Financial Mathematics 32000

Lecture 7

Roger Lee

2023 May 10

MC for American options

Can MC be used for Americans?

The 1993 edition of a leading derivatives textbook:

> *"One limitation of the Monte Carlo simulation approach is that it can be used only for European-style derivative securities"*

In later editions, this statement was removed.

Francis Longstaff and Eduardo Schwartz

Approximate American by Bermudan

Consider an American option exercisable at any time in $[0, T]$.

- Let Y_t be the discounted (to time 0) payoff for a time-t exercise. Assume Y is some given function of a Markov state variable (or vector) X. Define Y_T to incorporate the correct time-T decision. Example: for American put on X, let $Y_T := e^{-rT}(K - X_T)^+$ and let $Y_t = e^{-rt}(K - X_t)$ for $t < T$.
- Approximate the American put by allowing exercise only at a finite number of times:

$$0 = t_0 < t_1 < \cdots < t_N = T$$

So we approximate the American by a *Bermudan*.

Bermudan facts

1. The time-0 value of the Bermudan is

$$\max_{\tau} \mathbb{E} Y_\tau$$

where τ ranges over stopping times taking values in $\{t_0, t_1, \ldots, t_N\}$. So τ is a random exercise time that can depend on X, but only in a non-anticipating way.

2. Define $\tau_N := T$. Idea is that τ_n is the optimal exercise time given no exercise prior to t_n; thus τ_n is a stopping time taking values in $\{t_n, t_{n+1}, \ldots, t_N\}$. Given τ_{n+1}, define τ_n by

$$\tau_n := t_n \mathbf{1}\Big(Y_{t_n} \geq \mathbb{E}(Y_{\tau_{n+1}} | X_{t_n})\Big) + \tau_{n+1} \mathbf{1}\Big(Y_{t_n} < \mathbb{E}(Y_{\tau_{n+1}} | X_{t_n})\Big).$$

Then τ_0 attains the max in 1.

Continuation value function

But we need to estimate the *continuation value function*

$$f_n(x) := \mathbb{E}(Y_{\tau_{n+1}} | X_{t_n} = x)$$

Longstaff-Schwartz: Use simulation and least-squares regression.

- Simulate M paths of X. Each path includes $t_0, t_1, \ldots, t_N$.
- Define $\hat{\tau}_N := T$. On each path $m = 1, \ldots, M$, given $\hat{\tau}^m_{n+1}$, define

  $$\hat{\tau}^m_n := t_n \mathbf{1}\Big(Y^m_{t_n} \geq \hat{f}_n(X^m_{t_n})\Big) + \hat{\tau}^m_{n+1} \mathbf{1}\Big(Y^m_{t_n} < \hat{f}_n(X^m_{t_n})\Big)$$

  where $\hat{f}_n$ is a "least-squares" estimate of continuation value.

  If f_n known to be ≥ 0, can use $\hat{f}_n(X^m_{t_n})^+$ instead of $\hat{f}_n(X^m_{t_n})$.
- Estimate of the time-0 option price is

  $$\hat{C}^{\text{LS}}_M := \frac{1}{M} \sum_m Y^m_{\hat{\tau}^m_0}$$

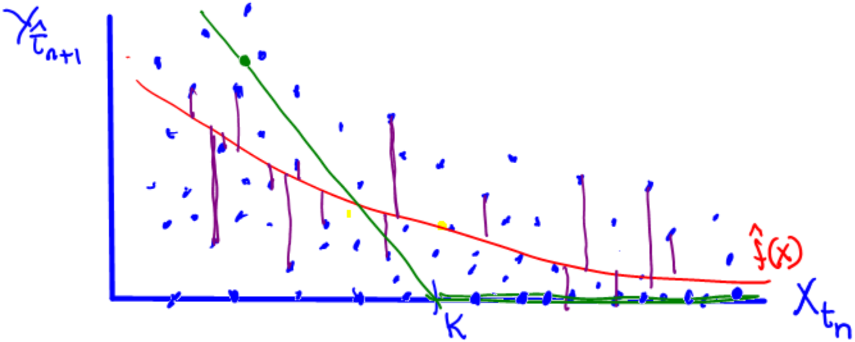
Estimated continuation value function

How to obtain the *estimated continuation value function* $\hat{f}_n$:

Each of the M paths of X generates a pair $(X^m_{t_n}, Y^m_{\hat{\tau}^m_{n+1}})$.

Intuition: Consider a plot of all M pairs.

We want to extract the conditional expectation from the cloud.



Estimated continuation value function

- Choose $\hat{f}_n(x)$ as a linear combination of some set of B basis functions. (Example: with $B = 3$ and basis functions 1, x, x^2 and coefficients $\hat{\beta}$, we have $\hat{f}_n(x) = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2$)
- Choose coefficients $\hat{\beta}$ to minimize the sum of squared residuals

$$\sum_m \left(\hat{f}_n(X^m_{t_n}) - Y^m_{\hat{\tau}^m_{n+1}} \right)^2 .$$

  To implement this, run an ordinary least squares linear regression of the simulated $Y_{\hat{\tau}_{n+1}}$ on the basis functions (evaluated at the simulated X_{t_n}).

  For $n = 0$, regression unnecessary; $\hat{f}_0(X^m_{t_0})$ is just $\frac{1}{M} \sum_{m=1}^{M} Y^m_{\hat{\tau}^m_1}$.
- (Open questions: Optimal choice of basis? Alternatives to OLS?)

Implementation

In Python: `sklearn.linear_model.LinearRegression` or `statsmodels.api.OLS` or `scipy.linalg.lstsq(A, y)`

where `A` is the $M \times B$ matrix of regressors, and `y` is the length-M vector of simulations of continuation payoffs $Y^m_{\hat{\tau}^m_{n+1}}$. Example:

$$\mathtt{A} = \begin{pmatrix} 1 & X^{(1)}_{t_n} & \left[X^{(1)}_{t_n}\right]^2 \\ 1 & X^{(2)}_{t_n} & \left[X^{(2)}_{t_n}\right]^2 \\ \vdots & \vdots & \vdots \\ 1 & X^{(M)}_{t_n} & \left[X^{(M)}_{t_n}\right]^2 \end{pmatrix} \qquad \mathtt{y} = \begin{pmatrix} Y^{(1)}_{\hat{\tau}^{(1)}_{n+1}} \\ Y^{(2)}_{\hat{\tau}^{(2)}_{n+1}} \\ \vdots \\ Y^{(M)}_{\hat{\tau}^{(M)}_{n+1}} \end{pmatrix}$$

In Excel: `LINEST(y, A)` but omit constant column from `A`

Also note that Excel outputs the coefficients in the *opposite* order.

Ignore at each time step the OTM data?

L-S: at each time step, include in the regression only the paths which are ITM at that time.

- We don't care about the continuation value function f_n at points X which are OTM, because we won't exercise there anyway.
- We care about getting a good estimate of f_n at points ITM.
- Discarding OTM points allows us to fit better the ITM points. This can help especially if our parametric form does not capture well the behavior of f_n across the whole range of X.
  But throwing away OTM observations also reduces sample size, magnifying the effect of the randomness in the ITM observations.

Bias

Let $\hat{C}^{\mathrm{LS}}$ be the MC estimate. Bias is by definition

$$\mathbb{E}\hat{C}^{\mathrm{LS}} - C$$

where C be the true value of the American.

- Bias due to pricing Bermudan instead of American: $\downarrow$
- Bias in Bermudan price due to using the *same paths* to estimate optimal exercise policy and to price the Bermudan: $\uparrow$
  Can eliminate by segregating paths: $\hat{f}$-estimation and pricing.
  (Similar idea as *out-of-sample* testing and *cross-validation*)
- Bias in Bermudan price due to failure of basis functions to span the true continuation value function, and due to randomness in the $\hat{f}$-estimation simulations: $\downarrow$