Advanced Monte Carlo Methods: American Options

Prof. Mike Giles

mike.giles@maths.ox.ac.uk

Oxford University Mathematical Institute

Early Exercise

Perhaps the biggest challenge for Monte Carlo methods is the accurate and efficient pricing of options with optional early exercise:

- Bermudan options: can exercise at a finite number of times t_j
- American options: can exercise at any time

The challenge is to find/approximate the optimal strategy (i.e. when to exercise) and hence determine the price and Greeks.

Early Exercise

Approximating the optimal exercise boundary introduces new approximation errors:

- An approximate exercise boundary is inevitably sub-optimal
  $\Longrightarrow$ under-estimate of "true" value, but accurate value for the sub-optimal strategy
- For the option buyer, sub-optimal price reflects value achievable with sub-optimal strategy
- For the option seller, "true" price is best a purchaser might achieve
- Can also derive an upper bound approximation

Early Exercise

Why is early exercise so difficult for Monte Carlo methods?

- leads naturally to a dynamic programming formulation working backwards in time
- fairly minor extension for finite difference methods which already march backwards in time
- doesn't fit well with Monte Carlo methods which go forwards in time

Problem Formulation

Following description in Glasserman's book, the Bermudan problem has the dynamic programming formulation:

$$\begin{aligned} \widetilde{V}_m(x) &= \widetilde{h}_m(x) \\ \widetilde{V}_{i-1}(x) &= \max\left(\widetilde{h}_{i-1}(x), \mathbb{E}[D_{i-1,i}\, \widetilde{V}_i(X_i) \mid X_{i-1} = x]\right) \end{aligned}$$

where

- X_i is the underlying at exercise time t_i
- $\widetilde{V}_i(x)$ is option value at time t_i assuming not previously exercised
- $\widetilde{h}_i(x)$ is exercise value at time t_i
- $D_{i-1,i}$ is the discount factor for interval $[t_{i-1}, t_i]$

Problem Formulation

By defining

$$\begin{aligned} h_i(x) &= D_{0,i}\, \widetilde{h}_i(x) \\ V_i(x) &= D_{0,i}\, \widetilde{V}_i(x) \end{aligned}$$

where

$$D_{0,i} = D_{0,1}\ D_{1,2}\ \ldots\ D_{i-1,i}$$

can simplify the formulation to

$$\begin{aligned} V_m(x) &= h_m(x) \\ V_{i-1}(x) &= \max\left(h_{i-1}(x), \mathbb{E}[V_i(X_i) \mid X_{i-1} = x]\right) \end{aligned}$$

Problem Formulation

An alternative point of view considers stopping rules τ, the time at which the option is exercised.

For a particular stopping rule, the initial option value is

$$V_0(X_0) = \mathbb{E}[h_\tau(X_\tau)],$$

the expected value of the option at the time of exercise.

The best that can be achieved is then

$$V_0(X_0) = \sup_\tau \mathbb{E}[h_\tau(X_\tau)]$$

giving an optimisation problem.

Problem Formulation

The continuation value is

$$C_i(x) = \mathbb{E}[V_{i+1}(X_{i+1}) \mid X_i = x]$$

and so the optimal stopping rule is

$$\tau = \min\{i : h_i(X_i) > C_i(X_i)\}$$

Approximating the continuation value leads to an approximate stopping rule.

Longstaff-Schwartz Method

The Longstaff-Schwartz method (2001) is the one most used in practice.

Start with N path simulations, each going from initial time $t=0$ to maturity $t=T=t_m$.

Problem is to assign a value to each path, working out whether and when to exercise the option.

This is done by working backwards in time, approximating the continuation value.

Longstaff-Schwartz Method

At maturity, the value of an option is

$$V_m(X_m) = h_m(X_m)$$

At the previous exercise date, the continuation value is

$$C_{m-1}(x) = \mathbb{E}[V_m(X_m) \mid X_{m-1} = x]$$

This is approximated using a set of R basis functions as

$$\widehat{C}_{m-1}(x) = \sum_{r=1}^{R} \beta_r \, \psi_r(x)$$

Longstaff-Schwartz Method

The coefficients β_r are obtained by a least-squares minimisation, minimising

$$\mathbb{E}\left\{\left(\mathbb{E}[V_m(X_m) \mid X_{m-1}] - \widehat{C}_{m-1}(X_{m-1})\right)^2\right\}$$

Setting the derivative w.r.t. β_r to zero gives

$$\mathbb{E}\left\{\left(\mathbb{E}[V_m(X_m) \mid X_{m-1}] - \widehat{C}_{m-1}(X_{m-1})\right) \psi_r(X_{m-1})\right\} = 0$$

and hence

$$\begin{aligned} \mathbb{E}[V_m(X_m)\, \psi_r(X_{m-1})] &= \mathbb{E}[\widehat{C}_{m-1}(X_{m-1})\, \psi_r(X_{m-1})] \\ &= \sum_s \mathbb{E}[\psi_r(X_{m-1})\, \psi_s(X_{m-1})]\, \beta_s \end{aligned}$$

Longstaff-Schwartz Method

This set of equations can be written collectively as

$$B_{\psi\psi}\,\beta = B_{V\psi}$$

where

$$(B_{V\psi})_r = \mathbb{E}[V_m(X_m)\,\psi_r(X_{m-1})]$$

$$(B_{\psi\psi})_{rs} = \mathbb{E}[\psi_r(X_{m-1})\,\psi_s(X_{m-1})]$$

Therefore,

$$\beta = B_{\psi\psi}^{-1}\,B_{V\psi}$$

Longstaff-Schwartz Method

In the numerical approximation, each of the expectations is replaced by an average over the values from the N paths.

For example,

$$\mathbb{E}[\psi_r(X_{m-1})\,\psi_s(X_{m-1})]$$

is approximated as

$$N^{-1}\sum_{n=1}^{N}\psi_r(X_{m-1}^{(n)})\;\psi_s(X_{m-1}^{(n)})$$

Assuming that the number of paths is much greater than the number of basis functions, the main cost is in approximating $B_{\psi\psi}$ with a cost which is $O(N\,R^2)$.

Longstaff-Schwartz Method

Once we have the approximation for the continuation value, what do we do?

- if $\widehat{C}(X_{m-1}) < h_{m-1}(X_{m-1})$, exercise the option and set

$$V_{m-1} = h_{m-1}(X_{m-1})$$

- if not, then <u>either</u> set

$$V_{m-1} = \widehat{C}(X_{m-1})$$

  (Tsitsiklis & van Roy, 1999), <u>or</u>

$$V_{m-1} = V_m$$

  (Longstaff & Schwartz, 2001)

Longstaff-Schwartz Method

The Longstaff-Schwarz treatment only uses the continuation estimate to decide on the exercise boundary – no loss of accuracy for paths which are not exercised.

The Tsitsiklis-van Roy treatment introduces more error, especially for American options where it gets applied each timestep.

Also, Longstaff-Schwarz can do least squares fit only for paths which are in-the-money (i.e. $h(X) > 0$) – leads to improved accuracy.

Because of the optimality condition, the option value is insensitive to small perturbations in the exercise boundary, so can assume that exercise of paths is not affected when computing first order Greeks.

Longstaff-Schwartz Method

Provided the basis functions are chosen suitably, the approximation

$$\widehat{C}_{m-1}(x) = \sum_{r=1}^{R} \beta_r \, \psi_r(x)$$

gets increasingly accurate as $R \to \infty$. Longstaff & Schwartz used 5-20 basis functions in their paper – I don't know what is standard now in practice.

Having completed the calculation for t_{m-1}, repeat the procedure for t_{m-2} then t_{m-3} and so on. Could use different basis functions for each exercise time – the coefficients β will certainly be different.

Longstaff-Schwartz Method

The estimate will tend to be biased low because of the sub-optimal exercise boundary, however might be biased high due to using the same paths for decision-making and valuation.

To be sure of being biased low, should use two sets of paths, one to estimate the continuation value and exercise boundary, and the other for the valuation.

However, in practice the difference is quite small.

This leaves the problem of computing an upper bound.

Upper Bounds

In Glasserman's Bermudan version of Roger's continuous time result (2002), he lets M_m be a martingale with $M_0 = 0$.

For any stopping rule τ, we have

$$\mathbb{E}[h_\tau(X_\tau)] \; = \; \mathbb{E}[h_\tau(X_\tau) - M_\tau] \; \leq \; \mathbb{E}[\max_k (h_k(X_k) - M_k)]$$

This is true for all martingales M and all stopping rules τ and hence

$$V_0(X_0) \; = \; \sup_\tau \mathbb{E}[h_\tau(X_\tau)] \; \leq \; \inf_M \mathbb{E}[\max_k (h_k(X_k) - M_k)]$$

Upper Bounds

The key duality result is that in fact there is equality

$$\sup_\tau \mathbb{E}[h_\tau(X_\tau)] \;=\; \inf_M \mathbb{E}[\max_k (h_k(X_k) - M_k)]$$

so that

- an arbitrary τ gives a lower bound
- an arbitrary M gives an upper bound
- making both of them “better” shrinks the gap between them to zero

Upper Bounds

Glasserman proves by induction that the optimal martingale M is equal to

$$M_k = \sum_1^k \Big(V_i(X_i) - \mathbb{E}[V_i(X_i) \mid X_{i-1}] \Big)$$

To get a good upper bound we approximate this martingale.

Upper Bounds

The approximate martingale for a particular path is defined as

$$\widehat{M}_k = \sum_1^k \left(V_i(X_i) - P^{-1} \sum_p V_i(X_i^{(p)}) \right)$$

where the $X_i^{(p)}$ are values for X_i from P different mini-paths starting at X_{i-1}, and

$$V_i(X_i) = \max(h_i(X_i), \widehat{C}_i(X_i))$$

with $\widehat{C}_i(X_i)$ being the approximate continuation value given by the Longstaff-Schwartz algorithm.

Glasserman suggests up to 100 mini-paths may be needed.

Final Words

- Bermudan and American options are important applications
- Longstaff-Schwartz method is popular, but still plenty of scope for improvement?
- suspect that finite difference method is used for Greeks?
- is independent second set of paths used in practice?
- are upper bounds used in practice?