 1/h$ is the number of time steps.

Investigate how the bias and variance of your estimator scale with respect to h and N. You can estimate the bias by comparing $\mathbb{E}[Y_h^{(1)}]$ with $\mathbb{E}[Y_{2h}^{(1)}]$. Use the same sample to estimate both expectations. Propose an adaptive algorithm to estimate $\mathbb{E}[Y^{(1)}]$ satisfying $MSE(Y^{(1)}) \leq \epsilon^2$.

(c) Consider now a 2-level MLMC, and again payoff 1. Estimate N_1/N_0 based on a pilot run using two grids $h_1 = 0.1$ and $h_0 = 0.2$, considering that $C_1 = 2C_0$. Implement the 2-level MLMC and quantify the variance reduction obtained by $\hat{\mu}^{\text{2-level}}$ with respect to the crude Monte Carlo estimator of point (b) with $h = h_1$ that has comparable cost. What can you say about the bias of $\hat{\mu}^{\text{2-level}}$ in comparison to that of the crude Monte Carlo estimator of point (b)?

(d) We now move to a multi-level approach. Consider again payoff 1 and the hierarchy of meshes h_ℓ, $\ell = 0, 1, \ldots$, with $h_\ell = 0.2 \times 2^{-\ell}$, so that the unit costs are $C_\ell = C_0 2^\ell$. Moreover, assume that $V_\ell =: \mathbb{V}ar[Y_l^{(1)} - Y_{l-1}^{(1)}] \tilde{=} V_0 2^{-\beta\ell}$, $\beta = 1$, $\ell = 1, 2, \ldots, L$, and $E_\ell := |\mathbb{E}[Y_\ell^{(1)} - Y_{\ell-1}^{(1)}] \tilde{=} E_0 2^{-\alpha\ell}$, $\alpha = 1$. Estimate $\tilde{V}_0, \tilde{E}_0$ from a pilot run (i.e, from a relatively small sample size at each level for the first few levels). Then, implement the multi-level Monte Carlo estimator using (5) and choosing appropriately L and $\{N_l\}_{l=0}^{L}$ based on your estimated weak errors E_ℓ and variances V_ℓ to achieve a MSE less than $2\epsilon^2$. Compare the computational work required for the MLMC estimator to that of a standard Monte Carlo estimator that achieves the same MSE for different tolerance values ϵ. Comment on the obtained results.

(e) Consider now the payoff 2 with $S_{max} = 1.5$ and its approximation

$$Y_h^{(2)} = e^{-r} \max\left\{0, S^M - K\right\} \mathbb{1}_{\{\max_{m=0,\dots,M} S^m < S_{max}\}}, \quad M = 1/h, \tag{7}$$

where $\{S^m\}_{m=0}^{M}$ are the values of the process at $t_m = mh$. Repeat the previous point, estimating α and β numerically using the first few levels. Report α and β and comment your results.

(f) Consider now a higher strike price $K = 2$ and $S_{max} = 2.5$. For each of the two payoffs, compute a (crude) Monte Carlo estimator of $\mathbb{E}[Y]$. Propose and implement a Variance Reduction Technique (VRT) for such an estimator and report your results. Can this VRT be used in the context of MLMC as well?

References

[1] Giles, M. B. (2015). Multilevel Monte Carlo Methods. *Acta Numerica, 24,* 259-328.