

# Approximate Dynamic Programming (ADP) and Dual Methods for Pricing American Options

---

It is only recently that simulation has begun to play an important role in pricing high-dimensional<sup>1</sup> American options. This was due to the fact that since Monte Carlo simulation generally works forward in time while dynamic programming works backwards, it was generally believed that the two were somewhat incompatible. Research in recent years has shown that this is not the case and that Monte Carlo simulation is an important tool for pricing American options. In these notes we describe some of the latest methods for pricing American options, including the ADP based methods for computing lower bounds on the true option price, and the dual based methods for computing upper bounds. First, however, we describe the properties of the financial market and define the American option pricing problem.

**The Financial Market:** We assume that we have a *dynamically complete*<sup>2</sup> financial market that is driven by a vector-valued Markov process,  $X_t = (X_t^1, \dots, X_t^n)$ .  $X_t$  represents the time  $t$  vector of risky asset prices as well as the values of any relevant state variables in the market. We also assume there exists a risk-free security whose time  $t$  price is  $B_t = e^{rt}$ , where  $r$  is the continuously compounded risk-free rate of interest. Finally, since markets are assumed to be dynamically complete, there exists a unique risk-neutral valuation measure,  $Q$ .

**Option Payoff:** Let  $h_t = h(X_t)$  be a nonnegative adapted process representing the payoff of the option so that if it is exercised at time  $t$  the holder of the option will then receive  $h_t$ .

**Exercise Dates:** The American feature of the option allows the holder of the option to *exercise* it at any of the pre-specified exercise dates in  $\mathcal{T} = \{0, 1, \dots, T\}$ . Strictly speaking, the security under consideration is a *Bermudan* option as it has only a finite number of possible exercise dates. However, we assume that the price of the Bermudan option approaches the price of the corresponding American option by letting the number of exercise dates in  $[0, T]$  go to infinity. In these notes we will use the term “American option” instead of the more correct “Bermudan option”.

**Stopping Time:** A stopping time is an *exercise strategy* for the American option where the decision to exercise or not at any time  $t$  can only depend on information that is available at that time.

**Option Price:** The value process of the American option,  $V_t$ , is the price process of the option conditional on it not having been exercised before  $t$ . It satisfies

$$V_t = \sup_{\tau \geq t} E_t^Q \left[ \frac{B_t h_\tau}{B_\tau} \right]. \quad (1)$$

where  $\tau$  is any stopping time with values in the set  $\mathcal{T} \cap [t, T]$ .

If  $X_t$  is high-dimensional, then standard solution techniques such as dynamic programming become impractical

---

<sup>1</sup>Low-dimensional problems may be solved using standard numerical techniques. These techniques become impractical from a computational standpoint as the dimensionality goes beyond 3 or 4 state variables.

<sup>2</sup>A financial market is dynamically complete if any random variable,  $W_T$ , representing a terminal cash-flow, can be attained from a *self-financing* trading strategy. A trading strategy is said to be self-financing if changes in the value of the portfolio are only due to capital gains or losses. In particular, no addition or withdrawal of funds is allowed after date  $t = 0$  and any new purchases of securities must be financed by the sale of other securities. You do not need to understand the notions of “completeness” or “self-financing” for these notes. They are required only to justify the expression in (1) for the price of the American option.

and we cannot hope to solve the optimal stopping problem (1) exactly. Fortunately, efficient ADP algorithms for addressing this problem have recently been developed<sup>3</sup>. These algorithms use the idea of *cross-path regressions* to construct approximations to the American option price,  $V_0$ . These approximations naturally give rise to a lower bound,  $\underline{V}_0$ , on the true price,  $V_0$ . In section 2 we will show how dual methods may be used to construct an upper bound on  $V_0$ .

## 1 Computing Lower Bounds using Cross-Path Regressions

The pricing problem at time  $t = 0$  is to compute

$$V_0 := \sup_{\tau \in \mathcal{T}} \mathbb{E}_0^Q \left[ \frac{h_\tau}{B_\tau} \right]$$

and in theory this problem is easily solved using value<sup>4</sup> iteration. In particular, we would obtain

$$\begin{aligned} V_T &= h(X_T) \quad \text{and} \\ V_t &= \max \left( h(X_t), \mathbb{E}_t^Q \left[ \frac{B_t}{B_{t+1}} V_{t+1}(X_{t+1}) \right] \right). \end{aligned}$$

The price of the option is then given by  $V_0(X_0)$  where  $X_0$  is the initial state of the economy. As an alternative to value iteration we could use *Q-value iteration*. If the Q-value function is defined to be the value of the option conditional on it not being exercised today (i.e. the continuation value of the option) then we also have

$$Q_t(X_t) = \mathbb{E}_t^Q \left[ \frac{B_t}{B_{t+1}} V_{t+1}(X_{t+1}) \right].$$

The value of the option at time  $t + 1$  is then

$$V_{t+1}(X_{t+1}) = \max(h(X_{t+1}), Q_{t+1}(X_{t+1}))$$

so that we can also write

$$Q_t(X_t) = \mathbb{E}_t^Q \left[ \frac{B_t}{B_{t+1}} \max(h(X_{t+1}), Q_{t+1}(X_{t+1})) \right]. \quad (2)$$

Equation (2) clearly gives a natural analog to value iteration, namely *Q-value iteration*. As stated earlier, if  $n$  is large so  $X_t$  is high dimensional, then both value iteration and Q-value iteration are not feasible in practice. However, we could perform an *approximate* and efficient version of Q-value iteration, and we now describe how to do this.

The first step is to choose a set of *basis functions*,  $\phi_1(X), \dots, \phi_m(X)$ . These basis functions define the *linear architecture* that will be used to approximate the Q-value functions. In particular, we will approximate  $Q_t(X_t)$  with

$$\tilde{Q}_t(X_t) = r_t^1 \phi_1(X_t) + \dots + r_t^m \phi_m(X_t)$$

where  $r_t := (r_t^1, \dots, r_t^m)$  is a vector of time  $t$  parameters that is determined by the algorithm which proceeds as follows:

<sup>3</sup>See Longstaff and Schwartz (2001) and Tsitsiklis and Van Roy (2001).

<sup>4</sup>The term “value iteration” refers to the dynamic programming approach of computing the value function iteratively by working backwards in time. This is how we compute the price of American options in the binomial model.

**Algorithm for Approximate Q-Value Iteration**

```

generate  $N$  paths of state vector,  $X$ 
set  $\tilde{Q}_T(X_T^i) = 0$  for all  $i = 1$  to  $N$ 
for  $t = T - 1$  Down to  $1$ 
    Estimate  $r_t = (r_t^1, \dots, r_t^m)$ 
    set  $\tilde{Q}_t(X_t^i) = \sum_k r_t^k \phi_k(X_t^i)$  for all  $i$ 
end for
set  $\tilde{V}_0(X_0) = \max(h(X_0), \tilde{Q}_0(X_0))$ 

```

Two steps require further explanation.

1. First, we estimate  $r_t$  by regressing  $\alpha \max(h(X_{t+1}), \tilde{Q}(X_{t+1}))$  on  $(\phi_1(X_t), \dots, \phi_m(X_t))$  where  $\alpha = B_t/B_{t+1}$  is the discount factor. We have  $N$  observations for this regression and  $N$  is usually taken to be somewhere between 10,000 and 50,000, though it will of course depend on the problem at hand.
2. Second, since all  $N$  paths have the same starting point,  $X_0$ , we can estimate  $\tilde{Q}_0(X_0)$  by averaging and discounting  $\tilde{Q}_1(\cdot)$  evaluated at the  $N$  successor points of  $X_0$ .

**Remark 1** Obviously many more details are required to fully specify the algorithm. In particular, parameter values and basis functions need to be chosen and specific implementation details can vary.

In practice, it is quite common for an alternative estimate,  $\underline{V}_0$ , of  $V_0$  to be obtained by simulating the exercise strategy that is defined implicitly by the sequence of Q-value function approximations. That is, we define  $\tilde{\tau} = \min\{t \in \mathcal{T} : \tilde{Q}_t \leq h_t\}$  and

$$\underline{V}_0 = E_0^Q \left[ \frac{h_{\tilde{\tau}}}{B_{\tilde{\tau}}} \right].$$

$\underline{V}_0$  is then an unbiased lower bound on the true value of the option as it is the price that corresponds to a feasible exercise strategy.

**Exercise 1** Given how the lower bound,  $\underline{V}_0$ , is computed, can you guess why we prefer to do an approximate Q-value iteration instead of an approximate value-iteration?

These algorithms have performed surprisingly well on realistic high-dimensional problems<sup>5</sup> and there has also been considerable theoretical<sup>6</sup> work explaining why this is so. The quality of  $\underline{V}_0$ , for example, can be explained in part by noting that exercise errors are never made as long as  $Q_t(\cdot)$  and  $\tilde{Q}_t(\cdot)$  lie on the *same* side of the optimal exercise boundary. This means in particular, that it is possible to have large errors in  $\tilde{Q}_t(\cdot)$  that do not impact the quality of  $\underline{V}_0$ .

Clearly, simulation plays an important role in these ADP algorithms as it is required to generate the  $N$  sample paths of  $X$  and to estimate  $\underline{V}_0$ . There are also opportunities for simulation techniques to significantly improve the efficiency of these ADP algorithms as well as the dual-based methods of Section 2 that can be used to evaluate ADP solutions.

<sup>5</sup>See Longstaff and Schwartz (2001) for numerical examples.

<sup>6</sup>See Tsitsiklis and Van Roy (2001).

## 2 Computing Upper Bounds using Dual Methods

While ADP methods have been very successful, an important weakness is their inability to determine how far the ADP solution is from the optimal solution in any given problem. Dual-based methods have recently been developed<sup>7</sup> for evaluating any approximate solution by using it to construct an upper bound on the true value function. This is of value since we can then compute lower<sup>8</sup> and upper bounds on the optimal price of the American option and therefore determine how far the approximate solution is from optimality. We now describe these dual based methods.

For an arbitrary supermartingale,  $\pi_t$ , the value of an American option,  $V_0$ , satisfies

$$\begin{aligned} V_0 = \sup_{\tau \in \mathcal{T}} E_0^Q \left[ \frac{h_\tau}{B_\tau} \right] &= \sup_{\tau \in \mathcal{T}} E_0^Q \left[ \frac{h_\tau}{B_\tau} - \pi_\tau + \pi_\tau \right] \\ &\leq \sup_{\tau \in \mathcal{T}} E_0^Q \left[ \frac{h_\tau}{B_\tau} - \pi_\tau \right] + \sup_{\tau \in \mathcal{T}} E_0^Q [\pi_\tau] \\ &\leq \sup_{\tau \in \mathcal{T}} E_0^Q \left[ \frac{h_\tau}{B_\tau} - \pi_\tau \right] + \pi_0 \\ &\leq E_0^Q \left[ \max_{t \in \mathcal{T}} \left( \frac{h_t}{B_t} - \pi_t \right) \right] + \pi_0 \end{aligned} \quad (3)$$

where the second inequality follows from the optional sampling theorem for supermartingales. Taking the infimum over all supermartingales,  $\pi_t$ , on the right hand side of (3) implies

$$V_0 \leq U_0 := \inf_{\pi} E_0^Q \left[ \max_{t \in \mathcal{T}} \left( \frac{h_t}{B_t} - \pi_t \right) \right] + \pi_0 \quad (4)$$

On the other hand, it is a known fact that the process  $V_t/B_t$  is itself a supermartingale, which implies

$$U_0 \leq E_0^Q \left[ \max_{t \in \mathcal{T}} (h_t/B_t - V_t/B_t) \right] + V_0.$$

Since  $V_t \geq h_t$  for all  $t$ , we conclude that  $U_0 \leq V_0$ . Therefore,  $V_0 = U_0$ , and equality is attained when  $\pi_t = V_t/B_t$ .

This shows that an upper bound on the price of the American option can be constructed simply by evaluating the right-hand-side of (3) for a given supermartingale,  $\pi_t$ . In particular, if such a supermartingale satisfies  $\pi_t \geq h_t/B_t$ , the option price  $V_0$  is bounded above by  $\pi_0$ .

When the supermartingale  $\pi_t$  in (3) coincides with the discounted option value process,  $V_t/B_t$ , the upper bound on the right-hand-side of (3) equals the true price of the American option. This suggests that a tight upper bound can be obtained by using an accurate approximation,  $\tilde{V}_t$ , to define  $\pi_t$ . One possibility<sup>9</sup> is to define  $\pi_t$  as a martingale:

$$\pi_0 = \tilde{V}_0 \quad (5)$$

$$\pi_{t+1} = \pi_t + \frac{\tilde{V}_{t+1}}{B_{t+1}} - \frac{\tilde{V}_t}{B_t} - E_t \left[ \frac{\tilde{V}_{t+1}}{B_{t+1}} - \frac{\tilde{V}_t}{B_t} \right]. \quad (6)$$

Let  $\bar{V}_0$  denote the upper bound we get from (3) corresponding to our choice of supermartingale in (5) and (6). Then it is easy to see that the upper bound is explicitly given by

$$\bar{V}_0 = \tilde{V}_0 + E_0^Q \left[ \max_{t \in \mathcal{T}} \left( \frac{h_t}{B_t} - \frac{\tilde{V}_t}{B_t} + \sum_{j=1}^t E_{j-1}^Q \left[ \frac{\tilde{V}_j}{B_j} - \frac{\tilde{V}_{j-1}}{B_{j-1}} \right] \right) \right]. \quad (7)$$

<sup>7</sup>See Haugh and Kogan (2001) and Rogers (2002).

<sup>8</sup>As we saw in Section 1, a lower bound is easy to compute.

<sup>9</sup>See Haugh and Kogan (2002) and Andersen and Broadie (2002) for further comments related to the choice of  $\pi_t$ .

As may be seen from (7), obtaining an accurate estimate of  $\bar{V}_0$  is computationally demanding. First, a number of sample paths must be simulated to estimate the outermost expectation on the right-hand-side of (7). While this number need not be large in practice, we also need to accurately estimate a conditional expectation at each time period along each simulated path. This requires some effort and clearly variance reduction methods would be useful in this context. Low discrepancy sequences are also very useful for this task.

Variations and extensions of these algorithms have also been developed recently and are a subject of ongoing research.<sup>10</sup>

### 3 Portfolio Optimization

These notes have concentrated on the problem of pricing American options. However, Monte Carlo methods may also be used for another important class of *control* problems in financial engineering, namely portfolio optimization problems. They can be used to approximately solve these problems when explicit solutions are unavailable and the dimensionality is too high for standard numerical techniques.

In particular, the cross-path regressions method for pricing American options has been extended<sup>11</sup> to portfolio optimization problems. This in turn has been extended<sup>12</sup> to certain classes of portfolio optimization problems when there are constraints on the types of portfolios that may be held, e.g. no-short selling or no-borrowing constraints.

### References

- Andersen, L. and M. Broadie. 2001. A Primal-Dual Simulation Algorithm for Pricing Multi-Dimensional American Options. Working Paper, Columbia University.
- Brandt, M.W., A. Goyal, P. Santa-Clara, and J.R. Stroud. 2001. A Simulation Approach to Dynamic Portfolio Choice with an Application to Learning in Discrete Time. Working paper, Duke University.
- Glasserman, P. 2003. *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, New York.
- Haugh, M.B., and L. Kogan. 2004. Pricing American Options: A Duality Approach. *Operations Research*, Vol. 52, No.2, March-April 2004, pp 258-270
- Haugh, M.B., L. Kogan and Z. Wu. 2004. Portfolio Optimization: A Duality Approach. Working paper, Department of IE & OR, Columbia University.
- Longstaff, F., and E. Schwartz. 2001. Valuing American Options by Simulation: A Simple Least-Squares Approach. *Review of Financial Studies* 14:113-147.
- Rogers, L.C.G. 2002. Monte-Carlo Valuation of American Options. *Mathematical Finance*. 12(3):271-286.
- Tsitsiklis, J., and B. Van Roy. 2001. Regression Methods for Pricing Complex American-Style Options. *IEEE Transactions on Neural Networks* 12(4):694-703.

<sup>10</sup>See Chapter 8 of Glasserman (2003) for a detailed treatment of Monte Carlo methods for pricing American options.

<sup>11</sup>See Brandt et al (2001).

<sup>12</sup>See Haugh et al (2004).