J.R. Birge and V. Linetsky (Eds.), *Handbooks in OR & MS, Vol. 15*

DOI: 10.1016/S0927-0507(07)15022-2

Chapter 22

Duality Theory and Approximate Dynamic Programming for Pricing American Options and Portfolio Optimization[1]

Martin B. Haugh
Department of IE and OR, Columbia University, New York, NY 10027, USA
E-mail: martin.haugh@columbia.edu

Leonid Kogan
Sloan School of Management, MIT, Cambridge, MA 02139, USA
E-mail: lkogan@mit.edu

Abstract

This chapter describes how duality and approximate dynamic programming (ADP) methods can be used in financial engineering. It focuses on American option pricing and portfolio optimization problems when the underlying state space is high-dimensional. In general, it is not possible to solve these problems exactly due to the so-called "curse of dimensionality" and as a result, approximate solution techniques are required. ADP and dual-based methods have been proposed for constructing and evaluating good approximate solutions to these problems. In this chapter we describe these methods. Some directions for future research are also outlined.

1 Introduction

Portfolio optimization and American option pricing problems are among the most important problems in financial engineering. Portfolio optimization problems occur throughout the financial services as pension funds, mutual funds, insurance companies, endowments and other financial entities all face the fundamental problem of dynamically allocating their resources across different securities in order to achieve a particular goal. These problems are often

[1] This chapter is a revised and extended version of Haugh [Haugh, M.B. (2003). Duality theory and simulation in financial engineering. In: Chick, S., Sánchez, P.J., Ferrin, D., Morrice, D.J. (Eds.), *Proceedings of the 2003 Winter Simulation Conference*, IEEE Press, Piscataway, NJ, pp. 327–334].

very complex owing to their dynamic and stochastic nature, their high dimensionality and the complexity of real-world constraints. While researchers have developed a number of sophisticated models for addressing these problems, the current state-of-the-art is such that explicit solutions are available only in very special circumstances. (See, for example, Merton, 1990; Cox and Huang 1989; Karatzas and Shreve, 1997; Liu, 1998.)

American option pricing has also presented several challenges to the field of financial engineering. Even in the simple Black–Scholes framework (Black and Scholes, 1973), a closed form expression for the price of an American put option is not available and so it must therefore be computed numerically. This does not present a challenge when there is just one or two underlying securities. However, as pricing an American option amounts to solving an optimal stopping problem, Bellman's curse of dimensionality implies that pricing high-dimensional American options using standard numerical techniques is not practically feasible. Unfortunately, the same conclusion also applies to solving general high-dimensional portfolio optimization problems.

Because these high-dimensional problems occur frequently in practice, they are of considerable interest to both researchers and practitioners. In recent years there has been some success in tackling these problems using approximate dynamic programming (ADP) and dual-based methods. ADP methods (see, for example, Bertsekas and Tsitsiklis, 1996) have had considerable success in tackling large-scale complex problems and have recently been applied successfully to problems in financial engineering (Brandt et al., 2005; Longstaff and Schwartz, 2001; Tsitsiklis and Van Roy, 2001). One difficulty with ADP, however, is in establishing how far the sub-optimal ADP solution to a given problem is from optimality. In the context of optimal stopping problems and pricing American options, Haugh and Kogan (2004) and Rogers (2002) developed dual formulations[2] which allows one to evaluate sub-optimal strategies, including those obtained from ADP methods (see, for example, Haugh and Kogan, 2004; Anderson and Broadie, 2004; Glasserman and Yu, 2004; Chen and Glasserman, 2007). A stochastic duality theory also exists for portfolio optimization problems and this has been developed by many researchers in recent years (see, for example, Shreve and Xu, 1992a, 1992b; He and Pearson, 1991; Cvitanic and Karatzas, 1992; Karatzas and Shreve, 1997). While this theory has had considerable success in characterizing optimal solutions, explicit solutions are still rare (see Rogers, 2003). Recently Haugh et al. (2003) have shown how some of these dual formulations can be used to evaluate suboptimal policies by constructing lower and upper bounds on the true optimal value function. These suboptimal policies could be simple heuristic policies or policies resulting from some approximation techniques such as ADP.

[2] The dual formulation in these papers relies on an alternative characterization of an optimal stopping problem, which can be traced back to Davis and Karatzas (1994).

Another promising approach to approximate solution of dynamic portfolio choice problems is based on the equivalent linear programming formulation of the dynamic program (see, for example, de Farias and Van Roy, 2003, 2004). The approximate linear programming formulation provides an approximation to the portfolio policy as well as an upper bound on the value function. This approach is computationally intensive and its ability to handle large-scale practical problems still needs to be evaluated. Some encouraging results in this direction are obtained by Han (2005).

Simulation techniques play a key role in both the ADP and dual-based evaluation methods that have been used to construct and evaluate solutions to these problems. While it has long been recognized that simulation is an indispensable tool for financial engineering (see the surveys of Boyle et al., 1997; Staum, 2002), it is only recently that simulation has begun to play an important role in solving *control* problems in financial engineering. These control problems include portfolio optimization and the pricing of American options, and they are the focus of this paper.

The remainder of the paper is outlined as follows. Section 2 describes the American option pricing problem. We briefly describe the ADP methods in Section 2.1 after which we will focus on the duality theory for optimal stopping in Section 2.2. Section 2.3 discusses extensions and other applications of these dual-based ideas. We then conclude Section 2 by outlining some directions for future research in Section 2.4. Section 3 concentrates on portfolio optimization. In Sections 3.1 and 3.2, respectively, we describe the portfolio optimization framework and review the corresponding duality theory. In Section 3.3 we show how an upper bound on the portfolio optimization problem can be obtained and we summarize the algorithm for obtaining lower and upper bounds in Section 3.4. We conclude by outlining directions for future research in Section 3.5. Results will not be presented in their full generality, and technical details will often be omitted as we choose to focus instead on the underlying concepts and intuition.

2 Pricing American options

The financial market. We assume there exists a dynamically complete financial market that is driven by a vector-valued Markov process, $X_t = (X_t^1, \ldots, X_t^n)$. In words, we say a financial market is dynamically complete if any random variable, W_T, representing a terminal cash-flow can be attained by using a *self-financing trading strategy*. (A self-financing trading strategy is a strategy where changes in the value of the portfolio are only due to accumulation of dividends and capital gains or losses. In particular, no net 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.) 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.[3] Finally, since markets are assumed to be dynamically complete, there exists (see Duffie, 1996) a unique risk-neutral valuation measure, $\mathcal{Q}$.

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 $\mathcal{T} = \{0, 1, \ldots, T\}$.[4]

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

$$V_t = \sup_{\tau \geqslant t} \mathrm{E}_t^{\mathcal{Q}}\left[\frac{B_t h_\tau}{B_\tau}\right], \tag{1}$$

where τ 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 and we cannot hope to solve the optimal stopping problem (1) exactly. Fortunately, efficient ADP algorithms for addressing this problem have recently been developed independently by Longstaff and Schwartz (2001) and Tsitsiklis and Van Roy (2001).[5] We now briefly describe the main ideas behind these algorithms, both of which rely on the ability to simulate paths of the underlying state vectors.

2.1 ADP for pricing American options

Once again, the pricing problem at time $t = 0$ is to compute

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

[3] Note that we can easily handle the case where $r_t = r(X_t)$ is stochastic.

[4] Strictly speaking, we consider Bermudan options that may only be exercised at one of a finite number of possible dates. While American options can be exercised at any time in a continuum of dates, in practice it is necessary to discretize time when pricing them numerically. As a result, we do not distinguish between Bermudan and American options in this chapter.

[5] See also Carriére (1996) for the original introduction of regression-based ideas for pricing American options.

and in theory this problem is easily solved using value iteration. In particular, a standard recursive argument implies

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

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

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

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

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

so that we can also write

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

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 this is precisely what the ADP algorithms of Longstaff and Schwartz (2001) and Tsitsiklis and Van Roy (2001) do. We now describe their main contribution, omitting some of the more specific details that can nevertheless have a significant impact on performance.

The first step is to choose a set of *basis functions*, $\phi_1(\cdot), \ldots, \phi_m(\cdot)$. 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

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

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

Approximate Q-value iteration

generate N paths of state vector, X, conditional on initial state, X_0
set $\tilde{Q}_T(X_T^i) = 0$ for all $i = 1$ to N
for $t = T-1$ *downto* 1
  regress $B_t \tilde{V}_{t+1}(X_{t+1}^i)/B_{t+1}$ on $(\phi_1(X_t^i), \ldots, \phi_m(X_t^i))$ where
    $\tilde{V}_{t+1}(X_{t+1}^i) := \max\big(h(X_{t+1}^i), \tilde{Q}(X_{t+1}^i)\big)$

set $\tilde{Q}_t(X_t^i) = \sum_k r_t^k \phi_k(X_t^i)$ where the r_t^ks are the estimated regression coefficients

end for

generate M samples of state vector, X_1, conditional on initial state, X_0

set $\tilde{V}_0(X_0) = (\sum_{j=1}^M \max(h(X_1^j), \tilde{Q}_1(X_1^j)))/MB_1$.

The key idea in this algorithm is the use of regression methods to estimate $\tilde{Q}_t(X_t^i)$. In practice, standard least squares regression is used and because this technique is so fast the resulting Q-value iteration algorithm is also very fast. For typical problems that arise in practice, N is often taken to be on the order of 10,000 to 50,000. Obviously, many more details are required to fully specify the algorithm. In particular, parameter values and basis functions need to be chosen. It is generally a good idea to use problem-specific information when choosing the basis functions. For example, if the value of the corresponding European option is available in closed form then that would typically be an ideal candidate for a basis function. Other commonly used basis functions are the intrinsic value of the option and the prices of other related derivative securities that are available in closed from.

Specific implementation details can also vary. While the algorithm described above is that of Tsitsiklis and Van Roy (2001), Longstaff and Schwartz (2001) omit states X_t^i where $h(X_t^i) = 0$ when estimating the regression coefficients, r_t^k for $k = 1, \ldots, m$. They also define $\tilde{V}_{t+1}(X_{t+1}^i)$ so that

$$\tilde{V}_{t+1}(X_{t+1}^i) = \begin{cases} h(X_{t+1}^i), & h(X_{t+1}^i) \geqslant \tilde{Q}(X_{t+1}^i), \\ \tilde{V}_{t+2}(X_{t+2}^i)B_{t+1}/B_{t+2}, & h(X_{t+1}^i) < \tilde{Q}(X_{t+1}^i). \end{cases}$$

In particular, they take $\tilde{V}_{t+1}(X_{t+1}^i)$ to be the *realized* discounted payoff on the ith path as determined by the exercise policy, $\tilde{\tau}$, implicitly defined by $\tilde{Q}_l(\cdot)$, for $l = t+1, \ldots, T$.

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 by $\tilde{\tau}$. Formally, we define $\tilde{\tau} = \min\{t \in \mathcal{T}\colon \tilde{Q}_t \leqslant h_t\}$ and

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

$\underline{V}_0$ is then an unbiased lower bound on the true value of the option. That the estimator is a lower bound follows from the fact $\tilde{\tau}$ is a feasible adapted exercise strategy. Typically, $\underline{V}_0$ is a much better estimator of the true price than $\tilde{V}_0(X_0)$ as the latter often displays a significant upwards bias. Glasserman (2004, Section 8.7) provides a very nice intuition for determining when $\tilde{V}_0(X_0)$ performs poorly as an estimator and, using the duality ideas of Section 2.2, he relates the quality of $\tilde{V}_0(X_0)$ to the quality of the chosen basis functions.

These ADP algorithms have performed surprisingly well on realistic high-dimensional problems (see Longstaff and Schwartz, 2001 for numerical examples) and there has also been considerable theoretical work (e.g. Tsitsiklis and Van Roy, 2001; Clemént et al., 2002) justifying this. Clemént et al. (2002), for example, show that the Longstaff and Schwartz algorithm converges as the number of paths, N, goes to infinity and that the limiting price, $\tilde{V}_0(X_0)$, equals the true price, V_0, if Q_t can be written as a linear combination of the chosen basis functions.

Haugh and Kogan (2004) also show that for any approximation, $\tilde{Q}_t$, the quality of the lower bound, $\underline{V}_0$, satisfies

$$V_0 - \underline{V}_0 \leqslant \mathrm{E}_0^{\mathcal{Q}}\left[\sum_{t=0}^{T} \frac{|\tilde{Q}_t - Q_t|}{B_t}\right].$$

While this may suggest that the quality of the lower bound deteriorates linearly in the number of exercise periods, in practice this is not the case. 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 $\underline{V}_0$.

More recently, it has been shown (see Glasserman, 2004; Glasserman and Yu, 2004) how the ADP – regression methods relate to the *stochastic mesh* method of Broadie and Glasserman (1997). In addition, Glasserman and Yu (2004) also study the trade-off between the number of paths, N, and the number of basis functions, m, when there is a finite computational budget available.

To complete this section, it is worth mentioning that there is an alternative to approximating the value function or Q-value function when using ADP methods (or indeed any other approximation methods). That is, we could choose instead to approximate the optimal exercise frontier. The exercise frontier is the boundary in X-space whereby it is optimal to exercise on one side of the boundary and to continue on the other side. It is possible to construct ADP methods that directly approximate this exercise boundary without directly approximating the value function. These methods often require work that is quadratic in the number of exercise periods. That said, in general it is very difficult to conduct a formal comparison between methods that approximate the exercise frontier and methods that approximate the value function.

2.2 Duality theory for American options

While ADP methods have been very successful, a notable weakness is their inability to determine how far the ADP solution is from optimality in any given problem. Haugh and Kogan (2004) and Rogers (2002) independently developed duality-based methods that can be used for constructing upper bounds on the true value function. Haugh and Kogan showed that any approximate solution, arising from ADP or other[6] methods, could be evaluated by using it to construct an upper[7] bound on the true value function. We also remark that Broadie and Glasserman (1997) were the first to demonstrate that tight lower and upper bounds could be constructed using simulation techniques. Their method, however, does not work with arbitrary approximations to the value function and does not appear to be as efficient as the dual-ADP techniques. We now describe these duality-based methods.

For an arbitrary adapted supermartingale, π_t, the value of an American option, V_0, satisfies

$$\begin{aligned} V_0 &= \sup_{\tau\in\mathcal{T}} \mathrm{E}_0^{\mathcal{Q}}\left[\frac{h_\tau}{B_\tau}\right] = \sup_{\tau\in\mathcal{T}} \mathrm{E}_0^{\mathcal{Q}}\left[\frac{h_\tau}{B_\tau} - \pi_\tau + \pi_\tau\right] \\ &\leqslant \sup_{\tau\in\mathcal{T}} \mathrm{E}_0^{\mathcal{Q}}\left[\frac{h_\tau}{B_\tau} - \pi_\tau\right] + \pi_0 \leqslant \mathrm{E}_0^{\mathcal{Q}}\left[\max_{t\in\mathcal{T}}\left(\frac{h_t}{B_t} - \pi_t\right)\right] + \pi_0, \end{aligned} \tag{3}$$

where the first inequality follows from the optional sampling theorem for supermartingales. Taking the infimum over all supermartingales, π_t, on the right-hand side of (3) implies

$$V_0 \leqslant U_0 := \inf_{\pi} \mathrm{E}_0^{\mathcal{Q}}\left[\max_{t\in\mathcal{T}}\left(\frac{h_t}{B_t} - \pi_t\right)\right] + \pi_0. \tag{4}$$

On the other hand, it is known (see e.g. Duffie, 1996) that the process V_t/B_t is itself a supermartingale, which implies

$$U_0 \leqslant \mathrm{E}_0^{\mathcal{Q}}\left[\max_{t\in\mathcal{T}}\left(\frac{h_t}{B_t} - \frac{V_t}{B_t}\right)\right] + V_0.$$

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

It is of interest to note that we could have restricted ourselves to the case where π_t is a strict martingale, as was the case with Rogers (2002). In that case, the Doob–Meyer decomposition theorem and the supermartingale property of V_t/B_t imply the existence of a martingale, M_t, and an increasing, predictable process, A_t, satisfying $A_0 = 0$ and

$$\frac{V_t}{B_t} = M_t - A_t.$$

Taking $\pi_t = M_t$ in (4) we again obtain $U_0 \leqslant V_0$ implying once again that $V_0 = U_0$. These results demonstrate that an upper bound on the price of the American option can be constructed simply by evaluating the right-hand side

[6] See, for example, the iterative technique of Kolodko and Schoenmakers (2006) who construct upper as well as lower bounds on the true option price using the dual formulations we describe in this section.
[7] As we saw in Section 2.1, a lower bound is easy to compute given an approximation to the value function.

of (3) for a given supermartingale, π_t. In particular, if such a supermartingale satisfies $\pi_t \geqslant h_t/B_t$, the option price V_0 is bounded above by π_0.

When the supermartingale π_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 π_t. One possibility[8] is to define π_t as the following martingale[9]

$$\pi_0 = \tilde{V}_0, \tag{5}$$

$$\pi_{t+1} = \pi_t + \frac{\tilde{V}_{t+1}}{B_{t+1}} - \frac{\tilde{V}_t}{B_t} - \mathrm{E}_t\left[\frac{\tilde{V}_{t+1}}{B_{t+1}} - \frac{\tilde{V}_t}{B_t}\right]. \tag{6}$$

Let $\overline{V}_0$ denote the upper bound we obtain 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

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

As may be seen from (7), obtaining an accurate estimate of $\overline{V}_0$ can be 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 can be quite small 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. Alternatively, if the initial value function approximation comes from ADP methods then, as suggested by Glasserman and Yu (2004), it might be possible to choose the basis functions in such a way that the conditional expectations in (7) can be computed analytically. In that case the need for conducting nested simulations would not arise.

2.3 Extensions

A number of variations and extensions of these algorithms have also been developed recently and are a subject of ongoing research. Andersen and Broadie (2004), for example, construct upper bounds by using an approximation to the optimal exercise frontier instead of an approximation to the Q-value function, while Meinshausen and Hambly (2004) use similar ideas to price options that may be exercised on multiple occasions. Jamshidan (2003)

[8] See Haugh and Kogan (2004) and Andersen and Broadie (2004) for further comments relating to the superiority of taking π_t to be a strict martingale.

[9] Haugh and Kogan (2004) also propose an alternative where π_t is constructed from $\tilde{V}_t$ in a multiplicative manner.

developed a multiplicative dual approach for constructing upper bounds and Chen and Glasserman (2007) compare this multiplicative dual approach with the additive approach of Haugh and Kogan (2004) and Rogers (2002). In this section we briefly describe these extensions. All of them, with the exception of Meinshausen and Hambly (2004), deal with pricing American options, or equivalently, optimal stopping problems. Brandt et al. (2005) extended the ADP ideas for optimal stopping to portfolio optimization problems. Haugh et al. (2003) showed how a duality theory that had already existed for portfolio optimization problems could be used in practice to create upper bounds on the solutions to these problems. These extensions to portfolio optimization problems will be described in Section 3.

2.3.1 Upper bounds from stopping rules

Approximations to the optimal exercise frontier can also be used to construct upper bounds. For example, suppose τ_i for $i = 1, \ldots, T$ is a sequence of stopping times with the property $\tau_i \geqslant i$ for all i. We interpret τ_i as the time at which the American option should be exercised (under some policy) given that it has not already been exercised before time i. These stopping times might, for example, be constructed from an approximation, $\tilde{Q}_t$, to the Q-value function so that $\tau_i := \min\{t \in \mathcal{T},\ t \geqslant i\colon\ \tilde{Q}_t \leqslant h_t\}$. Alternatively, τ_i may derive from a direct approximation to the exercise frontier. In this case,

$$\tau_i := \min\{t \in \mathcal{T},\ t \geqslant i\colon\ g_t = 1\}, \tag{8}$$

where $g_t = 1$ if the policy says "exercise" and $g_t = 0$ if the policy says "continue." Note that it is not necessary to have an approximation to the value function available when τ_i is defined in this manner.

Regardless of how τ_i is defined we can use it to construct martingales by setting $\tilde{M}_t := \sum_j^t \Delta_j$ where

$$\Delta_j := \mathrm{E}_j^{\mathcal{Q}}[h_{\tau_j}] - \mathrm{E}_{j-1}^{\mathcal{Q}}[h_{\tau_j}] = V_j - Q_{j-1}. \tag{9}$$

We can then take $\pi_t := \tilde{M}_t$ in (4) to construct upper bounds as before. It is necessary to simulate the stopping time τ_i as defined by (8) to estimate the Δ_js. This additional or *nested* simulation is required at each point along each simulated path when estimating the upper bound. This therefore suggests that the computational effort required to compute $\overline{V}_0$ when a stopping time is used is quadratic in the number of time periods. In contrast, it appears that the computational effort is linear when $\tilde{Q}_t$ is used to construct the upper bound of Section 2.2. However, an approximation to the optimal exercise frontier is likely to be more 'accurate' than an approximation to the Q-value function and so a more thorough analysis would be required before the superiority of one approach over the other could be established.

The stopping rule approach was proposed by Andersen and Broadie (2004) but see also Glasserman (2004) for further details. It is also worth mentioning that it is straightforward to combine the two approaches. In particular, an explicit approximation, $\tilde{Q}_t$, could be used in some regions of the state space to estimate $\tilde{M}_t$ while a nested simulation to estimate the Δ_js could be used in other regions.

2.3.2 The multiplicative dual approach

An alternative dual formulation, the *multiplicative dual*, was recently formulated by Jamshidian (2003) for pricing American options. Using the *multiplicative* Doob–Meyer decomposition for supermartingales, Jamshidan showed that the American option price, V_0, could be represented as

$$V_0 = \inf_{M \in \mathcal{M}^+} \mathrm{E}_0^M\left[\max_{0 \leqslant t \leqslant T} \frac{h_t}{M_t}\right] := \inf_{M \in \mathcal{M}^+} \mathrm{E}_0^Q\left[\max_{0 \leqslant t \leqslant T} \frac{h_t}{M_t} M_T\right], \tag{10}$$

where $\mathcal{M}^+$ is the set of all positive martingales, M_t, with $M_0 = 1$. Equation (10) suggests that if we choose a 'good' martingale, $\tilde{M}_t \in \mathcal{M}^+$ with $\tilde{M}_0 = 1$, then

$$\overline{V}_0 := \mathrm{E}_0^Q\left[\max_{0 \leqslant t \leqslant T} \frac{h_t}{\tilde{M}_t} \tilde{M}_T\right]$$

should provide a good upper bound on V_0. As was the case with the *additive* approaches of Section 2.2, it is possible to construct a candidate martingale, $\tilde{M}_t$, using an approximation, $\tilde{V}_t$, to the true value function, V_t. As usual, this upper bound can be estimated using Monte Carlo methods.

Chen and Glasserman (2007) compare this multiplicative dual formulation with the additive-dual formulations of Rogers (2002) and Haugh and Kogan (2004). They show that neither formulation dominates the other in the sense that any multiplicative dual can be improved by an additive dual and that any additive dual can be improved by a multiplicative dual. They also compare the bias and variance of the two formulations and show that either method may have a smaller bias. The multiplicative method, however, typically has a variance that grows much faster than the additive method. While the multiplicative formulation is certainly of theoretical interest, in practice it is likely to be dominated by the additive approach.

Bolia et al. (2004) show that the dual formulation of Jamshidian may be interpreted using an importance sampling formulation of the problem. They then use importance sampling techniques and nonnegative least square methods for function approximation in order to estimate the upper bound associated with a given approximation to the value function. Results are given for pricing an American put option on a single stock that follows a geometric Brownian motion. While they report some success, considerable work remains before these ideas can be applied successfully to high-dimensional problems. In addition, since the multiplicative dual formulation tends to have a much higher variance than the additive formulation, importance sampling methods might have more of an impact if they could be successfully applied to additive formulations. More generally, importance sampling methods should also be of interest when constructing value function approximations in the first place.

2.3.3 Multiple exercise opportunities

In an interesting extension, Meinshausen and Hambly (2004), extend the ADP-dual techniques to *multiple exercise* options. If $\mathcal{T} = \{0, 1, \ldots, T\}$ are the possible exercise dates, then a multiple exercise option with n exercise opportunities may be exercised at any of n dates in $\mathcal{T}$. Clearly $n \leqslant T+1$ and the case where $n = 1$ corresponds to a standard American option. The standard examples of a multiple exercise option is a *swing* option that is traded in the energy derivatives industry. A swing option gives the holder a fixed number of exercise opportunities when electricity may be purchased at a given price. Meinshausen and Hambly (2004) apply ADP methods to construct an approximation to (and lower bound on) the price of swing options.[10] They then use this approximation and a dual formulation to construct an upper bound on the true price of the option. This is completely analogous to the methods for pricing high-dimensional American options though the computational requirements appear to be much greater.

2.4 Directions for future research

There are many possible directions for future research. First, it should be possible to employ ADP and duality ideas to other classes of problems. There has of course already been some success in this direction. As described in Section 2.3.2, Meinshausen and Hambly (2004) have extended these results to option pricing problems where multiple exercises are permitted. Haugh et al. (2003) also developed analogous results for dynamic portfolio optimization problems.[11] It should therefore be possible to extend and develop these techniques for solving other classes of control problems. Of particular interest are *real options* problems, which typically embed American-style or multiple-exercise features.

Because ADP-dual techniques require simulation methods and therefore often demand considerable computational effort, it is clear that variance reduction techniques should be of value, particularly as ever more complex problems are solved. We expect importance sampling to be especially useful in this regard. First, it may be used for estimating the value associated with a given approximately optimal policy.[12] Second, and perhaps more interesting, it should prove especially valuable in actually constructing the approximately optimal policy itself. This is because importance sampling ideas can be used to focus the computational effort on the more 'important' regions of the state space when approximating the value function. While these ideas are not new, they have certainly not been fully explored within the ADP literature.

[10] They also price *chooser flexible caps*, fixed income derivative securities that give the holder the right to exercise a given number of caplets over a given horizon.
[11] See Section 3.
[12] See Bolia et al. (2004) and Section 2.3.2.

Estimating the so-called 'greeks' is also a particularly interesting problem. While ADP-dual ideas have now been very successful for pricing high-dimensional American options, efficiently computing the greeks for these problems remains a difficult[13] task.

Another future research direction is to use the dual formulation, possibly in conjunction with the primal formulation, to construct approximate value functions or exercise frontiers. The resulting 'dual' or 'primal–dual' algorithms would be of theoretical interest and we expect they would, in some instances, be superior to 'primal' algorithms that have already been developed. While it is straightforward to design admittedly simple primal–dual style algorithms,[14] there appears to have been little work done on this topic. This, presumably, is due to the great success that ADP-regression algorithms have had in quickly generating good approximations to the true option price. It is possible, however, that this will become a more active research area as more challenging classes of problems are tackled with ADP techniques.

3 Portfolio optimization

Motivated by the success of ADP methods for pricing American options, Brandt et al. (2005) apply similar ideas to approximately solve a class of high-dimensional portfolio optimization problems. In particular, they simulate a large number of sample paths of the underlying state variables and then working backwards in time, they use *cross path regressions* (as we described in the approximate Q-value iteration algorithm) to efficiently compute an approximately optimal strategy. Propagation of errors is largely avoided, and though the price for this is an algorithm that is quadratic in the number of time periods, their methodology can comfortably handle problems with a large number of time periods. Their specific algorithm does not handle portfolio constraints and certain other complicating features, but it should be possible to tackle these extensions using the ADP methods that they and others have developed.

As was the case with ADP solutions to optimal stopping problems, a principal weakness of ADP solutions to portfolio optimization problems is the difficulty in determining how far a given solution to a given problem is from optimality. This issue has motivated in part the research of Haugh et al. (2005) who use portfolio duality theory to evaluate the quality of suboptimal solutions to portfolio optimization problems by constructing lower and upper bounds on the optimal value function. These bounds are evaluated by simulating the stochastic differential equations (see Kloeden and Platen, 1992) that describe the evolution of the state variables in the model in question. In Section 3.2 we

[13] See Kaniel et al. (2006) for an application where dual methods are employed to estimate the greeks of Bermudan-style options.

[14] See, for example, Haugh and Kogan (2004).

describe the particular portfolio duality theory that was used in Haugh et al. (2005) and that was developed by Xu (1990), Shreve and Xu (1992a, 1992b), Karatzas et al. (1991), and Cvitanic and Karatzas (1992).

Before doing so, we remark that the duality theory of Section 3.2 applies mainly to problems in continuous time. ADP techniques, on the other hand, are generally more suited to a discrete time framework. This inconsistency can be overcome by extrapolating discrete-time ADP solutions to construct continuous-time solutions.

3.1 The model

We now state a portfolio choice problem under incomplete markets and portfolio constraints. The problem is formulated in continuous time and stock prices follow diffusion processes.

The investment opportunity set. There are N stocks and an instantaneously riskfree bond. The vector of stock prices is denoted by $P_t = (P_{1t}, \dots, P_{Nt})$ and the instantaneously riskfree rate of return on the bond is denoted by r_t. Without loss of generality, stocks are assumed to pay no dividends. The instantaneous moments of asset returns depend on the M-dimensional vector of state variables X_t:

$$r_t = r(X_t), \tag{11a}$$

$$\mathrm{d}P_t = P_t\big[\mu_P(X_t)\,\mathrm{d}t + \Sigma_P(X_t)\,\mathrm{d}B_t\big], \tag{11b}$$

$$\mathrm{d}X_t = \mu_X(X_t)\,\mathrm{d}t + \Sigma_X(X_t)\,\mathrm{d}B_t, \tag{11c}$$

where $P_0 = 1$, $X_0 = 0$, $B_t = (B_{1t}, \dots, B_{Nt})$ is a vector of N independent Brownian motions, μ_P and μ_X are N- and M-dimensional drift vectors, and Σ_P and Σ_X are diffusion matrices of dimension N by N and M by N, respectively. The diffusion matrix of the stock return process Σ_P is lower-triangular and nondegenerate: $x^\top \Sigma_P \Sigma_P^\top x \geqslant \epsilon \|x\|^2$ for all x and some $\epsilon > 0$. Then, one can define a process η_t, given by

$$\eta_t = \Sigma_{Pt}^{-1}(\mu_{Pt} - r_t).$$

In a market without portfolio constraints, η_t corresponds to the vector of instantaneous market prices of risk of the N stocks (see, e.g., Duffie, 1996, Section 6.G). The process η_t is assumed to be square-integrable so that

$$\mathrm{E}_0\left[\int_0^T \|\eta_t\|^2\,\mathrm{d}t\right] < \infty.$$

Portfolio constraints. A portfolio consists of positions in the N stocks and the riskfree bond. The proportional holdings of risky assets in the total portfolio value are denoted by $\theta_t = (\theta_{1t}, \dots, \theta_{Nt})$. The portfolio policy is assumed to satisfy a square integrability condition: $\int_0^T \|\theta_t\|^2\,\mathrm{d}t < \infty$ almost surely. The value of the portfolio changes according to

$$\mathrm{d}W_t = W_t\big\{\big[r_t + \theta_t^\top(\mu_{Pt} - r_t)\big]\,\mathrm{d}t + \theta_t^\top \Sigma_{Pt}\,\mathrm{d}B_t\big\}. \tag{12}$$

The portfolio weights are restricted to lie in a closed convex set, $\mathbf{K}$, containing the zero vector:

$$\theta_t \in \mathbf{K}. \tag{13}$$

For example, if short sales are not allowed, then the constraint set takes the form $\mathbf{K} = \{\theta\colon\ \theta \geqslant 0\}$. If in addition to prohibiting short sales, borrowing is not allowed, then $\mathbf{K} = \{\theta\colon\ \theta \geqslant 0,\ 1^\top\theta \leqslant 1\}$ where $1^\top = (1, \ldots, 1)$. The set $\mathbf{K}$ can be constant, or it can depend on time and the values of the exogenous state variables.

The objective function. For simplicity, the portfolio policy is chosen to maximize the expected utility of wealth at the terminal date T, $\mathrm{E}_0[U(W_T)]$. Preferences over intermediate consumption would be easy to incorporate. The function $U(W)$ is assumed to be strictly monotone with positive slope, concave, and smooth. Moreover, it is assumed to satisfy the Inada conditions at zero and infinity: $\lim_{W\to 0} U'(W) = \infty$ and $\lim_{W\to\infty} U'(W) = 0$. For instance, a common choice is a constant relative risk aversion (CRRA) utility function $U(W) = W^{1-\gamma}/(1-\gamma)$.

In summary, the portfolio choice problem is to solve for

$$V_0 := \sup_{\{\theta_t\}} \mathrm{E}_0\big[U(W_T)\big] \quad \text{subject to} \quad \text{(11), (12) and (13)}, \tag{$\mathcal{P}$}$$

where V_0 denotes the value function at 0.

3.2 Review of the duality theory

In this section we review the duality theory for the constrained portfolio optimization problem. In particular, the version of duality used in Haugh et al. (2005) is based on the work of Cvitanic and Karatzas (1992).

Starting with the portfolio choice problem ($\mathcal{P}$), one can define a fictitious problem ($\mathcal{P}^{(\nu)}$), based on a different financial market and without the portfolio constraints. First, define the *support function* of $\mathbf{K}$, $\delta(\cdot) : \mathbb{R}^N \to \mathbb{R} \cup \infty$, by

$$\delta(\nu) := \sup_{x\in\mathbf{K}} \big(-\nu^\top x\big). \tag{14}$$

The effective domain of the support function is given by

$$\tilde{\mathbf{K}} := \big\{\nu\colon\ \delta(\nu) < \infty\big\}. \tag{15}$$

Because the constraint set $\mathbf{K}$ is convex and contains zero, the support function is continuous and bounded from below on its effective domain $\tilde{\mathbf{K}}$. Then, one can define the set $\mathbf{D}$ of $\mathcal{F}_t$-adapted $\mathbb{R}^N$ valued processes to be

$$\mathbf{D} := \Bigg\{\nu_t,\ 0 \leqslant t \leqslant T\colon\ \nu_t \in \tilde{\mathbf{K}},\ \mathrm{E}_0\Bigg[\int_0^T \delta(\nu_t)\,\mathrm{d}t\Bigg] + \mathrm{E}_0\Bigg[\int_0^T \|\nu_t\|^2\,\mathrm{d}t\Bigg] < \infty\Bigg\}. \tag{16}$$

For each process, $\nu \in \mathbf{D}$, consider a fictitious market, $M^{(\nu)}$, in which the N stocks and the riskfree bond are traded. There are no constraints in the fictitious market. The diffusion matrix of stock returns in $M^{(\nu)}$ is the same as in the original market. However, the riskfree rate and the vector of expected stock returns are different. In particular, the riskfree rate process and the market price of risk in the fictitious market are defined respectively by

$$r_t^{(\nu)} = r_t + \delta(\nu_t), \tag{17a}$$

$$\eta_t^{(\nu)} = \eta_t + \Sigma_{Pt}^{-1}\nu_t, \tag{17b}$$

where $\delta(\nu)$ is the support function defined in (14). Assume that $\eta_t^{(\nu)}$ is square-integrable.

Because the number of Brownian motions, N, is equal to the number of stocks in the financial market described by (11) and the diffusion matrix is nondegenerate, it can be shown that the unconstrained fictitious market is *dynamically complete*. Dynamic completeness would imply the existence of a unique market-price-of-risk process, η_t and a unique *state-price-density* (SPD) process, π_t. $\pi_t(\omega)$ may be interpreted as the price per-unit-probability of \$1 at time t in the event ω occurs. (See Duffie, 1996 for further details.) It so happens that a portfolio optimization problem in complete markets is particularly easy to solve using martingale methods. Following Cox and Huang (1989), the state-price density process $\pi_t^{(\nu)}$ in the fictitious market is given by

$$\pi_t^{(\nu)} = \exp\Bigg(-\int_0^t r_s^{(\nu)}\,\mathrm{d}s - \frac{1}{2}\int_0^t \eta_s^{(\nu)\top}\eta_s^{(\nu)}\,\mathrm{d}s - \int_0^t \eta_s^{(\nu)\top}\,\mathrm{d}B_s\Bigg), \tag{18}$$

and the vector of expected returns is given by

$$\mu_{Pt}^{(\nu)} = r_t^{(\nu)} + \Sigma_{Pt}\,\eta_t^{(\nu)}.$$

The dynamic portfolio choice problem in the fictitious market without position constraints can be equivalently formulated in a static form (e.g., Cox and Huang, 1989; Karatzas and Shreve, 1997, Section 3):

$$V^{(\nu)} := \sup_{\{W_T\}} \mathrm{E}_0\big[U(W_T)\big] \quad \text{subject to} \quad \mathrm{E}_0\big[\pi_T^{(\nu)} W_T\big] \leqslant W_0. \tag{$\mathcal{P}^{(\nu)}$}$$

Once the optimal terminal wealth is computed, it can then be supported by a dynamic trading strategy. Due to its static nature, the problem $(\mathcal{P}^{(\nu)})$ is easy to solve. For example, when the utility function is of the CRRA type with relative risk aversion γ so that $U(W) = W^{1-\gamma}/(1-\gamma)$, the corresponding value function in the fictitious market is given explicitly by

$$V_0^{(\nu)} = \frac{W_0^{1-\gamma}}{1-\gamma}\mathrm{E}_0\left[\pi_T^{(\nu)\,\frac{\gamma-1}{\gamma}}\right]^{\gamma}. \tag{19}$$

It is easy to see that for any admissible choice of $\nu \in \mathbf{D}$, the value function in (19) gives an upper bound for the optimal value function of the original problem. In the fictitious market, the wealth dynamics of the portfolio are given by

$$\mathrm{d}W_t^{(\nu)} = W_t^{(\nu)}\left[\left(r_t^{(\nu)} + \theta_t^{\top}\Sigma_{Pt}\eta_t^{(\nu)}\right)\mathrm{d}t + \theta_t^{\top}\Sigma_{Pt}\,\mathrm{d}B_t\right], \tag{20}$$

so that

$$\begin{aligned}\frac{\mathrm{d}W_t^{(\nu)}}{W_t^{(\nu)}} - \frac{\mathrm{d}W_t}{W_t} &= \left[\left(r_t^{(\nu)} - r_t\right) + \theta_t^{\top}\Sigma_{Pt}\left(\eta_t^{(\nu)} - \eta_t\right)\right]\mathrm{d}t \\ &= \left[\delta(\nu_t) + \theta_t^{\top}\nu_t\right]\mathrm{d}t.\end{aligned}$$

The last expression is nonnegative according to (14) since $\theta_t \in \mathbf{K}$. Thus, $W_t^{(\nu)} \geqslant W_t\ \forall t \in [0, T]$ and

$$V_0^{(\nu)} \geqslant V_0. \tag{21}$$

Results in Cvitanic and Karatzas (1992) and Schroder and Skiadas (2003) imply that if the original optimization problem has a solution, then the upper bound is "tight," i.e., the value function of the fictitious problem $(\mathcal{P}^{(\nu)})$ coincides with the value function of the original problem $(\mathcal{P})$ at an optimally chosen ν^*:

$$V_0^{(\nu^*)} := \inf_{\{\nu\}} V^{(\nu)} = V_0 \tag{22}$$

(see Schroder and Skiadas, 2003, Proposition 3(b) and Theorems 7 and 9). The above equality holds for all times, and not just at time 0, i.e., $V_t^{(\nu^*)} = V_t$. Cvitanic and Karatzas (1992) have shown that the solution to the original problem exists under additional restrictions on the utility function, most importantly that the relative risk aversion does not exceed one. Cuoco (1997) proves a more general existence result, imposing minimal restrictions on the utility function.

3.3 The performance bound

The theoretical duality results of Section 3.2 suggest that one can construct an upper bound on the value function of the portfolio choice problem $(\mathcal{P})$ by computing the value function of any fictitious problem $(\mathcal{P}^{(\nu)})$. The fictitious market is defined by the process ν_t as in (17). Of course, one can pick any fictitious market from the admissible set **D** to compute an upper bound. Such a bound is uninformative if it is too loose. Since the objective is to evaluate a particular candidate policy, one should construct a process $\hat{\nu}_t$ based on such a policy to obtain tighter bounds. The solution to the portfolio choice problem under the fictitious market defined by $\hat{\nu}_t$ then provides a performance bound on the candidate policy.

In order to construct the fictitious market as defined by $\hat{\nu}_t$, Haugh et al. (2005) first use the solution to the dual problem to establish the link between the optimal policy, θ^*, and value function, $V_0 = V_0^{(\nu^*)}$, and the corresponding fictitious asset price processes, as defined by ν^*. Not knowing the optimal portfolio policy and value function, they instead use approximations to obtain the candidate process for ν^*, which is denoted by $\tilde{\nu}$. This candidate process in general does not belong to **D** and cannot be used to define a fictitious problem. Instead, one must search for a qualified process $\hat{\nu} \in \mathbf{D}$, which is close to $\tilde{\nu}$. Haugh et al. (2005) then use $\hat{\nu}$ as an approximation to ν^* to define the fictitious problem ($\mathcal{P}^{(\hat{\nu})}$). Since $\hat{\nu} \in \mathbf{D}$, the solution to the corresponding unconstrained problem in $M^{(\hat{\nu})}$ provides a valid performance bound for the candidate policy.

The state-price density process is related via an envelope theorem to the value function by

$$d \ln \pi_t^{(\nu^*)} = d \ln \frac{\partial V_t}{\partial W_t}. \tag{23}$$

In particular, the stochastic part of $d \ln \pi_t^{(\nu^*)}$ is equal to the stochastic part of $d \ln \partial V_t / \partial W_t$. If V_t is smooth, Itô's lemma and Equations (18) and (12) imply that

$$\eta_t^* := \eta_t^{(\nu^*)} = -W_t \left(\frac{\partial^2 V_t / \partial W_t^2}{\partial V_t / \partial W_t} \right) \Sigma_{Pt}^\top \theta_t^* - \left(\frac{\partial V_t}{\partial W_t} \right)^{-1} \Sigma_{Xt}^\top \left(\frac{\partial^2 V_t}{\partial W_t \partial X_t} \right), \tag{24}$$

where θ_t^* denotes the optimal portfolio policy for the original problem. In the special but important case of a CRRA utility function the expression for $\eta_t^{(\nu^*)}$ simplifies. In particular, the first term in (24) simplifies to $\gamma \Sigma_{Pt}^\top \theta_t^*$, where γ is the relative risk aversion coefficient of the utility function, and one only needs to compute the first derivative of the value function with respect to the state variables X_t to evaluate the second term in (24). This simplifies the numerical implementation, since it is generally easier to estimate first-order than second-order partial derivatives of the value function.

Given an approximation to the optimal portfolio policy $\tilde{\theta}_t$, one can compute the corresponding approximation to the value function, $\tilde{V}_t$, defined as the conditional expectation of the utility of terminal wealth, under the portfolio policy $\tilde{\theta}_t$. One can then construct a process $\hat{\nu}$ as an approximation to ν^*, using (24). Approximations to the portfolio policy and value function can be obtained using a variety of methods, e.g., the ADP method (see Brandt et al., 2005). Haugh et al. (2005) take $\tilde{\theta}_t$ as given and use it to construct an upper bound on the unknown true value function V_0.

Assuming that the approximate value function $\tilde{V}$ is sufficiently smooth, one can replace V_t and θ_t^* in (24) with $\tilde{V}_t$ and $\tilde{\theta}_t$ and obtain

$$\tilde{\eta}_t := \eta_t^{(\tilde{\nu})} = -W_t \left(\frac{\partial^2 \tilde{V}_t / \partial W_t^2}{\partial \tilde{V}_t / \partial W_t} \right) \Sigma_{Pt}^\top \tilde{\theta}_t - \left(\frac{\partial \tilde{V}_t}{\partial W_t} \right)^{-1} \Sigma_{Xt}^\top \left(\frac{\partial^2 \tilde{V}_t}{\partial W_t \partial X_t} \right). \tag{25}$$

$\tilde{\nu}_t$ is then defined as a solution to (17b).

Obviously, $\tilde{\eta}_t$ is a candidate for the market price of risk in the fictitious market. However, there is no guarantee that $\tilde{\eta}_t$ and the corresponding process $\tilde{\nu}_t$ belong to the feasible set **D** defined by (16). In fact, for many important classes of problems the support function $\delta(\nu_t)$ may be infinite for some values of its argument. Haugh et al. (2005) look for a price-of-risk process $\hat{\eta}_t \in \mathbf{D}$ that is close to $\tilde{\eta}_t$. They choose a Euclidian norm as the measure of distance between the two processes to make the resulting optimization problem tractable.

The requirement that $\hat{\eta}_t \in \mathbf{D}$ is not straightforward to implement computationally. Instead, Haugh et al. (2005) impose a set of tighter uniform bounds,

$$\|\hat{\eta} - \eta\| \leqslant A_1, \tag{26a}$$

$$\delta(\hat{\nu}) \leqslant A_2, \tag{26b}$$

where A_1 and A_2 are positive constants that can be taken to be arbitrarily large. The condition (26a) implies that the process $\hat{\nu}_t$ is square-integrable, since η_t is square integrable and $\|\hat{\eta} - \eta\|^2 = \hat{\nu}^\top ({\Sigma_P}^{-1})^\top {\Sigma_P}^{-1} \hat{\nu} \geqslant A \|\hat{\nu}\|^2$ for some $A > 0$. Haugh et al. (2005) provide a discussion on the choice of constants A_1 and A_2.

In summary, $\hat{\eta}_t$ and $\hat{\nu}_t$ are defined as a solution of the following problem:

$$\min_{\hat{\nu}, \hat{\eta}} \|\hat{\eta} - \tilde{\eta}\|^2, \tag{27}$$

subject to

$$\hat{\eta} = \eta + \Sigma_P^{-1} \hat{\nu}, \tag{28a}$$

$$\delta(\nu) < \infty, \tag{28b}$$

$$\|\hat{\eta} - \eta\| \leqslant A_1, \tag{28c}$$

$$\delta(\hat{\nu}) \leqslant A_2. \tag{28d}$$

The value of $\hat{\eta}_t$ and $\hat{\nu}$ can be computed quite easily for many important classes of portfolio choice problems. The following two examples are taken from Haugh et al. (2005).

Incomplete markets. Assume that only the first L stocks are traded so that the positions in the remaining $N - L$ stocks are restricted to the zero level. In this case the set of feasible portfolio policies is given by

$$\mathbf{K} = \{\theta \mid \theta_i = 0 \text{ for } L < i \leqslant N\} \tag{29}$$

and hence the support function $\delta(\nu)$ is equal to zero if $\nu_i = 0, 1 \leqslant i \leqslant L$ and is infinite otherwise. Thus, as long as $\nu_i = 0, 1 \leqslant i \leqslant L$, the constraint (26b) does not need to be imposed explicitly. To find $\hat{\eta}$ and $\hat{\nu}$, one must solve

$$\min_{\hat{\eta},\hat{\nu}} \|\hat{\eta} - \tilde{\eta}\|^2, \tag{30}$$

subject to

$$\hat{\eta} = \eta + \Sigma_P{}^{-1}\hat{\nu},$$
$$\hat{\nu}_i = 0, \quad 1 \leqslant i \leqslant L,$$
$$\|\hat{\eta} - \eta\|^2 \leqslant A_1^2.$$

The diffusion matrix Σ_P is lower triangular and so is its inverse. Using this, the solution can be expressed explicitly as

$$\hat{\eta}_i = \eta_i, \quad 1 \leqslant i \leqslant L,$$
$$\hat{\eta}_j = \eta_j + a(\tilde{\eta}_j - \eta_j), \quad L < j \leqslant N,$$
$$\hat{\nu} = \Sigma_P(\hat{\eta} - \eta),$$

where

$$a = \min\left[1, \left(\frac{A_1^2 - \|\tilde{\eta} - \eta\|_{(L)}^2}{\|\tilde{\eta} - \eta\|^2 - \|\tilde{\eta} - \eta\|_{(L)}^2}\right)^{1/2}\right], \quad \|\tilde{\eta}\|_{(L)}^2 = \sum_{i=1}^{L} \tilde{\eta}_i^2.$$

Incomplete markets, no short sales, and no borrowing. The market is the same as in the previous case, but no short sales and borrowing are allowed. Then the set of admissible portfolios is given by

$$\mathbf{K} = \left\{\theta \mid \theta \geqslant 0,\ 1^\top\theta \leqslant 1,\ \theta_i = 0 \text{ for } L < i \leqslant N\right\}. \tag{31}$$

The support function is given by $\delta(\nu) = \max(0, -\nu_1, \ldots, -\nu_L)$, which is finite for any vector ν. Because in this case $\delta(\nu) \leqslant \|\nu\|$, the relation $\|\nu\| = \|\Sigma_P(\hat{\eta} - \eta)\| \leqslant \|\Sigma_P\| A_1$ implies that as long as A_2 is sufficiently large compared to A_1, one only needs to impose (26a) and (26b) is redundant. We therefore need to solve the following problem:

$$\min_{\hat{\eta},\hat{\nu}} \|\hat{\eta} - \tilde{\eta}\|^2, \tag{32}$$

subject to

$$\hat{\eta} = \eta + \Sigma_P{}^{-1}\hat{\nu},$$
$$\|\hat{\eta} - \eta\|^2 \leqslant A_1^2.$$

Then the fictitious market is described by

$$\hat{\eta} = \eta + \min\left(1, \frac{A_1}{\|\tilde{\eta} - \eta\|}\right)(\tilde{\eta} - \eta),$$
$$\hat{\nu} = \Sigma_P(\hat{\eta} - \eta).$$

3.4 Summary of the algorithm

Bounds on the optimal value function can be computed by simulation in several steps.

1. Start with an approximation to the optimal portfolio policy of the original problem and the corresponding approximation to the value function. Both can be obtained using an ADP algorithm, as in Brandt et al. (2005) or Haugh et al. (2005). Alternatively, one may start with an approximate portfolio policy and then estimate the corresponding value function by simulation.
2. Use the approximate portfolio policy and partial derivatives of the approximate value function to construct a process $\tilde{\eta}_t$ according to the explicit formula (25). The process $\tilde{\eta}_t$ is a candidate for the market price of risk in the fictitious market.
3. Construct a process $\hat{\eta}_t$ that is close to $\tilde{\eta}_t$ and satisfies the conditions for the market price risk of a fictitious market in the dual problem. This involves solving the quadratic optimization problem (27).
4. Compute the value function from the static problem $(\mathcal{P}^{(\nu)})$ in the resulting fictitious market defined by the market price of risk process $\hat{\eta}_t$. This can be accomplished efficiently using Monte Carlo simulation. This results in an upper bound on the value function of the original problem.

The lower bound on the value function is obtained by simulating the terminal wealth distribution under the approximate portfolio strategy.

Successful practical implementation of the above algorithm depends on efficient use of simulation methods. For instance, the expectation in (19) cannot be evaluated explicitly and so it has to be estimated by simulating the underlying SDE's. This is a computationally intensive task, particularly when $\tilde{\nu}_t$ cannot be guaranteed in advance to be well-behaved. In such circumstances it is necessary to solve a quadratic optimization problem at each discretization point on each simulated path in order to convert $\eta_t^{(\tilde{\nu})}$ and $\tilde{\nu}_t$ into well-behaved versions that can then be used to construct an upper bound on V_0. (See Haugh et al., 2005 for further details.)

Besides the actual ADP implementation that constructs the initial approximate solution, simulation is also often necessary to approximate the value function and its partial derivatives in (25). This occurs when we wish to evaluate a given portfolio policy, $\tilde{\theta}_t$, but do not know the corresponding value function, $\tilde{V}_t$. In such circumstances, it seems that it is necessary to simulate the policy, $\tilde{\theta}_t$, in order to approximate the required functions. Once again, this is computationally demanding and seeking efficient simulation techniques for all of these tasks will be an important challenge as we seek to solve ever more complex problems.

Haugh, Kogan and Wu (2005) present several numerical examples illustrating how approximate portfolio policies can be evaluated using duality techniques. While relatively simple, these examples illustrate the potential usefulness of the approach. Haugh et al. (2005) apply the above algorithm to evaluate an ADP solution of the portfolio choice problem in incomplete markets with no-borrowing constraints. Haugh and Jain (2006) use these duality techniques to evaluate other classes of portfolio strategies and to study in further detail the strategies studied by Haugh et al. (2005). Finally, Haugh and Jain (2007) show how path-wise Monte Carlo estimators can be used with the cross-path regression approach to estimate a given portfolio policy's value function as well as its partial derivatives.

3.5 Directions for further research

There are several remaining problems related to the use of duality-based methods in portfolio choice. On the theoretical side, there is room for developing new algorithms designed to tackle more complex and realistic problems. For example, the results summarized above apply to portfolio choice with constraints on proportions of risky assets. However, some important finance problems, such as asset allocation with illiquid/nontradable assets do not fit in this framework. Additional work is required to tackle such problems.

Another important class of problems involve a different kind of market frictions: transaction costs and taxes. Problems of this type are inherently difficult, since the nature of frictions often makes the problem path-dependent and leads to a high-dimensional state space. Some duality results are known for problems with transaction costs (see Rogers, 2003 for a review), while problems with capital gains taxes still pose a challenge. However, note that it is not sufficient to have a dual formulation in order to derive a useful algorithm for computing solution bounds. It is necessary that a high-quality approximation to the optimal dual solution can be recovered easily from an approximate value function. Not all existing dual formulations have such a property and further theoretical developments are necessary to address a broader class of problems.

On the computational side, there is a need for efficient simulation algorithms at various stages of practical implementation of the duality-based algorithms. For instance, motivated by the promising application of importance sampling methods to pricing American options, one could conceivably develop similar techniques to improve performance of portfolio choice algorithms. In particular, in a problem with dynamically complete financial markets and position constraints, approximation errors would tend to accumulate when portfolio constraints are binding. By sampling more frequently from the "problematic" areas in the state space, one could achieve superior approximation quality.

The above discussion has been centered around the problem of evaluating the quality of approximate solutions. A major open problem both on the theoretical and computational fronts is how to use dual formulations to direct the search for an approximate solution. While a few particularly tractable problems have been tackled by duality methods, there are no efficient general algorithms that could handle multi-dimensional problems with nontrivial dynamics and constraints or frictions. Progress on this front may be challenging, but would significantly expand our ability to address outstanding problems in financial engineering theory and practice.

References

Andersen, L., Broadie, M. (2004). A primal–dual simulation algorithm for pricing multi-dimensional American options. *Management Science* 50 (9), 1222–1234.

Bertsekas, D., Tsitsiklis, J. (1996). *Neuro-Dynamic Programming*. Athena Scientific, Belmont, MA.

Black, F., Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–654.

Bolia, N., Glasserman, P., Juneja, S. (2004). Function-approximation-based importance sampling for pricing American options. In: Ingalls, R.G., Rossetti, M.D., Smith, J.S., Peters, B.A. (Eds.), *Proceedings of the 2004 Winter Simulation Conference*. IEEE Press, Piscataway, NJ, pp. 604–611.

Boyle, P., Broadie, M., Glasserman, P. (1997). Monte Carlo methods for security pricing. *Journal of Economic Dynamics and Control* 21, 1267–1321.

Brandt, M.W., Goyal, A., Santa-Clara, P., Stroud, J.R. (2005). A simulation approach to dynamic portfolio choice with an application to learning about return predictability. *Review of Financial Studies* 18, 831–873.

Broadie, M., Glasserman, P. (1997). A stochastic mesh method for pricing high-dimensional American options. *Working paper*, Columbia Business School, Columbia University, New York.

Carriére, J. (1996). Valuation of early-exercise price of options using simulations and non-parametric regressions. *Insurance: Mathematics and Economics* 19, 19–30.

Chen, N., Glasserman, P. (2007). Additive and multiplicative duals for American option pricing. *Finance and Stochastics* 11 (2), 153–179.

Clemént, E., Lamberton, D., Protter, P. (2002). An analysis of a least-squares regression algorithm for American option pricing. *Finance and Stochastics* 6, 449–471.

Cox, J., Huang, C.-F. (1989). Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory* 49, 33–83.

Cuoco, D. (1997). Optimal consumption and equilibrium prices with portfolio constraints and stochastic income. *Journal of Economic Theory* 72, 33–73.

Cvitanić, J., Karatzas, I. (1992). Convex duality in constrained portfolio optimization. *Annals of Applied Probability* 2, 767–818.

de Farias, D., Van Roy, B. (2003). The linear programming approach to approximate dynamic programming. *Operations Research* 51, 850–865.

de Farias, D., Van Roy, B. (2004). On constraint sampling in the linear programming approach to approximate dynamic programming. *Mathematics of Operations Research* 29, 462–478.

Davis, M.H.A., Karatzas, I. (1994). A deterministic approach to optimal stopping, with applications. In: Kelly, F. (Ed.), *Probability, Statistics and Optimization: A Tribute to Peter Whittle*. John Wiley & Sons, New York/Chichester, pp. 455–466.

Duffie, D. (1996). *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, NJ.

Glasserman, P. (2004). *Monte-Carlo Methods in Financial Engineering*. Springer, New York.

Glasserman, P., Yu, B. (2004). Pricing American options by simulation: Regression now or regression later. In: Niederreiter, H. (Ed.), *Monte Carlo and Quasi-Monte Carlo Methods*. Springer-Verlag, Berlin.
Han, J. (2005). Dynamic portfolio management: An approximate linear programming approach. *Ph.D. thesis*, Stanford University.
Haugh, M.B., Jain, A. (2006). On the dual approach to portfolio evaluation. *Working paper*, Columbia University.
Haugh, M.B., Jain, A. (2007). Pathwise estimators and cross-path regressions: An application to evaluating portfolio strategies. In: Henderson, S.G., Biller, B., Hsieh, M.H., Shortle, J., Tew, J.D., Barton, R.R. (Eds.), *Proceedings of the 2007 Winter Simulation Conference*. IEEE Press, in press.
Haugh, M.B., Kogan, L. (2004). Pricing American options: A duality approach. *Operations Research* 52 (2), 258–270.
Haugh, M.B., Kogan, L., Wang, J. (2003). Portfolio evaluation: A duality approach. *Operations Research*, in press, available at: http://www.columbia.edu/~mh2078/Research.html.
Haugh, M.B., Kogan, L., Wu, Z. (2005). Approximate dynamic programming and duality for portfolio optimization. *Working paper*, Columbia University.
He, H., Pearson, N. (1991). Consumption and portfolio policies with incomplete markets and short-sale constraints: The infinite dimensional case. *Journal of Economic Theory* 54 (2), 259–304.
Jamshidian, F. (2003). Minimax optimality of Bermudan and American claims and their Monte-Carlo upper bound approximation. *Working paper*.
Kaniel, R., Tompaidis, S., Zemlianov, A. (2006). Efficient computation of hedging parameters for discretely exercisable options. *Working paper*.
Karatzas, I., Shreve, S.E. (1997). *Methods of Mathematical Finance*. Springer-Verlag, New York.
Karatzas, I., Lehocky, J.P., Shreve, S.E., Xu, G.L. (1991). Martingale and duality methods for utility maximization in an incomplete market. *SIAM Journal Control Optimization* 259, 702–730.
Kloeden, P., Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag, Berlin.
Kolodko, A., Schoenmakers, J. (2006). Iterative construction of the optimal Bermudan stopping time. *Finance and Stochastics* 10, 27–49.
Liu, J. (1998). Dynamic portfolio choice and risk aversion. *Working paper*, Stanford University, Palo Alto.
Longstaff, F., Schwartz, E. (2001). Valuing American options by simulation: A simple least-squares approach. *Review of Financial Studies* 14, 113–147.
Meinshausen, N., Hambly, B.M. (2004). Monte Carlo methods for the valuation of multiple exercise options. *Mathematical Finance* 14, 557–583.
Merton, R.C. (1990). *Continuous-Time Finance*. Basil Blackwell, New York.
Rogers, L.C.G. (2002). Monte-Carlo valuation of American options. *Mathematical Finance* 12 (3), 271–286.
Rogers, L.C.G. (2003). Duality in constrained optimal investment and consumption problems: A synthesis. *Working paper*, Statistical Laboratory, Cambridge University. http://www.statslab.cam.ac.uk/~chris/.
Schroder, M., Skiadas, C. (2003). Optimal lifetime consumption-portfolio strategies under trading constraints and generalized recursive preferences. *Stochastic Processes and Their Applications* 108, 155–202.
Shreve, S.E., Xu, G.L. (1992a). A duality method for optimal consumption and investment under short-selling prohibition, Part I: General market coefficients. *Annals of Applied Probability* 2, 8–112.
Shreve, S.E., Xu, G.L. (1992b). A duality method for optimal consumption and investment under short-selling prohibition, Part I: Constant market coefficients. *Annals of Applied Probability* 2, 314–328.
Staum, J. (2002). Simulation in financial engineering. In: Yücesan, E., Chen, C.-H., Snowdon, J.L., Charnes, J.M. (Eds.), *Proceedings of the 2002 Winter Simulation Conference*. IEEE Press, Piscataway, NJ, pp. 1481–1492.
Tsitsiklis, J., Van Roy, B. (2001). Regression methods for pricing complex American-style options. *IEEE Transactions on Neural Networks* 12 (4), 694–703.
Xu, G.L. (1990). A duality method for optimal consumption and investment under short-selling prohibition. *Ph.D. dissertation*, Carnegie Mellon University.