Monte Carlo Methods for American Options

Luca Capriotti

Lecture for the Master of Finance
Baruch College, New York, April 26, 2013

European vs. American/Bermudan Options

- Contrary to European options, American and Bermudan options can be exercised on multiple days up to trade expiry.
- Bermudan options can be exercised on a discrete set of dates while American options can be exercised in continuous time intervals.
- More choices, more value:

$$\text{European} \leq \text{Bermudan} \leq \text{American}$$

- In the following we will mainly restrict our discussion to the case of Bermudan options as it is the one of more relevance in practice.
- The case of American options can be obtained as a limiting case when the number of exercise dates per unit time tends to infinity.

Exercise Policies and Stopping Times

- We indicate with $T_1, \ldots, T_B$, the exercise dates of the option and with $\mathcal{D}(t)$ the deterministic set of exercise dates T_i larger or equal to time t, namely $\mathcal{D}(t) = \{T_i \geq t\}$.
- An *exercise policy* is represented mathematically by a *stopping time* taking values in $\mathcal{D}(t)$. Recall that a random variable τ is a stopping time if the event $\{\tau \leq t\}$ can be determined using only the information available up to time t.
- We indicate with $\mathcal{T}(t)$ the set of stopping times taking values in $\mathcal{D}(t)$.
- We assume $t < T_{B-1}$ as in the last period the Bermudan option becomes European.

Optimal Exercise

- A rational investor will exercise the option that she holds in such a way to maximize its economic value.
- As a result, the value of a Bermudan option is the supremum of the option value over all the possible exercise policies, namely

$$\frac{V(t)}{N(t)} = \sup_{\tau \in \mathcal{T}(t)} \mathbb{E}_t \Big[\frac{E(\tau)}{N(\tau)} \Big]$$

  where $E(t)$ is the exercise value of the option, and $N(t)$ is the chosen *numeraire*.
- In this equation $V(t)$ is the value of the option with early exercise *conditional on exercise not taking place strictly before time t*.

Continuation or Hold Value

- Indicate with $\eta(t)$ the smallest integer such that $T_{\eta(t)+1} > t$.
- The *hold value* $H_\eta(t)$ ($T_{\eta(t)} \leq t < T_{\eta(t)+1}$) is the value of the Bermudan option when the exercise dates are restricted to $\mathcal{D}(T_{\eta(t)+1})$

$$\frac{H_\eta(t)}{N(t)} = \mathbb{E}_t\left[\frac{V(T_{\eta+1})}{N(T_{\eta+1})}\right]$$

- Clearly $H_\eta(t) = V(t)$ for $T_\eta < t < T_{\eta+1}$, since there are no exercise opportunities in this interval.
- Instead $H_\eta(T_\eta)$ can be interpreted as the hold value of the Bermudan option at time T_η, i.e., the value of the option if we decide *not* to exercise at time T_η.

Optimal Exercise and Dynamic Programming

- The option holder following an optimal exercise policy will exercise her option if the exercise value is larger than the hold value

$$V(T_\eta) = \max(E(T_\eta), H_\eta(T_\eta))$$

- This, when combined with the definition of hold value, leads to the so-called *dynamic programming* or Bellman principle formulation, namely,

$$\frac{H_\eta(t)}{N(t)} = \mathbb{E}_t\Big[\max\Big(\frac{E(T_{\eta+1})}{N(T_{\eta+1})}, \frac{H_{\eta+1}(T_{\eta+1})}{N(T_{\eta+1})}\Big)\Big]$$

for $T_\eta \leq t < T_{\eta+1}$, and $\eta = 1, \ldots, B-1$.

Optimal Exercise Time

- Starting from the terminal condition $H_B(T_B) \equiv 0$, this defines a backward iteration in time for $H_\eta(T_\eta)$.
- By definition, this is also equal to $V(t)$ if t is not an exercise date.
- Conversely, if t is an exercise date, $t = T_\eta$, then

$$V(T_\eta) = \max(E(T_\eta), H_\eta(T_\eta))$$

- The dynamic programming formulation above implies that the stopping time defining optional exercise (as seen as time t) is given by

$$\tau^\star = \inf[T_i \geq t : E(T_i) \geq H_i(T_i)]$$

Example: American put option

- Consider an American put option struck at K on a stock $S(t)$:

$$\frac{dS(t)}{S(t)} = rdt + \sigma dW_t$$

  where r is the (constant) instantaneous risk free rate of interest, σ is the volatility and W_t is a standard Brownian motion.
- The value of the Bermudan put option can be expressed as

$$V(t) = \sup_{\tau \in \mathcal{T}(t)} \mathbb{E}_t \left[e^{-r(\tau - t)} (K - S(\tau))^+ \right]$$

Example: American put option (cont'd)

- The optimal exercise time is given by

$$\tau^\star = \inf[t \in \mathcal{D}(t) : (K - S(t))^+ \geq H(t)]$$

- Since $H(t)$ is function itself of $S(t)$ the latter condition can be expressed equivalently as

$$\tau^\star = \inf[t \in \mathcal{D}(t) : S(t) < b^\star(t)]$$

for a deterministic function $b^\star(t)$, assuming the natural interpretation of *exercise boundary*.

Example: American put option (cont'd)

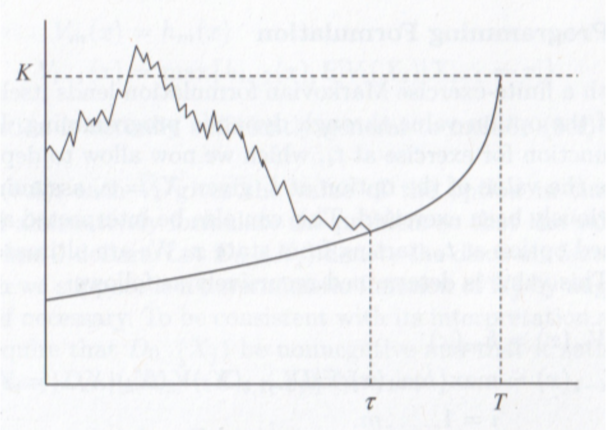

Figure: Exercise Boundary for the American Put option on a single stock (taken from Ref. [1]).

Monte Carlo challenges

- The dynamic programming recursion can be easily implemented in the context of deterministic numerical methods (multinomial trees or PDEs), which are based on *backward induction*.
- These are limited by the *curse of dimensionality*.
- On the other hand, in the context of Monte Carlo methods, the paths describing the time evolution of the underlying risk factors are generated *forward in time*, thus making the direct application of backward induction impossible.
- This makes pricing Bermudan options, whose dimensionality is too high to be treated by deterministic numerical methods, very challenging.

Variational Principle and Lower Bound Methods

- An immediate consequence of

$$\frac{V(t)}{N(t)} = \sup_{\tau \in \mathcal{T}(t)} \mathbb{E}_t\left[\frac{E(\tau)}{N(\tau)}\right]$$

is that for any stopping time $\tau \in \mathcal{T}(t)$

$$\frac{V(t)}{N(t)} \geq \mathbb{E}_t\left[\frac{E(\tau)}{N(\tau)}\right]$$

- As a result, a *lower bound* for the value of the Bermudan option $V(t)$ can be computed by means of Monte Carlo through any exogenous guess for the optimal exercise strategy $\tau^{\star}$.

Parametric Exercise Boundary Methods

- Parametric lower bound methods involve a user supplied specification of a parametric stopping rule $\tau_\theta \in \mathcal{T}_\theta(t)$ where $\theta \in \Theta \subset \mathbb{R}^{N_\theta}$ is a N_θ dimensional parameter vector.
- Due to the variational principle, for any value of θ the stopping rule generates a lower bound, $V_\theta(t)$, of the true value of the Bermudan option, $V(t)$.
- Clearly, the best approximation of such value within the chosen class, is the one for which the value $V_\theta(t)$ is the largest, namely

$$V_{\theta^\star}(t) = \sup_\theta V_\theta(t)$$

Algorithm

Step 1 Generate n independent paths of the underlying Markov process $X^{(k)} = (X^{(k)}(T_{\eta(t)+1}), \ldots, X^{(k)}(T_B))$, for $k = 1, \ldots, n$. For path k, let $\tau^{(k)}(\theta)$ be the exercise time suggested by the stopping rule for the given value of the parameter θ.

Step 2 For each path k, set:

$$E_\theta^{(k)} = E(X(\tau_\theta^{(k)}))$$

$$V_\theta^{(k)} = \frac{E_\theta^{(k)}}{N(X(\tau_\theta^{(k)}))}$$

Algorithm (cont'd)

Step 3 Return:

$$\bar{V}_\theta(t) = N(t) \times \frac{1}{n} \sum_{k=1}^{n} V_\theta^{(k)}$$

Step 4 Find:

$$\theta^\star = \arg \sup_{\theta \in \Theta} \bar{V}_\theta(t)$$

Step 5 Return:

$$\tilde{V}_{\theta^\star}(t) = \sup_{\theta \in \Theta} \bar{V}_\theta(t)$$

Lower Bound?

- The value $\tilde{V}_{\theta^\star}(t)$ is an estimate of the true value of the Bermudan option $V(t)$. *However we do not know whether it is a lower bound or an upper bound.*
- This is because while $V_{\theta^\star}(t) \leq V(t)$, because of the variational principle, the *estimator* $\tilde{V}_{\theta^\star}(t)$ is instead *biased high* with respect to $V_{\theta^\star}(t)$.
- Indeed, as a result of Jensen's inequality:

$$\mathbb{E}[\tilde{V}_{\theta^\star}(t)] = \mathbb{E}[\sup_{\theta\in\Theta} \bar{V}_\theta(t)] \geq \sup_{\theta\in\Theta} \mathbb{E}[\bar{V}_\theta(t)] = \sup_{\theta\in\Theta} V_\theta(t) = V_{\theta^\star}(t)$$

so that:

$$\mathbb{E}[\tilde{V}_{\theta^\star}(t)] \geq V_{\theta^\star}(t) \leq V(t)$$

Algorithm: Lower Bound!

- In order to avoid this problem on can replace Step 5 with the following two steps:

  Step 5′ Draw N *independent* paths for $X^{(k)}$ and compute for each path:

$$E_{\theta^\star}^{(k)} = E(X(\tau_{\theta^\star}^{(k)}))$$

$$V_{\theta^\star}^{(k)} = \frac{E_{\theta^\star}^{(k)}}{N(X(\tau_{\theta}^{(k)}))}$$

  Step 6 Return:

$$\bar{V}_{\theta^\star}(t) = N(t) \times \frac{1}{N} \sum_{k=1}^{N} V_{\theta^\star}^{(k)}$$

- Since $\mathbb{E}[\bar{V}_{\theta^\star}(t)] = V_{\theta^\star}(t) \leq V(t)$, this provides a genuine lower bound estimator.

A Useful Parameterization of the Exercise Boundary

- The optimization in Step 4 above can be simplified whenever the parameter θ decomposes into $B-1$ sub-components, with the i-th subset parametrizing the exercise decision at time T_i, i.e. $\theta = (\theta_{\eta(t)+1}, \ldots, \theta_{B-1})$ with each θ_i possibly a vector itself.
- For instance, this is the case for the so-called *moneyness stopping rule*

$$\tau(\theta) = \inf_{i>\eta(t)} [T_i : E(X(T_i)) > \theta_i]$$

prescribing early exercise whenever the option is in the money 'deeply enough'.

A Useful Parameterization of the Exercise Boundary (cont'd)

- This can be seen as a simplification of the optimal stopping rule

$$\tau^\star = \inf[T_i \geq t : E(T_i) \geq H_i(T_i)]$$

in which the hold value, $H_i(X(T_i))$, is replaced by the constant θ_i.

- In this case, Step 4 above is replaced by the following step:

Step 4′ Proceeding backwards in time, for $i = B - 1, \ldots, \eta(t) + 1$ find $\tilde{\theta}_i$ by keeping $(\tilde{\theta}_{i+1}, \ldots, \tilde{\theta}_{B-1})$ fixed and maximizing:

$$\tilde{V}(t) = N(t) \times \sum_{k=1}^{n} \frac{E(X(\tilde{\tau}_i))}{N(\tilde{\tau}_i)} \tag{1}$$

where $\tilde{\tau}_i = \tilde{\tau}(\theta_i, \tilde{\theta}_{i+1} \ldots, \tilde{\theta}_{B-1}) \in \mathcal{D}(T_i)$.

A Useful Parameterization of the Exercise Boundary (cont'd)

Note that:

- Step $4'$ does not involve repeating the Monte Carlo simulation from scratch. Rather the same set of n paths can be reused.
- With a finite number of paths the one in Step $4'$ is a non-smooth optimization problem and is best solved by an iterative search rather than a derivatives based approach.
- There is no guarantee the algorithm in Step $4'$ produces the optimum $\theta^\star = \arg\sup_{\theta\in\Theta} \bar{V}_\theta(t)$. This is because each $\tilde{\theta}_i$ is optimized assuming exercise only after T_i. In reality, only a subset of the paths would arrive to T_i without triggering early exercise beforehand.

Exercise 1

Consider a Bermudan option on the maximum of two assets following a geometric Brownian motion of the form

$$\frac{dS_i(t)}{S_i(t)} = (r - \delta)dt + \sigma dW_t^i$$

where $r = 5\%$, $\delta = 10\%$, $\sigma = 20\%$ and $S_1(0) = S_2(0)$. The two assets are assumed independent. The (undiscounted) payoff upon exercise at time T_i is

$$(\max(S_1(t), S_2(t)) - K)^+$$

where $K = 100$ is the strike price. The maturity of the option is $T_B = 3$ and can be exercised at nine equally spaced dates $T_i = i/3$ with $i = 1, \ldots, 9$. The exact option prices obtained are 13.90, 8.08 and 21.34 for $S_i(0) = 100, 110$ and 90, respectively. Use the parametric lower bound method and the moneyness stopping rule to estimate the value of the option (use both a single value of θ and a different value of θ for each exercise date). Compare the results obtained with Steps 1 to 5 and Steps 1 to 6 (through Step 5′) of the algorithm above.

Regression Approaches

- The optimal exercise strategy defined by

$$\tau^\star = \inf[T_i \geq t : E(T_i) \geq H_i(T_i)]$$

  can be approximated by constructing an estimate of the hold value $H_i(T_i)$, $i = \eta(t) + 1, \ldots, B - 1$.
- In general, in a Markov setting the hold value is a function of the state vector at time T_i

$$H_i(x) = N(x) \times \mathbb{E}\Big[\frac{V(X(T_{i+1}))}{N(X(T_{i+1}))}\Big|X(T_i) = x\Big]$$

Regression Approaches (cont'd)

- Regression based approaches are based on the following *ansatz* for the hold value

$$\hat{H}_i(x) = \sum_{j=1}^{d} \beta_{ij}\psi_j(x)$$

for a set of d *basis functions* $\psi_j(x)$ and coefficients β_{ij}. Equivalently, this can be written in matrix form as

$$\hat{H}_i(x) = \beta_i^T \psi(x) = \psi^T(x)\beta_i$$

where $\beta_i = (\beta_1, \ldots, \beta_d)^T$ and $\psi(x) = (\psi_1(x), \ldots, \psi_d(x))^T$.

Regression Approaches (cont'd)

- Multiplying by $\psi^T(X(T_i))$, using the definition of hold value and taking unconditional expectations one gets

$$\mathbb{E}\Big[\beta_i^T \psi(X(T_i))\psi^T(X(T_i))\Big] = \mathbb{E}\Big[\frac{N(X(T_i))V(X(T_{i+1}))}{N(X(T_{i+1}))}\psi^T(X(T_i))\Big]$$

which gives

$$\beta_i = \Psi_i^{-1}\Omega_i$$

where we define the $d \times d$ matrix

$$\Psi_i = \mathbb{E}\Big[\psi(X(T_i))\psi^T(X(T_i))\Big]$$

and the $d \times 1$ vector

$$\Omega_i = \mathbb{E}\Big[\frac{N(X(T_i))V(X(T_{i+1}))}{N(X(T_{i+1}))}\psi(X(T_i))\Big]$$

Regression Approaches (cont'd)

- These equations provide a straightforward recipe to compute the regression coefficients β_i by substituting Ψ and Ω with their sample average over n Monte Carlo replications, $\bar{\Psi}_i$ and $\bar{\Omega}_i$.
- More explicitly, considering a set of Monte Carlo paths of the Markov state variable X sampled on the Bermudan exercise dates

$$\left(X^{(k)}(T_{\eta(t)+1}), \ldots, X^{(k)}(T_{B-1})\right)$$

for $k = 1, \ldots, n$ one could compute the sample averages

$$\bar{\Psi}_i = \frac{1}{n}\sum_{k=1}^{n} \psi(X^{(k)}(T_i))\psi^T(X^{(k)}(T_i))$$

$$\bar{\Omega}_i = \frac{1}{n}\sum_{k=1}^{n} \frac{N(X^{(k)}(T_i))\hat{V}(X^{(k)}(T_{i+1}))}{N(X^{(k)}(T_{i+1}))}\psi(X^{(k)}(T_i))$$

Regression Approaches (cont'd)

- In the last equation $\hat{V}$ is given by

$$\hat{V}(X^{(k)}(T_i)) = \max\left(E(X^{(k)}(T_i), \hat{H}_i(X^{(k)}(T_i))\right)$$

for $i = \eta(t) + 1, \ldots B - 1$, where we have replaced the true value H with the estimate $\hat{H}$ according to

$$\hat{H}_i(X^{(k)}(T_i)) = \sum_{j=1}^{d} \beta_{ij} \psi_j(X^{(k)}(T_i))$$

- However this depends on the *yet to be determined coefficients* β_i.

Backward Induction

- As a result, the calculation of the sample averages $\bar{\Psi}_i$ and $\bar{\Omega}_i$ needs to be performed backwards. Indeed, starting from the penultimate Bermudan exercise date T_{B-1} on which

  $$\hat{V}(X^{(k)}(T_B)) = \max\left(E(X^{(k)}(T_B)), 0\right)$$

  one can compute

  $$\bar{\beta}_{B-1} = \bar{\Psi}_{B-1}^{-1}\bar{\Omega}_{B-1}$$

  which allows one to compute for $i = B - 2$ the estimate of the hold value

  $$\hat{H}_{B-1}(X^{(k)}(T_{B-1})) = \sum_{j=1}^{d} \bar{\beta}_{B-1j}\psi_j(X^{(k)}(T_{B-1}))$$

  required to compute the estimate $\hat{V}(X^{(k)}(T_{B-1}))$. This can be iterated until we get to $i = \eta(t) + 1$...

Algorithm

Step 1 Simulate n independent paths ($k = 1, \ldots, n$)

$$\left(X^{(k)}(T_{\eta(t)+1}), \ldots, X^{(k)}(T_{B-1})\right)$$

Step 2 At final expiry compute the value:

$$\hat{V}(X^{(k)}(T_B)) = \max(E(X^{(k)}(T_B)), 0)$$

Step 3 For $i = B-1, \ldots, \eta(t)+1$

a) Compute:

$$\bar{\Psi}_i = \frac{1}{n}\sum_{k=1}^{n} \psi(X^{(k)}(T_i))\psi^T(X^{(k)}(T_i))$$

b) Compute:

$$\bar{\Omega}_i = \frac{1}{n}\sum_{k=1}^{n} \frac{N(X^{(k)}(T_i))\hat{V}(X^{(k)}(T_{i+1}))}{N(X^{(k)}(T_{i+1}))}\psi(X^{(k)}(T_i))$$

using the value of $\hat{V}(X^{(k)}(T_{i+1}))$ computed in the previous time step.

Algorithm (cont'd)

c) Compute by matrix inversion and multiplication:

$$\bar{\beta}_i = \bar{\Psi}_i^{-1} \bar{\Omega}_i$$

d) Set for the estimate of the hold value at time T_i:

$$\hat{H}_i(X^{(k)}(T_i)) = \bar{\beta}_i \psi_i(X^{(k)}(T_i))$$

e) Set for the estimate of the Bermudan option value at time T_i:

$$\hat{V}(X^{(k)}(T_i)) = \max(E(X^{(k)}(T_i)), \hat{H}_i(X^{(k)}(T_i)))$$

Step 4 Compute:

$$\bar{V}(t) = N(t) \times \frac{1}{n} \sum_{i=1}^{n} \frac{\hat{V}(X^{(k)}(T_{\eta(t)+1}))}{N(X^{(k)}(T_{\eta(t)+1}))}$$

Longstaff-Schwartz Method

- A modification of this algorithm was proposed by Longstaff and Schwartz and entails replacing

$$\hat{V}(X^{(k)}(T_i)) = \max(E(X^{(k)}(T_i)), \hat{H}_i(X^{(k)}(T_i)))$$

in Step 3e, with

$$\hat{V}(X^{(k)}(T_i)) = \begin{cases} E(X^{(k)}(T_i)) & \text{if } \; E(X^{(k)}(T_i)) > \hat{H}_i(X^{(k)}(T_i))) \\ \hat{V}(X^{(k)}(T_{i+1}))N(X^{(k)}(T_i))/N(X^{(k)}(T_{i+1})) & \text{otw} \end{cases}$$

which in the examples considered was shown to lead to more accurate results.

- Similarly to the parametric exercise boundary methods, regression based approaches produce lower bound estimate of the true Bermudan option value if this is computed by means of a second simulation in which the continuation value estimated in the first simulation is used to determine early exercise.

Exercise 2

Consider the Bermudan option of Exercise 1. Compare the results obtained with the regression based approach (including the modification of Longstaff and Schwartz). Use basis functions of the form $S_1^{\alpha} S_2^{\beta}$. Include results obtained directly from the backward induction Steps 1-4 in this Section and those obtained by means of a second independent simulation using the hold value estimated by means of the backward induction in the first simulation.

Super-martingale Property

- The pricing problem for American or Bermudan option admit a ‘dual’ formulation in which it can be expressed as a *minimization problem*.
- The dynamic programming equations

$$\frac{H_\eta(t)}{N(t)} = \mathbb{E}_t\left[\frac{V(T_{\eta+1})}{N(T_{\eta+1})}\right]$$
$$V(T_\eta) = \max(E(T_\eta), H_\eta(T_\eta))$$

imply

$$\frac{V(X(T_i))}{N(X(T_i))} \geq \mathbb{E}\left[\frac{V(X(T_{i+1}))}{N(X(T_{i+1}))}\middle| X(T_i)\right]$$

Super-martingale Property

- Hence the discounted value of a Bermudan option, $V(X(T_i))/N(X(T_i))$, is a *super-martingale*.
- In addition one also has:

$$\frac{V(X(T_i))}{N(X(T_i))} \geq \frac{E(X(T_i))}{N(X(T_i))}$$

- In fact, the discounted value of a Bermudan option is the *minimal* super-martingale dominating the discounted exercise value $E(X(T_i))/N(X(T_i))$.

Duality and Upper Bounds

- Consider a martingale M with $M(t) = \mathbb{E}_t[M(s)] = 0$.

$$\frac{V(t)}{N(t)} = \sup_{\tau \in \mathcal{T}(t)} \mathbb{E}_t\Big[\frac{E(\tau)}{N(\tau)}\Big]$$

$$= \sup_{\tau \in \mathcal{T}(t)} \mathbb{E}_t\left[\frac{E(\tau)}{N(\tau)} + M(\tau) - M(\tau)\right] = \sup_{\tau \in \mathcal{T}(t)} \mathbb{E}_t\Big[\frac{E(\tau)}{N(\tau)} - M(\tau)\Big]$$

- Hence, using Jensen's inequality:

$$\frac{V(t)}{N(t)} = \sup_{\tau \in \mathcal{T}(t)} \mathbb{E}_t\Big[\frac{E(\tau)}{N(\tau)} - M(\tau)\Big] \leq \mathbb{E}_t\Big[\max_{\tau \in \mathcal{D}(t)} \left(\frac{E(\tau)}{N(\tau)} - M(\tau)\right)\Big]$$

As a result an *upper bound* on the value of the Bermudan option is

$$\frac{V(t)}{N(t)} \leq \inf_{M(t)=0} \mathbb{E}_t\left[\max_{\tau \in \mathcal{D}(t)} \left(\frac{E(\tau)}{N(\tau)} - M(\tau)\right)\right]$$

Duality and Upper Bounds (cont'd)

- Interestingly, we can show that

$$\frac{V(t)}{N(t)} \leq \inf_{M(t)=0} \mathbb{E}_t \left[\max_{\tau \in \mathcal{D}(t)} \left(\frac{E(\tau)}{N(\tau)} - M(\tau) \right) \right]$$

holds with equality i.e., it is possible to find $M^\star$ such that

$$\frac{V(t)}{N(t)} = \mathbb{E}_t \left[\max_{\tau \in \mathcal{D}(t)} \left(\frac{E(\tau)}{N(\tau)} - M^\star(\tau) \right) \right]$$

Duality and Upper Bounds (cont'd)

- This is a consequence of the so called *Doob-Mayer decomposition* for super-martingales:

$$\frac{V(s)}{N(s)} = \frac{V(t)}{N(t)} + M^\star(s) - A(s)$$

where $A(s)$ is a non decreasing predictable process with $A(t) = 0$. Indeed, choosing for M the martingale component above

$$\frac{V(t)}{N(t)} \leq \mathbb{E}_t \left[\max_{\tau \in \mathcal{D}(t)} \left(\frac{E(\tau)}{N(\tau)} - M^\star(\tau) \right) \right]$$
$$= \frac{V(t)}{N(t)} + \mathbb{E}_t \left[\max_{\tau \in \mathcal{D}(t)} \left(\frac{E(\tau)}{N(\tau)} - \frac{V(\tau)}{N(\tau)} - A(\tau) \right) \right] \leq \frac{V(t)}{N(t)}$$

Duality and Upper Bounds (cont'd)

- As a result the Bermudan option value can be obtained by finding the martingale component of the deflated option price.
- Of course this is in general as difficult as the original problem. Nonetheless any approximation of such martingale component will provide an upper bound of the option price through

$$\frac{V(t)}{N(t)} \leq \inf_{M(t)=0} \mathbb{E}_t \left[\max_{\tau \in \mathcal{D}(t)} \left(\frac{E(\tau)}{N(\tau)} - M(\tau) \right) \right]$$

- Conceivably the better the approximation the tighter the upper bound.

Martingales from Approximate Option Values

- Given an approximation of the deflated option price $V(X_s)/N(X_s)$ one can extract the corresponding martingale component to construct an approximation of $M^\star(s)$
- One natural choice is

$$\hat{M}^\star(t) = \sum_{j=\eta(t)+1}^{B} \Delta(T_j)$$

with $M(t) = 0$, and

$$\Delta(T_j) = \frac{V(X(T_j))}{N(X(T_j))} - \mathbb{E}\Big[\frac{V(X(T_j))}{N(X(T_j))}\Big| X(T_{j-1})\Big]$$

for $j > \eta(t) + 1$ and

$$\Delta(T_{\eta(t)+1}) = \frac{V(X(T_{\eta(t)+1})))}{N(X(T_{\eta(t)+1})))} - \mathbb{E}\Big[\frac{V(X(T_{\eta(t)+1})))}{N(X(T_{\eta(t)+1})))}\Big| X(t)\Big]$$

Martingales from Approximate Option Values (cont'd)

- By applying the 'Tower law' of conditional expectations it is immediate to see that $\hat{M}^\star$ is indeed a martingale.
- The same is true if we replace V with the estimators $\hat{V}$defined e.g., by a regression based approach.
- However, replacing the true deflated hold value $\mathbb{E}\Big[\frac{V(X(T_j))}{N(X(T_j))}\Big|X(T_{j-1})\Big]$ with an approximate one $\hat{H}/N$ would not guarantee the martingale property for $\Delta(T_j)$.
- As a result, the deflated hold value needs to be valued by means of a single time-step nested simulation of N' paths spun out of $X(T_{j-1})$ namely

$$\mathbb{E}\Big[\frac{V(X(T_j))}{N(X(T_j))}\Big|X(T_{j-1})\Big] \simeq \frac{1}{N'}\sum_{m=1}^{N'}\frac{\hat{V}(X^{(m)}(T_j))}{N(X^{(m)}(T_j))}$$

Martingales from Approximate Stopping Times

- Alternatively one can employ approximations of the optimal stopping time.
- Denote with τ_j the stopping time as seen at time T_j, i.e. $\tau_j \in \mathcal{D}(T_j)$, and suppose that these stopping times are defined by approximate hold value functions $\hat{H}$, namely,

$$\tau_j = \min\left(k = j, \ldots, B : E(X(T_k)) \geq \hat{H}(X(T_k))\right)$$

  where $\hat{H}$ could be estimated for instance by means of regression.
- One can construct $\Delta(T_j)$ as

$$\Delta(T_j) = \mathbb{E}\Big[\frac{E(X(\tau_j))}{N(X(\tau_j))}\Big| X(T_j)\Big] - \mathbb{E}\Big[\frac{E(X(\tau_j))}{N(X(\tau_j))}\Big| X(T_{j-1})\Big]$$

Martingales from Approximate Stopping Times (cont'd)

- Also note that

$$\mathbb{E}\Big[\frac{E(X(\tau_j))}{N(X(\tau_j))}\Big|X(T_j)\Big]$$
$$= \begin{cases} E(X(T_j))/N(X(T_j)) & \text{if } \; E(X(T_j)) \geq \hat{H}(X(T_j)) \;, \\ \mathbb{E}\Big[E(X(\tau_{j+1}))/N(X(\tau_{j+1}))\Big|X(T_j)\Big] & \text{otherwise.} \end{cases}$$

- As a result the only quantities that need to be valued are those of the form

$$\mathbb{E}\Big[\frac{E(X(\tau_{j+1}))}{N(X(\tau_{j+1}))}\Big|X(T_j)\Big]$$

which can be computed by sub-simulation.

Upper Bond Algorithm

Step 1 Generate n independent paths of the underlying Markov process $X^{(k)} = (X^{(k)}(T_{\eta(t)+1}), \ldots, X^{(k)}(T_B))$, for $k = 1, \ldots, n$.

Step 2 For each path k, at each $X^{(k)}(T_j)$, $j = \eta(t) + 1, \ldots, B$

a) Evaluate:

$$
\begin{aligned}
E_j^{(k)} &= E(X^{(k)}(T_j)) \\
\hat{H}_j^{(k)} &= \hat{H}(X^{(k)}(T_j)) \\
N_j^{(k)} &= N(X^{(k)}(T_j))
\end{aligned}
$$

b) Simulate N' subpaths starting from $X^{(k)}(T_j)$ and compute

$$
\bar{E}_j^{(k)} = \frac{1}{N'} \sum_{m=1}^{N'} E(X^{(k,m)}(\tau_{j+1})) \ .
$$

c) Compute $\hat{\Delta}_j^{(k)}$ using the information above.

Upper Bond Algorithm

Step 3 For each path k, compute: $M^{(k)}(T_i)$ for $i = \eta(t) + 1, \ldots, B$.

Step 4 For each path k, evaluate: $E^{(k)}(T_i) - M^{(k)}(T_i)$ for $i = \eta(t) + 1, \ldots, B$.

Step 5 For each path k, evaluate:

$$U^{(k)} = \min_{i=\eta(t)+1,\ldots,B} \left(\frac{E^{(k)}(T_i)}{N(X(T_i))} - M^{(k)}(T_i) \right) .$$

Step 6 Return:

$$\bar{V}(t) = N(t) \times \frac{1}{N} \sum_{k=1}^{N} U^{(k)} .$$

Exercise 3

Consider the Bermudan option of Exercise 1. Apply the upper bound method described in this Section. Use the different estimators of the exercise time considered in Exercise 2 in order to construct approximate martingales components of the deflated option price.

References I

[1] P. Glasserman, *Monte Carlo methods in Financial Engineering*, Springer, New York (2004).

[2] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C++: the art of scientific computing*, Cambridge University Press, Cambridge (2002).

[3] P. Boyle, J. Evnine, and S. Gibbs, *Numerical evaluation of multivariate contingent claims*, Review of Financial Studies **2**, 241-250 (1989).

[4] J. Tsitsiklis and B. Van Roy, *Regression methods for complex American-style options*, IEEE Transactions on Neural Networks **12**, 694-703 (2001).

[5] F.A. Longstaff and E. S. Schwartz, *Valuing American options by simulation: a simple least-square approach*, Review of Financial Studies **14**, 113-147 (2001).

[6] S. E. Shreve, *Stochastic calculus for Finance*, Springer-Verlag, New York (2004).

[7] L. Andersen and M. Broadie, *A primal dual algorithm for pricing multidimensional American options*, Management Science **50**, 1222-1234 (2004).