Numerical Methods for Financial Derivatives

Hwan C. Lin

Department of Economics
University of North Carolina at Charlotte

Lecture 5: Monte Carlo Simulation for American Options (Ch. 3)

Introduction

- We discuss risk-neutral pricing of American options.
- We present two classes of Monte Carlo simulation methods for the pricing American options. methods.
  - Parametric Methods
  - Regression Methods

Risk-Neutral Valuation of European Options

- The risk-neutral valuation of European options is

$$V(S_0, 0) = e^{-rT} E_Q[\Psi(S_T)|S_0] \tag{1}$$

  where

  - Q is the risk neutral probability measure;
  - r is riskless interest rate;
  - $\Psi(S_T)$ denotes the payoff;
  - S_0 is the time 0 of the underlying asset (S);
  - S_T is the price of S at maturity date T.

- We have discussed how to apply Monte Carlo simulation for the pricing of European options.

Risk-Neutral Valuation of American Options

- The risk-neutral valuation of American options is

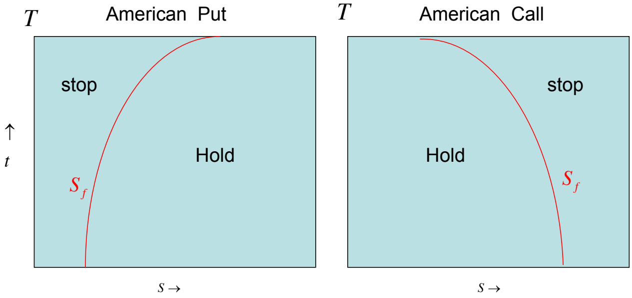
$$V(S_0, 0) = \sup_{0 \leq \tau \leq T} e^{-r\tau} E_Q[\Psi(S_\tau) | S_0] \tag{2}$$

- where
  - τ is stopping time.
  - $\Psi(S_\tau)$ is the payoff when the option is exercised optimally.

Definition (Stopping Time)

A stopping time τ with respect to a filtration $\mathcal{F}_t$ is a random variable, which is $\mathcal{F}_t$-measurable for all t.

Early Exercise Curves S_f



- At any moment $t < T$, if the American option is not exercised, its continuation value is

$$V(S_t, t)^{Cont} = \sup_{t \leq \tau \leq T} e^{-r\tau} E_Q[\Psi(S_\tau)|S_t]$$

- $V(S_t, t) = \Psi(S_t)$ if $\Psi(S_i) \geq V(S_t, t)^{Cont}$:
  - This occurs when S_t is *stopping region*.
  - One encounters the "stopping time."
- $V(S_t, t) = V(S_t, t)^{Cont}$ if $\Psi(S_i) < V(S_t, t)^{Cont}$:
  - This occurs when S_t is in *holding region*.

- **Lower Bound** to$V(S_0, 0)$: In terms of (2), the value of American options, $V(S_0, 0)$, is given by taking the supremum over all stopping times. Thus a lower bound to the $V(S_0, 0)$ can be obtained by taking a specific stopping strategy. To illustrate this idea, we choose the stopping strategy to be the level of parameterβ. If the resulting stopping time for each simulated path is denoted by $\tilde{\tau} \in [0, T]$. A lower bound to $V(S_0, 0)$ is given by

$$V^{low(\beta)}(S_0, 0) = E_Q[e^{-r\tilde{\tau}}\Psi(S_{\tilde{\tau}})|S_0] \tag{3}$$

- **Approximation:** $V(S_0, 0)$ can be approximated by:

$$\sup_{\beta} V^{low(\beta)} \approx V(S_0, 0)$$

Parametric Methods for American Options (2)

- **Application of the Parametric Method**:
  1. Construct a curve depending on a parameter vector β such that the curve approximates the early-exercise curve.
  2. The stopping strategy is to stop when the path S_t crosses the curve defined by β.
  3. For N such paths, evaluate the payoff and evaluate (approximate) the value $V^{low(\beta)}$.
  4. Next, attempt to maximize the value $V^{low(\beta)}$ by repeating the procedure for a "better" β vector.
  5. To complete the procedure, one should also construct an upper bound V^{up}. As a crude example, the entire path S_t for $t \in [0, T]$ may be simulated and the option is exercised **in retrospect** when
  $$e^{-rt}\Psi(S_t)$$
  is maximal.
- As a by-product of approximating $V(S_0, 0)$, the corresponding parameters β provide an approximation of the early exercise curve.
- But this is just a crude method and an optimization in the β parameter space is costly.
- **Literature**: P. Glasserman: Monte Carlo Methods in Financial Engineering (2004).

Introduction to Regression Methods

The central Ideas of approximating the value of American options using regression methods:

- Use the value of a **Bermuda** option as an approximation of the **American** option
- The value of the Bermudan option is calculated **recursively** in a backward fashion, and

$$V^{Am}(S_{t_j}, t_j) \approx V^{Be}(S_{t_j}, t_j) = \max\{\Psi(S_{t_j}), V^{cont}\} = \max\{excise, hold\}$$

- That is, at each t_j, the holder of the option decides which of the two possibilities *exercise*, *hold* is optimal based on the principle of **dynamic programming**.
- For a Bermuda option, we define the continuation value V^{cont} at t_j according to

$$C_j(x) \equiv e^{-r\triangle t} E_Q[V(S_{t_{j+1}}, t_{j+1}) | S_{t_j} = x],$$

and these functions $C_j(x)$ are approximated at each t_j by **least squares regression**.

Bermuda Options as Approximation of American Options

Definition (Bermuda Options)

A Bermudan option is an option that can be exercised only at a finite number M of discrete time instances t_j.

- Specifically, for $t_j = j \triangle t$, $\triangle t = \frac{T}{M}$, $j = 0, 1, \ldots, M$, we denote by $V^{Be(M)}$ the value of a Bermudan option. Because of the additional exercise possibilities, it holds that

$$V^{Eur} \leq V^{Be(M)} \leq V^{Am} \tag{4}$$

and one can show that

$$\lim_{M \to \infty} V^{Be(M)} = V^{Am}.$$

- For suitable M the value $V^{Be(M)}$ is used as approximation of V^{Am}.

Use of Dynamic Programming Principle

Principle 3.11 (Dynamic Programming)

- Set $V_M(x) = \Psi(x)$.
- For $i = M-1, \ldots, 1$, calculate $C_i(x)$ for $x > 0$ and

$$V_i(x) \equiv V(x, t_i) = \max\{\Psi(x), C_i(x)\}$$

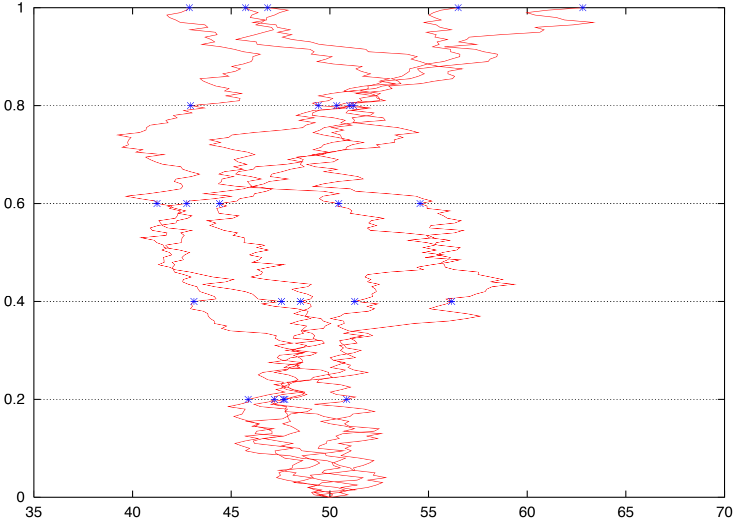
- For $i = 0$, calculate

$$V_0 \equiv V(S_0, 0) = C_0(S_0)$$

- Note:
  - The $C_i(x)$ are calculated by least squares (see Appendix C4 of the Textbook)
  - Sample paths $S_i = x \in [0, \infty)$ are calculated starting from S_0 according to the underlying risk-neutral model.

Monte Carlo Setting for Regression Methods

$N = 5$ trajectories; $M = 5$ exercise times; horizontal axis: S; vertical axis: t; * marked as S_{ik}



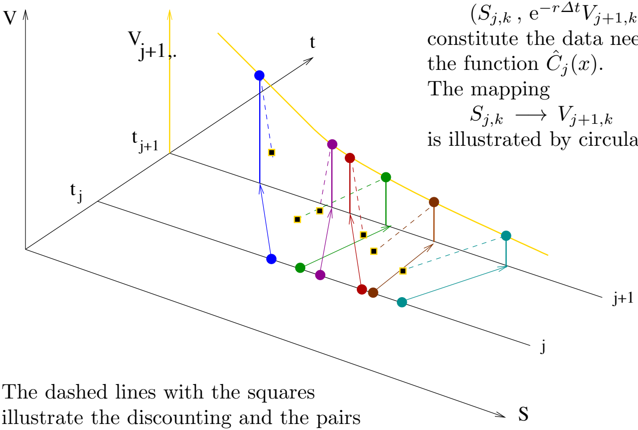
Regression I

Algorithm 3.12 (regression I)

1. Simulate N paths $S_1(t), \ldots, S_N(t)$ and store the values
$$S_{ik} \equiv S_k(t_i), \quad i = 1, \ldots, M, \quad k = 1, \ldots, N$$
2. For $i = M$, set $V_{MK} \equiv \Psi(S_{Mk})$ for all k.
3. For $i = M - 1, \ldots, 1$:
   1. Approximate $C_i(x)$ using suitable basis functions $\phi_0, \ldots, \phi_L$
   $$C_i(x) \approx \sum_{l=0}^{L} a_l \phi_l(x) \equiv \hat{C}_i(x) \tag{5}$$
   by least squares over the N points:
   $$(x_k, y_y) \equiv (S_{ik}, e^{-r\triangle t} V_{i+1,k}), \quad k = 1, \ldots N,$$
   2. and set $V_{ik} = \max\{\Psi(S_{ik}), C_i(\hat{S_{ik}})\}$.
4. Calculate $V_0 = e^{-r\triangle t} \frac{1}{N}(V_{11} + \ldots + V_{1N})$

Schematic Illustration of Regression I

Mapping C, illustration for $N = 6$



The N pairs $(k = 1, \ldots, N)$

$$(S_{j,k}\ ,\ \mathrm{e}^{-r\Delta t} V_{j+1,k})$$

constitute the data needed to calculate the function $\hat{C}_j(x)$.

The mapping

$$S_{j,k} \longrightarrow V_{j+1,k}$$

is illustrated by circular dots.

The dashed lines with the squares illustrate the discounting and the pairs $(S_{j,k}\ ,\ \mathrm{e}^{-r\Delta t} V_{j+1,k})$, which enter the least squares procedure to generate $\hat{C}_j(x)$.

- Source: Rüdiger Seydel, Tools of Computational Finance, chapter (version 2014)

Regression II

Algorithm 3.13 (regression II)

1. Simulate N paths $S_1(t), \ldots, S_N(t)$ and store the values

$$S_{ik} \equiv S_k(t_i), \quad i = 1, \ldots, M, \quad k = 1, \ldots, N$$

2. For $i = M$, set $g_K \equiv \Psi(S_{Mk})$, $\tau_k = M$ for all k.
3. For $i = M - 1, \ldots, 1$:
   1. For the subset of in-the-money points

   $$(x_k, y_y) \equiv (S_{ik}, e^{-r(\tau_k - i)\triangle t} g_k), \quad \{k\} \subseteq \{1, \ldots N\},$$

   approximate $C_i(x)$ by $\hat{C}_i(S_{ik})$ using (5),
   2. and if $\Psi(S_{ik}) \geq \hat{C}_i(S_{ik})$, update cash flow and exercise time:

   $$g_k = \Psi(S_{ik}), \quad \tau_k = i$$

4. Calculate $\hat{C}_0 = \frac{1}{N} \sum_{k=1}^{N} e^{-r\triangle t} g_k$, $V_0 = \max\{\Psi(S_0), \hat{C}_0\}$

- Denote

$$\hat{\mu} \equiv \frac{1}{N}\sum_{k=1}^{N} f(x_k), \qquad \hat{s}^2 = \frac{1}{N-1}\sum_{k=1}^{N}[f(x_k) - \hat{\mu}]^2,$$

  and $E(\hat{\mu}) = \mu$.
- According to the **central limit theorem**, the approximation $\hat{\mu}$ obeys $\mathcal{N}(\mu, \sigma)$ with distribution function F:

$$F(a) = P(\hat{\mu} - \mu \leq a\frac{\sigma}{\sqrt{N}}).$$

- In practice, σ^2 is replaced with its approximation $\hat{s}^2$ and the error behaves as $\hat{s}/\sqrt{N}$.
- How to reduce this sampling error $\hat{s}/\sqrt{N}$ in Monte Carlo simulation? There are choices:
  1. Reduce the numerator (*variance reduction*);
  2. Enlarge the denominator. This means to increase the number of simulations (N), and is very costly.

Accuracy of Monte Carlo Simulation(2)

Another error, the bias

- In several cases, the computation of $f(x_i)$ gives rise to another error, namely, the **bias**.
- Let $\hat{x}$ be an estimator of x, then the bias is defined as

$$bias(\hat{x}) = E[\hat{x}] - x$$

- Example 1: For a lookback option, the payoff involves the variable

$$x \equiv E\left[\max_{0 \leq t \leq T} \{S_t\}\right]$$

An approximation is

$$\hat{x} = \max_{0 \leq j \leq M} \{S_{t_j}\}$$

clearly, $\hat{x} \leq x$ almost surely; i.e., $E[\hat{x}] < x$. Hence $bias \neq 0$.

- Example 2: Compared to the analytic solution of GBM, the explicit Euler method provides biased results. For GBM,

$$S_{t_{j+1}} = S_{t_j} \exp\left[(r - \frac{1}{2}\sigma^2)\triangle t + \sigma \triangle W\right]$$

is unbiased, whereas the explicit Euler method

$$S_{t_{j+1}} = S_{t_j} + rS_{t_j}\triangle t + \sigma S_{t_j}\triangle W), \quad j = 0, 1, \ldots, M$$

is biased, with a bias due to the weak error.

Accuracy Enhancement

- The overall error is measured by the **mean squared error**:

$$\begin{aligned} MSE(\hat{x}) &= E[(\hat{x}-x)^2] \\ &= (E[\hat{x}]-x)^2 + E[(\hat{x}-E[\hat{x}])^2] \\ &= \underbrace{(bias(x))^2}_{approximation\ error} + \underbrace{Var(\hat{x})}_{sampling\ error} \end{aligned}$$

- To reduce the errors, one can apply one of the following possibilities, or all of them:

1. apply variance reduction $\Rightarrow \hat{s}$ smaller in $\hat{s}/\sqrt{N}$
2. increase N (i.e. more sample paths) $\Rightarrow N$ larger in $\hat{s}/\sqrt{N}$
3. increase M (i.e. $\triangle t$ smaller) $\Rightarrow$ reduce the bias.

- We must compare costs and benefits of accuracy.

There are several methods of variance reduction. The simplest (and maybe the least powerful) is the method of **antithetic variates**:

- Let us denote by $\hat{V}$ the MC approximation of an European option, for instance.
- **The idea of antithetic variates** is to use in parallel the numbers $-Z_1$, $-Z_2$, ..., which are also $\sim \mathcal{N}(0,1)$, to calculate "**mirror paths**" S_t^-, from which the payoff values $\Psi(S_T^-)$ are calculated. This leads to a second Monte Carlo value V^-. It turns out that the mean $V_{AV} \equiv \frac{1}{2}(\hat{V} + V^-)$ carries a much smaller variance:

$$Var(V_{AV}) < \frac{1}{2} Var(\hat{V})$$

Methods of Variance Reduction (2)

- Why does V_{AV} carry a much smaller variance?
- $$\begin{aligned} Var(V_{AV}) &= \frac{1}{4} Var(\hat{V} + V^-) \\ &= \frac{1}{4}\left[Var(\hat{V}) + Var(V^-) + 2Cov(\hat{V}, V^-)\right] \\ &= \frac{1}{2} Var(\hat{V}) + \frac{1}{2} Cov(\hat{V}, V^-), \; Var(\hat{V}) = Var(V^-) \end{aligned}$$
- Given $Cov(\hat{V}, V^-) < 0$, we have $Var(V_{AV}) < \frac{1}{2} Var(\hat{V})$.
- This approach at most **doubles** the costs. In comparison, an error reduction ($factor < \frac{1}{2}$) by merely increasing N requires at least **fourfold** costs.