Bermudan Options Pricing with Stochastic Grid Bundling Method

awalee

Summary

Introduction

Pricing American-style derivatives, i.e., contracts in which the holder can choose the time of exercise during the life of the derivative is still a challenge for banks and the complexity and cost of implementing the necessary infrastructure remains a tough task for most financial institutions. In order to comply with regulatory requirements, institutions need to compute supplementary metrics for risk management. These metrics, and particularly the Credit valuation adjustment (CVA), have a high cost in terms of computational resources as they need the product to be valued in a time bucket form valuation date to its maturity.

This becomes more challenging when it comes to pricing American or Bermudan options because of the early exercise feature. This requires an effective and particularly fast method for the implementation process.

There are traditional valuation methods such as lattices and tree-based techniques, that are used mainly in the low-dimensional cases. But they are often impractical and inaccurate when the dimension increases.

Several simulation-based methods have been proposed to price options with early-exercise features, which combine Monte Carlo paths generation and dynamic programming techniques to determine optimal policies. We refer to [1] for more details regarding this topic.

In this paper, we investigate the Stochastic Grid Bundling Method (SGBM). This is a simulation-based dynamic programming method that enhances the classic least square Monte Carlo method (LSM) which has been developed by Longstaff and Schwartz, see [2] for details. First it generates paths of the underlying variable forward in time then uses machine learning algorithms to create clusters of the values at a given time so that the regression, used to compute the continuation value, performs well on each cluster. The optimal early-exercise policy is obtained by moving backward in time.

SGBM belongs to the class of regression-based methods. Another approach is based on approximating the transition probabilities using either bundling, partitioning or quantization of the state space.

Since, owing to the central limit theorem, the convergence rate of methods based on simulation of the underlying state vector converges in proportion to the square root of the number of paths generated and is independent of the dimension of the problem, this makes simulation-based methods attractive for valuing path-dependent and multi-asset derivatives. But they can be complicated when the option has American-style features. In such a situation, an optimal exercise policy has to be determined via a dynamic programming approach. The difficulty then arises in combining the forward evolution of simulation paths with the backward induction of dynamic programming.

We will first give a formal presentation of the method which is a mix of partitioning and local regression. Then we will detail some of the important and most suited clustering algorithms. The efficiency of the method will then be emphasized in the case of CIR model. Finally, we will discuss a different approach to price Bermudan options which is a quantization-based method.

1 Presentation of the Stochastic Grid Bundling Method (SGBM)

In this section, we will make a formal presentation of SGBM. The method is a hybrid of regression and bundling-based approaches, and uses regressed value functions, together with bundling of the state space to approximate continuation values at different time steps.

Let $(\Omega, F, \mathbb{P})$ be a probability space, Ω is the set of all possible realizations of the stochastic variable X_t, $t \in [0, T]$, F is the filtration generated by this variable and $\mathbb{P}$ is the risk-neutral probability measure.

A Bermuda option is a type of exotic options contract that can only be exercised on predetermined dates.

Let $U(X_t)$ be the payoff function of a Bermudan option on the underlying X_t.

Suppose that the maturity of the option is T and exercise times are $\{t_1 < \cdots < t_M = T\}$ and we want to value the option at time $t_0 = 0$.
The value of the option at time T is simply:

$$V_T = U(X_T)$$

For some m such that $0 \leq m \leq M-1$, the conditional continuation value $C(t_m, X_{t_m})$ is given by:

$$C(t_m, X_{t_m}) = D_{t_m}\mathbb{E}[V_{t_{m+1}}(X_{t_{m+1}})|X_{t_m}]$$

where D_{t_m} is the discount factor defined as

$$D_{t_m} = e^{-\int_{t_m}^{t_{m+1}} r_s ds}$$

r_s is the instantaneous risk-free rate and $\mathbb{E}$ is the expectation with respect to $\mathbb{P}$.

The Bermudan option value at time t_m is given by:

$$V_{t_m}(X_{t_m}) = max\left(U(X_{t_m}), C(t_m, X_{t_m})\right)$$

We are seeking $V_{t_0}(X_{t_0})$.
To this end, we need to compute the continuation value $C(.,.)$.

We first generate N risk neutral Monte Carlo paths and then we construct the bundles. These are sub-sets that divide the domain into disjoint sub-domains that have very similar realized values of the risk-neutral samples. The next section will detail some of the most popular bundling techniques.

Denote by X_{t_m} some of the N realizations of the underlying variable X_t at time t_m. Now suppose we have the option value $V_{t_{m+1}}(X_{t_{m+1}})$ at time t_{m+1}. For each bundle B_j we perform a local regression to find the coefficients $\beta_1, \ldots, \beta_b$ of the regression by minimizing the sum:

$$\sum_{X_{t_{m+1}} \in B_j} \left(V_{t_{m+1}}(X_{t_{m+1}}) - \sum_{k=0}^{b} \beta_k \times \phi_k(X_{t_{m+1}}) \right)^2$$

where

$$\phi_k(.), 0 \leq k \leq b$$

are some basis functions.

Then the continuation function over each bundle B_j is estimated by:

$$\hat{C}_j(t_m, X_{t_m}) = \sum_{k=0}^{b} \beta_k \times \psi_k(X_{t_m})$$

with

$$\psi_k(X_{t_m}) = D_{t_m}\mathbb{E}[\phi_k(X_{t_{m+1}})|X_{t_m}].$$

For the paths in the bundle B_j, the option value at time t_m is then

$$V_{t_m}(X_{t_m}) = max\left(U(X_{t_m}), \hat{C}_j(t_m, X_{t_m})\right)$$

where $U(.)$ is the payoff function.

We iterate the process for all the bundles at each exercise date t_m going backward in time from maturity t_M until t_1, the first exercise date.

Once this is done, the option value V_{t_0} is therefore the average of the discounted cash flows at t_1:

$$V_{t_0} = D_{t_1} mean(V_{t_1}(X_{t_1})).$$

If the basis functions are polynomial, that is to say we have:

$$\phi_k(X) = X^k, 0 \leq k \leq b$$

and the characteristic function of the underlying process X_t is known, analytic formulas can be obtained for:

$$\psi_k(.), 0 \leq k \leq b.$$

We first compute the associated characteristic function of the underlying variable which is:

$$\Phi_X(y) = \mathbb{E}[exp(iyX_t)].$$

Consequently, we have:

$$\mathbb{E}[X^k] = \frac{1}{i^k}\frac{d^k\Phi_X(y)}{dy^k}|y = 0.$$

Therefore ψ_k can be derived by:

$$\begin{aligned}\psi_k(X_{t_m}) &= D_{t_m}\mathbb{E}[\phi_k(X_{t_{m+1}})|X_{t_m}] \\ &= D_{t_m}\mathbb{E}[(X_{t_{m+1}})^k|X_{t_m}] \\ &= D_{t_m}\tfrac{1}{i^k}\tfrac{d^k\Phi_X(y)}{dy^k}|y = 0.\end{aligned}$$

We refer to the appendix for CIR model moments calculation.

Classic LSM can be viewed as SGBM but with only one bundle including all the paths. Compared to LSM, SGBM may improve the local approximation in regression because we use disjoint sub-domains, and we can reduce the number of basis functions which allows to reduce computation time.

LSM is efficient for the computation of option value at time zero only. It doesn't give an accurate value of the option at future times (It overprices it) which is not desirable for CVA calculation.

The following section aims at detailing some of the most important bundling techniques.

2 Bundling techniques

SGBM employs bundling to approximate the conditional distribution using simulation. The aim of bundling in SGBM is to cluster grid points based on proximity. There are many bundling techniques, we will focus on the most important ones.

All of them aims at bundling the stochastic grid points at time t_m into non-overlapping sets

$$B_{t_m}(1), \ldots, B_{t_m}(n)$$

where n is the number of bundles. We generally use clustering algorithms for that aim.

In following subsections, we will consider two of them: K-means and equal-number bundling partitioning.

K-means

Given a set of observations at time t_m $(X_{t_m}(1), \ldots, X_{t_m}(N))$ where each observation is a d-dimensional real vector that, in our case, represent the values of the underlying variable along the N generated paths using Monte Carlo, K-means clustering aims at partitioning the N observations into $K \leq N$ sets

$$B_{t_m} = (B_{t_m}(1), \ldots, B_{t_m}(K))$$

so as to minimize the within-cluster variance. Formally, the objective is to find B_{t_m} that minimizes:

$$\sum_{i=1}^{K} |B_{t_m}(i)| Var(B_{t_m}(i))$$

where $|B_{t_m}(i)|$ denotes the cardinality of the set $B_{t_m}(i)$ and $Var(.)$ its variance.

The algorithm works as follows:

-First, we determine the number of clusters we want to build and eventually the maximum number of iterations.

-We initialize K points. Some of the various ways to do so are:

• Random generation of the first k centroids.

• Farthest selection: The first center is selected as a random case from the dataset. The second center should be the farthest from the first center and so on. Each time the newly added center should be the farthest from the centroid of the set.

-We categorize each item to its closest center and we update the center's coordinates, which becomes the centroid of the items categorized in that cluster so far.

-We repeat the process for a given number of iterations and at the end, we have our clusters.

Equal-number bundling technique

The simpler way to create bundles is to choose the number of bundles J, then rank the values of the paths at a given time and put N/J values in each bundle starting from the smallest ones. These bundles divide the domain into disjoint sub-domains.

The aim of the next section is to give an insight about a different approach to price Bermudan options. It is based on transition probabilities approximation using quantization.

3 Pricing Bermudan options with Quantization

The quantization provides an optimal spatial discretization of a random variable by a discrete variable taking at most N values called quantizers. Quantization is a deterministic alternative to Monte Carlo (MC) simulation, it can be more accurate and faster than MC (at least for dimension ≤ 5) for some models and derivatives payoffs.

Let $X \in R^d$, be some random variable and $x = (x^1, \ldots, x^N) \in (R^d)^N$.
A Borel partition $((C_i(x))_{i=1,\ldots,N}$ of R^d is a Voronoi partition of N-quantizer x, if for every $i = 1, \ldots, N$:

$$C_i(x) \subset \left\{ \xi \in R^d, ||\xi - x^i||_2 \leq \min_{i \neq j} ||\xi - x^j||_2 \right\}$$

where $||.||_2$ is the Euclidean norm.

The N-Quantizer of X is:

$$\hat{X}^x = \sum_{i=1}^{N} x^i 1_{X \in C_i(x)}$$

where the quantizers minimize:

$$||X - \hat{X}^x||_2 = \left(\mathbb{E}[\min_{1 \leq i \leq N} ||X - x^i||_2^2] \right)^{\frac{1}{2}}$$

Given that $\hat{X}^x$ can be a risk factor or any underlying asset, the expected value (for a payoff f) can be approximated by:

$$\mathbb{E}[f(X)] \approx \mathbb{E}[f(\hat{X}^x)] = \sum_{i=1}^{N} f(x^i) \mathbb{P}(X \in C_i(x))$$

where $\mathbb{P}$ is the probability distribution function of the variable X and $\mathbb{E}$ is the expectation with respect to $\mathbb{P}$.

Now in order to price Bermudan options with quantization, we need to do the following: We must first set the number of quantizers in each exercise date. Suppose we have N_k quantizers for each date $t_k \in [0, T]$ with $1 \leq k \leq n$ and $t_n = T$ is the maturity of the Bermudan option. Let $\hat{X}_k = \left(x_1^k, ..., x_{N_k}^k\right)$ be the quantized variable of X_t at $t = t_k$.

We define by backward induction the function:

$$\begin{cases} \hat{v}_n(x_i^n) = U(x_i^n), 1 \leq i \leq N_n \\ \hat{v}_k(x_i^k) = max\left(U(x_i^k), \sum_{j=1}^{N_{k+1}} \pi_{ij}^{k+1} \hat{v}_{k+1}(x_j^{k+1})\right), 1 \leq i \leq N_k, \\ 1 \leq k \leq n-1 \end{cases}$$

The option value at maturity is then:

$$\hat{V}_T = \sum_{i=1}^{N_n} \hat{v}_n(x_i^n) \mathbb{P}(X \in C_i(\hat{X}_n))$$

and for intermediate times t_k, the value is:

$$\hat{V}_{t_k} = \sum_{i=1}^{N_k} \hat{v}_k(x_i^k) \mathbb{P}(X \in C_i(\hat{X}_k)) \text{ for } 1 \leq k \leq n-1$$

where $\pi_{ij}^{k+1} = \mathbb{P}(\hat{X}_{k+1} = x_j^{k+1} | \hat{X}_k = x_i^k)$ are the transition probabilities.

These can be computed as:
We first define:

$$\begin{cases} \hat{p}_i^k = \mathbb{P}(\hat{X}_k = x_i^k) & 1 \leq k \leq n-1, 1 \leq i \leq N_k \\ \hat{p}_{ij}^k = \mathbb{P}(\hat{X}_{k+1} = x_j^{k+1}, \hat{X}_k = x_i^k) & 1 \leq j \leq N_{k+1} \end{cases}$$

then

$$\pi_{ij}^{k+1} = \frac{p_{ij}^k}{p_i^k}$$

The challenge is to compute such probabilities. If the underlying variable has a density, the problem becomes easier. Otherwise we need to estimate the transition probabilities by Monte Carlo. See [3] for more details regarding pricing Bermudan options using quantization.

Conclusion

In this paper, we have discussed the pricing of Bermudan options using the stochastic grid bundling method (SGBM) which belongs to the regression-based methods. We presented its difference with the classic Least square Monte Carlo then detailed some of the popular techniques used for bundling. These are generally machine learning algorithms. We also showed that SGBM can be very easy to implement when the characteristic function of the underlying variable has a closed formulae. This has been illustrated in the case of the CIR process which is one of the most used processes in practice especially in the modeling of interest rates dynamics. We also introduced a method based on the approximation of transition probabilities which is Quantization.

Pricing Bermudan options still be among the greatest practical challenges facing users of Monte Carlo methods in the early-exercise derivatives pricing industry. The proposed method is an alternative to the classic least square Monte Carlo method (LSM) that is generally used for this purpose.

Compared to LSM, the approximate option values computed using SGBM have lower numerical noise, not just at the initial step but also at the intermediate time steps; which makes it more accurate and faster for computations that require option values at intermediate time steps such as computing future exposures within the CVA context.

Another favorable property of SGBM is that it can be used to get fast approximations of the sensitivities or Greeks of the option price which is essential for hedging and risk management, but it typically requires substantially more computing time than pricing the derivative.

Appendix

The CIR model specifies that the instantaneous interest rate follows the below stochastic differential equation:

$$dX_t = -\lambda(X_t - \theta)dt + \sigma\sqrt{X_t}dB_t$$

where B_t is a Brownian Motion and the parameter λ corresponds to the speed of adjustment, θ to the mean and σ to volatility.
The standard deviation factor, $\sigma\sqrt{X_t}$ avoids the possibility of negative interest rates for all positive values of λ and θ. An

interest rate of zero is also avoided if the following condition is met:

$$2\lambda\theta \geq \sigma^2.$$

The aim of the following section is to find the characteristic function of such process. We write $u = iy$ and make the assumption that the characteristic function of X_t is of exponentially-affine form, that is to say, we will show that for some T, we have:

$$\mathbb{E}[exp(iyX_T)] = \mathbb{E}[exp(uX_T)] = exp\left(A(T,u) + B(T,u)X_0\right)$$

More precisely, we are seeking functions $A(t,u)$ and $B(t,u)$ with initial conditions:

$$A(0,u) = 0 \text{ and } B(0,u) = u,$$

such that:

$$M_t = f(t,X_t) = exp\left(A(T-t,u) + B(T-t,u)X_t\right)$$

is a martingale.
Once the above functions are found, we can write:

$$\mathbb{E}[M_T] = M_0.$$

By replacing M_T and M_0 by their expressions, the above equation becomes:

$$\mathbb{E}[exp(uX_T)] = exp\left(A(T,u) + B(T,u)X_0\right).$$

Now, we are going to demonstrate that.
Applying Ito's formulae to the function $f(t,X_t)$ yields:

$$\frac{df(t,X_t)}{f(t,X_t)} = -\left(\frac{dA_{T-t}}{dt} + X_t\frac{dB_{T-t}}{dt}\right)dt + B_{T-t}dX_t + \frac{1}{2}B_{T-t}^2\sigma^2 X_t dt$$

$$\begin{aligned}\frac{df(t,X_t)}{f(t,X_t)} &= -\left(\frac{dA_{T-t}}{dt} + X_t\frac{dB_{T-t}}{dt}\right)dt - B_{T-t}\lambda(X_t-\theta)dt \\ &+ B_{T-t}\sigma\sqrt{X_t}dB_t + \frac{1}{2}B_{T-t}^2\sigma^2 X_t dt\end{aligned}$$

$f(t,X_t)$ is a local martingale if the drift coefficient is null which means:

$$\left(\frac{dA_{T-t}}{dt} + X_t\frac{dB_{T-t}}{dt}\right) = -B_{T-t}\lambda(X_t-\theta) + \frac{1}{2}B_{T-t}^2\sigma^2 X_t$$

Since the previous equation is valid for all $t \geq 0$ and both sides are affine functions of X_t, we get:

$$\begin{cases} \frac{dA_{T-t}}{dt} = \lambda\theta B_{T-t} \\ \frac{dB_{T-t}}{dt} = -\lambda B_{T-t} + \frac{1}{2}B_{T-t}^2\sigma^2 \end{cases}$$

We need to solve the previous equations with initial conditions:

$$A(0,u) = 0 \text{ and } B(0,u) = u$$

The differential equation for B is called a Bernoulli equation which is a special case of Riccati equation. It is a nonlinear differential equation with known exact solution. We refer to [4] for its solution. Once B is found, A equation is straight forward.
Finally the solution is:

$$\begin{cases} B(t,u) = \frac{ue^{-\lambda t}}{1-\frac{\sigma}{2\lambda}u(1-e^{-\lambda t})} \\ A(t,u) = -\frac{2\lambda\theta}{\sigma^2}log\left(1-\frac{\sigma^2}{2\lambda}u(1-e^{-\lambda t})\right) \end{cases}$$

Thus, the characteristic function $\mathbb{E}[exp(iyX_t)]$ of the CIR process is given by:

$$\left(1-\frac{\sigma^2}{2\lambda}(1-e^{-\lambda t})iy\right)^{-\frac{2\lambda\theta}{\sigma^2}} exp\left(\frac{ue^{-\lambda t}}{1-\frac{\sigma}{2\lambda}iy(1-e^{-\lambda t})}X_0\right)$$

References

[1] Paul Glasserman. Monte carlo methods in financial engineering. *Springer*, 2004.

[2] Eduardo S. Schwartz Francis A. Longstaff. Valuing american options by simulation: A simple least-squares approach. *The Review of Financial Studies*, 2001.

[3] Jacques PRINTEMS Vlad BALLY, Gilles PAGES. A quantization tree method for pricing and hedging multidimensional american options. 2001.

[4] Michiel Hazewinkel. *Riccati equation.* Kluwer Academic Publishers, 2001.

A propos d'Awalee

Cabinet de conseil indépendant spécialiste du secteur de la Finance.

Nous sommes nés en 2009 en pleine crise financière. Cette période complexe nous a conduits à une conclusion simple : face aux exigences accrues et à la nécessité de faire preuve de souplesse, nous nous devions d'aider nos clients à se concentrer sur l'essentiel, à savoir leur performance.

Pour accomplir cette mission, nous nous appuyons sur trois ingrédients : habileté technique, savoir-faire fonctionnel et innovation.

Ceci au service d'une ambition : dompter la complexité pour simplifier la vie de nos clients.

«*Run the bank*» avec Awalee !

Contactez-nous

Ronald LOMAS
Partner
rlomas@awaleeconsulting.com
06 62 49 05 97