Stochastic Simulations

Autumn Semester 2021

Prof. Fabio Nobile *Assistant: Sundar Ganesh*

Project 6

submission deadline: 19 January 2022

Stochastic approximation in mathematical finance

1 Introduction and background

The goal of this project is to investigate the performance of stochastic approximation methods to compute the root or the minimum of an objective function representing some expected loss. The problem can be formalized as follows: let $X \in \mathbb{R}^d$ be a random variable and $f_\theta : \mathbb{R}^d \mapsto \mathbb{R}$ a family of functions parameterized by a parameter $\theta \in \mathbb{R}^p$. Then, we aim to

$$\text{find} \quad \theta^* : \quad \mathbb{E}[f_{\theta^*}(X)] = 0, \tag{1}$$

or

$$\text{find} \quad \theta^* = \underset{\theta}{\operatorname{argmin}}\, \mathbb{E}[f_\theta(X)] \implies \nabla_\theta \mathbb{E}[f_{\theta*}(X)] = 0. \tag{2}$$

In both cases, the exact evaluation of either $\mathbb{E}[f_\theta(X)]$ or $\nabla_\theta \mathbb{E}[f_\theta(X)]$ can be approximated by a Monte Carlo estimate. As an application, this project focuses on estimating the implied volatility in option pricing problems.

1.1 Robbins Monro algorithm

Let us denote $J(\theta) = \mathbb{E}[f_\theta(X)]$ or $J(\theta) = \nabla_\theta \mathbb{E}[f_\theta(X)]$, depending on the problem to be solved. The Robbins-Monro (RM) algorithm [2] is an iterative scheme of the form

$$\theta_{n+1} = \theta_n - \alpha_n \widehat{J}(\theta_n) \tag{3}$$

where $\widehat{J}(\theta)$ is an *unbiased* estimator of $J(\theta)$. For instance, in the case $J(\theta) = \mathbb{E}[f_\theta(X)]$, one could take a simple Monte Carlo estimator

$$\widehat{J}(\theta) = \frac{1}{N} \sum_{i=1}^{N} f_\theta(X^{(i)}) \tag{4}$$

with $X^{(i)} \overset{\text{iid}}{\sim} X$. In (3), α_n is the so-called *learning rate* of the RM algorithm. To guarantee convergence of the RM algorithm, α_n must satisfy:

$$\sum_n \alpha_n = +\infty, \quad \sum_n \alpha_n^2 < +\infty. \tag{5}$$

For more details on the properties of the RM algorithm and the proof of convergence, you are welcome to review Chapter 8 from [1]. For our purposes, we will choose the learning rate to have the form $\alpha_n = \frac{\alpha_0}{n^\rho}$, where $\alpha_0 \in \mathbb{R}_+$ and $\rho \in (1/2, 1]$ will be application-specific.

1.2 Extracting implicit volatility in option pricing

Consider the problem of pricing an option with maturity $T > 0$ based on the stock price $\{S_t,\ t \in [0,T]\}$, which is given as the solution to the stochastic differential equation

$$dS_t = rS_t dt + \sigma S_t dW_t, \tag{6}$$

Here S_0 represents the initial value of the underlying asset, r is the interest rate, σ is the volatility and W_t denotes a standard one-dimensional Wiener process. One can show that a solution to such SDE is given by

$$S_t = S_0 \exp\left((r - \sigma^2/2)t + \sigma W_t\right), \tag{7}$$

from which it follows that S_t has a log-normal distribution for any $t > 0$.

We consider in this project a possibly path-dependent option whose (discounted) payoff is a function of $S = (S_{\frac{T}{m}}, S_{\frac{2T}{m}}, \dots, S_T)$. Let $f(S) = f(S_{\frac{T}{m}}, S_{\frac{2T}{m}}, \dots, S_T)$ be such function and denote by $p(S)$ the joint probability distribution of the random vector S. Then, the option price is given by

$$I = \mathbb{E}[f(S)] = \int_{\mathbb{R}^m} f(S)\, p(S) dS. \tag{8}$$

Notice that the expected value in (8) depends on the parameters (σ, r, S_0, T). We are, in particular, interested in the dependence $I = I(\sigma)$ on the volatility and the possibility of estimating the *implied volatility* σ^* from available market price value I_{market} of the option, i.e., to determine σ^* s.t. $I_{\text{market}} - I(\sigma^*) = I_{\text{market}} - \mathbb{E}[f(S;\sigma^*)] = 0$. Notice that this problem can be set in the form (1):

$$\text{find}\ \ \sigma^*: \quad \mathbb{E}[\tilde{f}_{\sigma^*}(S)] = 0, \quad \tilde{f}_\sigma(S) = f(S;\sigma) - I_{\text{market}} \tag{9}$$

2 Goals of the project

2.1 Warm-up: European put option

As a warm-up, we begin our numerical experiments by considering a European put option for which

$$f(S) = f(S_T) = e^{-rT}\left(K - S_T\right)_+, \tag{10}$$

where $(x)_+ = \max\{x, 0\}$, with given interest rate $r = 5\%$, maturity time $T = 0.2$, initial asset price $S_0 = 100$ and strike price $K = 120$. On the other hand, the volatility σ is unknown and should be estimated from mark-to-market prices of the option. For this problem, a closed form for $I(\sigma)$ is available (Black-Scholes formula):

$$I(\sigma) = e^{-rT} K \Phi[w^*(\sigma)] - S_0 \Phi[w^*(\sigma) - \sigma\sqrt{T}], \tag{11}$$

where $\Phi[\cdot]$ is the cummulative distribution function of a standard Gaussian r.v. and

$$w^*(\sigma) = \frac{\log[K/S_0] - (r - \sigma^2/2)T}{\sigma\sqrt{T}}. \tag{12}$$

The function $\sigma \to I(\sigma)$ is increasing and satisfies $I(\sigma) \to \left(e^{-rT}K - S_0\right)_+$ as $\sigma \to 0$ and $I(\sigma) \to e^{-rT}K$ as $\sigma \to +\infty$. Hence, for any $I_{\text{market}} \in [(e^{-rT}K - S_0)_+, e^{-rT}K]$, the equation $I_{\text{market}} - I(\sigma) = 0$ has a unique solution that can be computed easily with some root-finding algorithm.

1. Let $I_{\text{market}} = 22 \in \left[\left(e^{-rT}K - S_0\right)_+, e^{-rT}K\right]$ be the consistent mark-to-market price for the put option with maturity T and strike price K. Use a root-finding algorithm to compute the "true" implied volatility σ^* such that:

$$I(\sigma^*) = I_{\text{market}} \iff \mathbb{E}[f(S_T;\sigma^*) - I_{\text{market}}] = 0, \tag{13}$$

2. Now use the RM algorithm to find the implied volatility $\hat{\sigma}$ using different sample sizes $N \in \{1, 10, 100, ...\}$ in Eq. (4). Try different values of $\rho = 0.8, 1$ (in the learning rate $\alpha_n = \alpha_0/n^\rho$). Estimate the mean squared error (MSE) $E_n = \mathbb{E}[(\hat{\sigma}_n - \sigma^*)^2]$ by repeating the simulation multiple times. Plot the estimated MSE E_n as a function of n (To be able to compare the cases with $N = 1, 10, 100...$, it is actually better to plot E_n versus the total number of random values generate upto iteration n, i.e., $N \times n$). What convergence rate do you observe? How does the MSE depend on N?

   Hint: We recommend taking $\alpha_0 = \frac{2}{K+S_0}$ in the learning rate.

2.2 Asian put option

Consider now the *path dependent* discounted payoff

$$f(S) = e^{-rT}\left(K - \bar{S}_T\right)_+ \tag{14}$$

where $\bar{S}_T = \frac{1}{m}\sum_{i=1}^m S_{iT/m}$ is the *discrete monitoring average* in the time interval $[0, T]$. The parameters (r, T, S_0, K) are as in Section 2.1 and $m = 50$. Then $\sigma \mapsto I(\sigma)$ is also increasing with respect to σ.

4. Show that $I(\sigma) \to e^{-rT}\left(K - \frac{S_0}{m}\sum_{i=1}^m e^{riT/m}\right)_+$ when $\sigma \to 0$ and $I(\sigma) \to e^{-rT}K$ when $\sigma \to \infty$.
   Hint: Use the Markov property to show that the joint density of $(S_{T/m}, S_{2T/m}, \ldots, S_T)$ is

$$p(x_1, \ldots, x_m) = \prod_{i=1}^m \frac{1}{x_i\sigma\sqrt{t_i - t_{i-1}}}\Phi\left(\frac{\log[x_i/x_{i-1}] - (r - \sigma^2/2)(t_i - t_{i-1})}{\sigma\sqrt{t_i - t_{i-1}}}\right), \tag{15}$$

   where $\Phi(\cdot)$ is the probability density function of a standard Gaussian r.v.

5. For $I_{\text{market}} = 22$ as in point 1, find again σ^* that satisfies (13) using the RM algorithm. Monitor the convergence of the algorithm and propose a suitable stopping criterion for the iterations.

6. Consider now the case $K = 150$. A crude Monte Carlo estimator for $I(\sigma)$ is rather inefficient and could be improved by some variance reduction technique.

(a) First, consider and implement an importance sampling (IS) strategy to approximate $I(\sigma)$ for a given σ, where the importance distribution is constructed by modifying the interest rate r to $\tilde{r}$ in the dynamics (6). Justify your choice of $\tilde{r}$ and quantify the variance reduction achieved for different values of σ.

(b) Modify the RM algorithm of point 5 to include importance sampling. Compare its performance with that of point 5.

7. Consider now a more sophisticated IS strategy. In a screening phase, we try to find the optimal IS drift $\tilde{r}(\sigma)$ as a function of σ. More precisely, let $p(S; r, \sigma)$ be the joint PDF of $S = (S_{\frac{T}{m}}, S_{\frac{2T}{m}}, \dots, S_T)$, when (r, σ) are the drift and volatility in (6). For a given value of σ, one could estimate the optimal importance sampling parameter $\tilde{r} = \tilde{r}(\sigma)$ by performing a pilot run $S^{(i)} \overset{\text{iid}}{\sim} p(\cdot; r, \sigma)$, $i = 1, 2, \dots, \bar{N}$, and minimizing the variance of the importance sampling estimator

$$\begin{aligned}
\tilde{r}(\sigma) &= \arg\min_{\eta} \mathbb{V}_{p(\cdot;\eta,\sigma)} \left[\frac{p(S; r, \sigma)}{p(S; \eta, \sigma)} f(S; r, \sigma) \right] & (16) \\
&= \arg\min_{\eta} \mathbb{E}_{p(\cdot;r,\sigma)} \left[f^2(S; r, \sigma) \frac{p(S; r, \sigma)}{p(S; \eta, \sigma)} \right] & (17) \\
&\approx \arg\min_{\eta} \frac{1}{\bar{N}} \sum_{i=1}^{\bar{N}} \frac{f^2(S^{(i)}; r, \sigma) p(S^{(i)}; r, \sigma)}{p(S^{(i)}; \eta; \sigma)}. & (18)
\end{aligned}$$

We can then estimate the optimal parameter $\tilde{r}$ for several values $\sigma \in [0, 5]$ and build a suitable approximate function $\tilde{r}(\sigma)$, e.g. by spline interpolation (It is wise in this context to always use the same generated paths $S^{(i)}$ when optimizing $\tilde{\sigma}(\sigma)$ for different values of σ). With such function $\tilde{r}(\sigma)$ we can now run the RM algorithm where at each iteration k we use an IS with modified drift $\tilde{r}(\sigma)$. Implement such a strategy and compare its performance with the one in point 4, which uses the same modified drift $\tilde{r}$ for all iterations of RM.

References

[1] S. Asmussen and P.W. Glynn, *Stochastic simulation: algorithms and analysis (vol. 57)*, Springer Science & Business Media, 2007.

[2] H. Robbins and S. Monro, *A stochastic approximation method, in herbert robbins selected papers (pp. 102-109)*, Springer, New York, NY, 1985.