Pricing High Dimensional American-Style Options A Classification - Monte Carlo (CMC) Approach

JORGE A. PICAZO
PhD Candidate
Department of Statistics
Stanford University

Advisor: Professor Art B. Owen

Some Basic Concepts

Call option:

Gives the right, but not the obligation, to buy an asset at a given exercise price K.

- European call
  Allows exercise ONLY at the expiration date. That is, at expiration T pays:
  $$\max(S_T - K, 0)$$
- American call
  Can be exercised at any time $t \leq T$, paying:
  $$\max(S_t - K, 0)$$
- Basket call option
  Its payoff is given by:
  $$\max(\text{Average over } m \text{ stocks} - K, 0)$$
  can be either American or European, and average can be arithmetic (weighted or unweighted) or geometric.

Valuation

- GBM model for the price of a stock with dividends:
  $$dS_t = (r - \delta)S_t dt + \sigma S_t dB_t^Q$$
  that is:
  $\log(S_t) - \log(S_0) \sim N\left((r - \delta - \frac{\sigma^2}{2})t, \sigma^2 t\right)$
  where
  $r =$ risk free rate, $\delta =$ dividend rate, $\sigma =$ volatility
- European Call price:
  $$\mathrm{E}^Q\left(e^{-rT}\max(S_T - K, 0)\right)$$
- American Call price:
  $$\mathrm{E}^Q\left(e^{-r\tau}\max(S_\tau - K, 0)\right)$$
  where τ is the “optimal” exercise (stopping) time.

Option Pricing Algorithms

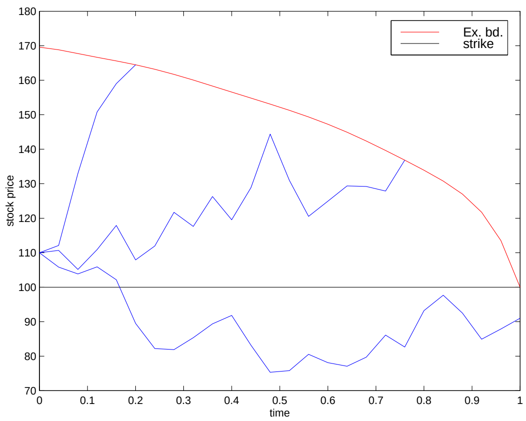
Table 1: Option pricing algorithms

Algorithms	European	American
One stock	Black Scholes	Binomial Trees Finite Difference
Basket	Monte Carlo	?

MC option pricing

- Assume the optimal exercise boundary is known. Then the price of the option can be estimated through:

$$\frac{1}{n}\sum_{i=1}^{n} e^{-r\tau_i}\max(S_{\tau_i}-K,0)$$



- However, the optimal exercise boundary is not only not known, but also it is not easy to estimate.

Characterization of the exercise boundary

At a fixed time t, define the value of continuation y at the current state of the underlying assets x as:

$y =$ (Avg.) discounted payoff - value of exercise (of the sampling path(s) starting from x.)

The exercise boundary is given by the set of points x such that $\mathrm{E}(y|x) = 0$.

Consequently, the exercise boundary is characterized by a function $F(x)$ such that:

- $F(x) > 0$ whenever $\mathrm{E}(y|x) > 0$ (hold), and
- $F(x) < 0$ whenever $\mathrm{E}(y|x) < 0$ (exercise).

Characterization of exercise boundary by Minimizing $J(F) = \mathrm{E}(e^{-yF(\mathbf{x})})$

At a fixed x, let $F^*(x) = \mathrm{argmin}_F \mathrm{E}(e^{-yF}|x)$ then:

$$e^{-F^*(x)\mathrm{E}(y|x)} \leq \mathrm{E}(e^{-yF^*(x)}|x) \leq 1.$$

Thus, $F^*(x)\mathrm{E}(y|x) \geq 0$.

Some properties of F^*:

- if $|F^*\mathrm{E}(y|x)| > 0$ then F^* gets the right sign.
- Thus F^* provides a characterization of the exercise boundary.
- This characterization is invariant under scaling of the $y's$.

Estimating $F^*(x) = \mathrm{argmin}_F \mathrm{E}(e^{-yF}|x)$ (Boosting Methodology)

1. Given a training sample $\{(x_i, y_i)\}_{i=1}^{N},\ y_i \in \Re$ start with weights $w_i = 1/N, i = 1, ..., N, F(x) = 0.$
2. Repeat for $m = 1, 2, ..., M$:
   (a) Fit the regression function $f_m(x) = \mathrm{E}_w(\frac{1}{y}|x)$ by weighted least-squares of $\frac{1}{y_i}$ to x_i with weights w_i.
   (b) Update $F(x) \leftarrow F(x) + f_m(x)$.
   (c) Recompute weights $w_i = y_i^2 e^{-y_i F(x_i)}$ and renormalize so that $\sum_{i=1}^{N} w_i = 1$.
3. Output the classifier $\mathrm{sign}[F(x)] = \mathrm{sign}[\sum_{m=1}^{M} f_m(x)]$.

Classification Monte Carlo (CMC) algorithm

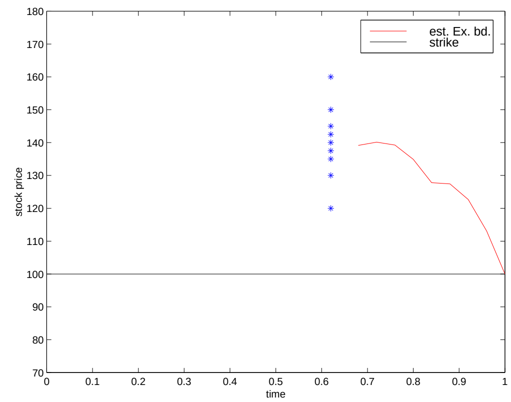
The CMC algorithm has two steps: A Classification step (**C** step), and a Monte Carlo step (**MC** step).

- **C** Step: Characterize Optimal Exercise Boundary
  - Discretize the problem, considering that there is a finite number m of exercising opportunities.
  - Following the Dynamic Programming approach, assume the optimal exercise boundary is known for the last $m - k + 1$ dates. Then at the k-th date:
    - ∗ Take some state points and simulate some paths starting from them.

- ∗ For each state point, define the Value of Continuation (y) as:

  y = Avg. disc. payoff - Value of Exercise
- ∗ Estimate the function $F^*(x)$ that characterizes the exercise boundary.

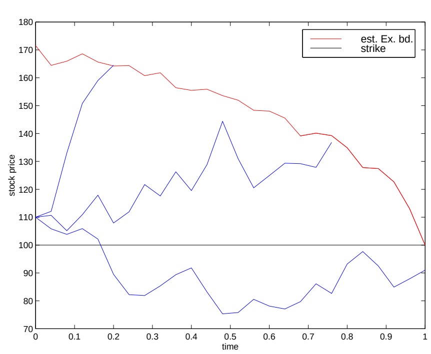
C Step



CMC algorithm

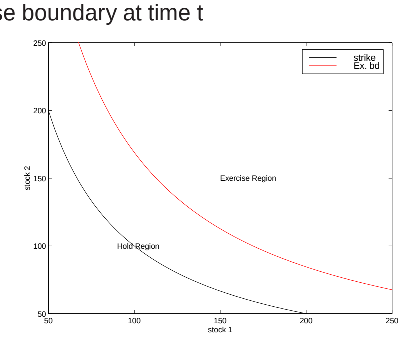
- **MC** Step: Simulate paths and price the option by Monte Carlo

MC Step

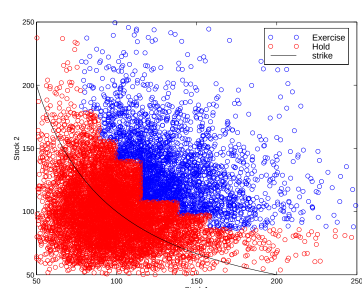


CMC algorithm Geometric average of two stocks

- Exercise boundary at time t



- Estimated Exercise boundary at time t



Some Numerical Results

Parameters:

$r = 3\%$, $\delta = 5\%$, $\sigma = 40\%$, $K = 100$ and $T = 1$ year.

Exercise opportunities: 10

Number of points per iteration: 3000

Branches per point: 300

Method: Boosted stumps with 150 iterations.

Monte Carlo scenarios: 100,000

Table 2: Call option on geometric average of 5 assets

S0	Euro	Ame	Est	Diff	% error
110	7.521	10.211	10.190	-0.021	0.2%
100	3.445	4.291	4.268	-0.023	0.5%
90	1.172	1.362	1.351	-0.011	0.8%

Table 3: Call option on geometric average of 7 assets

S0	Euro	Ame	Est	Diff	% error
110	6.201	10.000	10.000	0.000	0.0%
100	2.419	3.270	3.245	-0.025	0.8%
90	0.628	0.761	0.748	-0.013	1.7%

Note: The "American price" corresponds to a single asset binomial tree.
The "Estimated price" is the average over 20 independent scenarios.

Comments on CMC algorithm

- The algorithm grows linearly with the number of underlying stocks.
- An **advantage** of the CMC algorithm is that by estimating the optimal exercise boundary, the price of the option through its entire life can be recalculated using only the MC step. Also, the exercise boundary is independent of the starting stock('s) price(s).

Conclusions and Further Directions

- At a fixed time t, we are interested in finding the regions of the underlying assets space $\mathbf{X_t}$ such that $\mathrm{E}(y|x_t) > 0$ (and the regions where $\mathrm{E}(y|x_t) < 0$), where y is the value of continuation.
- The solution F^* that minimizes the loss function $J(F) = E(e^{-yF(x)})$, provides a good characterization of the exercise boundary.
- Boosting methodology seems to be giving promising results for this problem.
- Other methods for minimizing such loss function should be investigated.
- Importance sampling could play a very important role in improving the characterization of the exercise boundary, since within a computer simulation environment one is able to gather more data.

References

[1] Breiman L., Friedman J., Olshen, J.,Stone C. (1983) Classification and Regression Trees, *Wadsworth*

[2] Broadie M., Glasserman P. (1997) Pricing American-Style Securities Using Simulation, *Journal of Economic Dynamics and Control* **21**, 1323-1352.

[3] Broadie M., Glasserman P. (1997) A Stochastic Mesh Method for Pricing High-Dimensional American Options, Working Paper, Columbia University.

[4] Cohn D., Atlas L., Ladner R. (1994) Improving Generalization with Active Learning, *Machine Learning* **15**, 201-221.

[5] Freund, Y. Schapire, R. (1996) Experiments with a New Boosting Algorithm, *Machine Learning: Proceedings of the Thirteenth International Conference*, 148-156.

[6] Friedman J., Hastie T., Tibshirani R. (1998), Additive Logistic Regression: A Statistical View of Boosting, Stanford University, Dept. of Statistics, Technical Report.

[7] Fu M., Laprise S., Madan D., Su Y., Wu R. (1999) Pricing American Options: A Comparison of Monte Carlo Simulation Approaches, Working Paper, University of Maryland at College Park

[8] Longstaff F., Schwartz E. (1998) Valuing American Options By Simulation: A Simple Least-Squares Approach, Working Paper, University of California, Los Angeles