15 MARKOV DECISION PROCESSES IN FINANCE AND DYNAMIC OPTIONS

Manfred Schäl

Dedicated to Professor Dr. Dr.h.c. Karl Hinderer
on the occasion of his seventieth birthday

Abstract: In this paper a discrete-time Markovian model for a financial market is chosen. The fundamental theorem of asset pricing relates the existence of a martingale measure to the no-arbitrage condition. It is explained how to prove the theorem by stochastic dynamic programming via portfolio optimization. The approach singles out certain martingale measures with additional interesting properties. Furthermore, it is shown how to use dynamic programming to study the smallest initial wealth x^* that allows for super-hedging a contingent claim by some dynamic portfolio. There, a joint property of the set of policies in a Markov decision model and the set of martingale measures is exploited. The approach extends to dynamic options which are introduced here and are generalizations of American options.

15.1 INTRODUCTION AND SUMMARY

In discrete time $n = 0, 1, \cdots, N$ a financial market is studied which is free of arbitrage opportunities but incomplete. In the market $1 + d$ assets can be traded. One of them with price process $\{B_n, 0 \le n \le N\}$ is called the bond or savings account and is assumed to be nonrisky. The other d assets are called stocks and are described by the d-dimensional price process $\{S_n, 0 \le n \le N\}$. An investor is considered whose attitude towards risk is specified in terms of a utility function U. A dynamic portfolio is specified by a policy ϕ. The investor's objective is to maximize the expected utility of the discounted terminal wealth $X_N^\phi(x)$ when starting with an initial wealth x .

Utility optimization is now a classical subject (Bertsekas [1], Hakansson [16]). The present paper is intended as a bridge between Markov decision processes (MDPs) and modern concepts of finance. It is assumed the reader knows some basic facts about MDPs, but knowledge about Mathematical Finance is not needed. In particular, we here are interested in using the no-arbitrage condition for the construction of a particular martingale measure by use of the optimal policy ϕ^*. A condition close to the no-arbitrage condition was already used by Hakansson [16]. Martingale measures are used for option pricing. The paper makes use of an approach how to base option pricing on the optimal solution ϕ^* to the portfolio optimization problem. This approach is also explained by Davis [4].

In a one-period model, the construction of a martingale measure by use of an optimization problem is easily explained. Let us assume $d = 1$. If the action (portfolio) $a \in \mathbb{R}$ is chosen then the discounted terminal wealth of the portfolio has the form $x + a \cdot R$ where R can be interpreted as a return. If $a = a^*$ is optimal then one has $\partial E[U(x+a\cdot R)]/\partial a = 0 = E[U'(x+a\cdot R)\cdot R]$ for $a = a^*$. Upon defining a new measure Q by $dQ = \text{const } \cdot U'(x+a^*\cdot R)dP$, one obtains $\int R \, dQ = E_Q[R] = 0$ which is the martingale property. Since the martingale property is a local one both in time and in space, martingale measures can be constructed by local optimization problems (see Rogers 1994) whereas we will apply global (dynamic) optimization. This has the advantage that the resulting martingale measures have interesting interpretations. The case where the utility function U is only defined for positive values is treated in [39, 40].

A second interesting problem concerns values x for the initial wealth that allow for super-hedging discounted contingent claims $\breve{X}$ by some policy ϕ, i.e. $X_N^\phi(x) \geq \breve{X}$. The smallest value x^* coincides with the maximal expectation of $\breve{X}$ under (equivalent) martingale measures. The proof can make use of dynamic programming by exploiting an analogy between the set of all policies in a stochastic dynamic programming model and the set of martingale measures. A similar problem can be considered for American options where an optimal stopping time has to be chosen. It is well-known that an optimal stopping problem can be considered as a special stochastic dynamic programming problem. Therefore it is natural from the point of view of Markov decision theory to generalize the concept of an American option. This is done in the present paper and the generalization is called a dynamic option which has interesting applications.

15.2 THE FINANCIAL MARKET

On the market an investor can observe the prices of $1+d$ securities at the dates $n = 0, 1, \cdots, N$ where N is the *time horizon*. One of the securities is a *bond* (or savings account) with *interest rates* r_n, $1 \leq n \leq N$. It is essential for the theory that the interest rates for borrowing and lending are assumed to be the same. The *bond price process* is defined by

$$B_n := (1+r_1)\cdots(1+r_n), 0 \leq n \leq N, \text{ where } B_0 = 1 \,. \tag{15.1}$$

Here we assume that $\{B_n\}$ is a deterministic process. If $\{B_n\}$ is given as the initial term structure, then the interest rates r_n can be computed by (15.1).