EE 736 Project Report

Krutika Patil 12D100005

This report is based on the paper “MDP algorithms for portfolio optimization problems in pure jump markets” by Bäuerle, Nicole, and Ulrich Rieder. It was published in Finance and Stochastics 13.4 (2009)

Introduction

The term portfolio refers to the collection of assets such as stocks, bonds held by an investor. Portfolio optimization is the process of determining the composition a portfolio that meets a given objective such as maximize return for a given risk. The objective addressed in this paper is to maximize the expected utility of the terminal wealth of a portfolio.

Modelling the stock prices: The market is a continuous time pure jump market, i.e. jumps in the stock prices arrive according to a poisson process and the relative jump heights of the stock are independent identically distributed random vectors. The stock prices are deterministic between the jumps. Hence the price process falls in the class of Piecewise Deterministic Markov Processes (PDMP). Also some model extensions are discussed like regime switching of model parameters and additional risk constraints.

The optimization problem is formulated as a discrete-time Markov Decision Process. This implies interesting results like : 1) The characterization of the value function as the fixed point of a dynamic programming operator. The operator can be shown to be contracting which establishes the uniqueness of the fixed point. 2) Because of the contraction property of the operator, the convergence of numerical schemes can be shown 3) The existence of an optimal portfolio strategy 4) The use of value iteration or policy improvement to solve the problem rather than finding the solution to the Hamilton-Jacobi-Bellman equation. It was shown that howard's policy improvement algorithm is valid.

Model

Asset dynamics follow a Piecewise Deterministic Markov Process.

$T > 0$ is the time horizon of the model and we there is a probability space (Ω, F, P), and on this space there is a homogeneous càdlàg Poisson process $N = (N_t)$ with constant rate $\lambda > 0$ and a sequence of independent and identically distributed random vectors (Y_n) with values in $(-1, \infty)^d$. The Y_n are assumed to have a distribution Q and to be independent of (N_t). Thus, we can define the $\mathbb{R}^d$-valued compound Poisson process

$$C_t := \sum_{n=1}^{N_t} Y_n. \tag{1}$$

(F_t) is the filtration generated by $C = (C_t)$. (C_t^k) denotes the k^{th} component of this process. It is supposed that we have d risky assets and one riskless bond with the following dynamics for $t \in [0, T]$: The price process (S_t^0) of the riskless bond is given by $S_t^0 := e^{\rho t}$, where $\rho \geq 0$ denotes the fixed continuous interest rate. The price processes (S_t^k) of the risky assets $k = 1, ..., d$ satisfy the stochastic differential equations

$$dS_t^k = S_t^k - \mu_k dt + dCt^k \tag{2}$$

where $\mu_k \in R$ are given constants. The initial prices S_0^k are assumed to be fixed and strictly positive . $T_0 = 0$ and $T_0 < T_1 < T_2 < ...$ are the jump time points of the Poisson process. If $t \in [T_n, T_{n+1})$, then for $k = 1, ..., d$,

$$S_t^k = S_{T_n}^k exp\left(\mu_k(t - T_n)\right) \tag{3}$$

At the time of a jump we have

$$S_{T^n}^k - S_{T^{n-}}^k = S_{T^n}^k - Y_n^k. \tag{4}$$

Thus, Y_n^k gives the relative jump height of asset k at the n^{th} jump Π denotes the set of all admissible portfolio strategies. The dynamics of the wealth process is given by

$$dX_t^\pi = X_{t-}^\pi(\rho dt + \pi_t(\mu - \rho e)dt + \pi_t dC_t) \tag{5}$$

The wealth process is again a PDMP. We obtain for the wealth process the explicit expression

$$X_t^\pi = x_0 \; exp\left(\int_0^t (\rho + \pi_s.(\mu - \rho e))\, ds\right) \prod_{j=1}^{N_t} (1 + \pi_{T_j}.Y_j). \tag{6}$$

The aim of the investor is to maximize the expected utility of the terminal wealth. We denote by $U : (0, \infty) \to \mathbb{R}^+$ an increasing, concave utility function and define for $\pi \in \Pi$, $t \in [0, T]$, $x > 0$

$$V_\pi(t, x) := \mathbb{E}_{t,x} U(X_T^\pi), \tag{7}$$

the expected utility when we start at time t with wealth x and use the portfolio strategy π. The maximal utility when we start at time t with wealth x is given by

$$V(t, x) := sup_{\pi \in \Pi} V_\pi(t, x). \tag{8}$$

Formulation of the Discrete MDP

It can be shown that the value function V can be obtained as the value function of a suitably defined discrete-time Markov decision process which is obtained by looking at the embedded jump process. V can then be identified as the unique fixed point of the contracting dynamic programming operator . The optimal portfolio is the maximizer of V. The MDP is formulated as follows:

1. The state space $E := [0, T] \times (0, \infty)$ with the Borel σ -algebra $\mathcal{B}(E)$ A state (t, x) represents the jump time point and the wealth of the process directly after the jump. The embedded Markov Chain is denoted by $Z_n = (T_n, X_{T_n})$. Z_n is fixed and is defined by $\Delta \notin E$ whenever $T_n > T$ since the process upto time T is of interest.

2. The action space of the MDP is given by $A = \alpha : [0, T] \to U \; measurable$ Where$\alpha(t) = \alpha_t$ denotes the fractions of wealth invested in stocks t time units after the last jump. For$\alpha \in A$ the movement of wealth between jumps is defined by

$$\phi_t^\alpha := x \; exp(\int_0^t (\rho + \alpha_s.(\mu - \rho e))ds) \tag{9}$$

   where ϕ_t^α is the wealth t units after the last jump, when the state directly after the jump was x.

3. For $B \in \mathcal{B}(E)$ The transition probability q is given by

$$q(B|(t, x), \alpha) := \lambda \int_0^{T-t} e^{-\lambda s} \int 1_B(t + s, \phi_s^\alpha(x)(1 + \alpha_s.y))Q(dy)ds \tag{10}$$

$$q(\Delta|(t, x), \alpha) := e^{-\lambda(T-t)} = 1 - q(E|(t, x), \alpha) \tag{11}$$

$$q(\Delta|\Delta, \alpha) := 1 \tag{12}$$

4. The reward function is defined as a function of the state $(t, x) \in E$ and action $\alpha \in A$ by

$$r((t, x), \alpha) := e^{-\lambda(T-t)} U(\phi_{T-t}^\alpha) \tag{13}$$

$$r(\Delta, \alpha) := 0 \tag{14}$$

A measurable function f called the decision rule determines the action to be performed until the time of next jump as a function of the state directly after the jump $f \; : E \cup \Delta \; \to \; A$. F denotes the set of all decision rules. A Markov Policy is a sequence of decision rules (f_n) such that $f_n \in F$.
The expected reward of a Markov Policy is given by

$$J_{(f_n)}(t, x) := \mathbb{E}_{t,x}^{(f_n)} \left[\sum_{k=0}^\infty r(Z_k, f_k(Z_k))\right], \quad (t, x) \in E \tag{15}$$

The value function $J(t, x)$ is the supremum of $J_{(f_n)}(t, x)$ over all decision rules.
The Markovian portfolio strategy π_t is given by

$$\pi_t = f_n(T_n, X_{T_n}^\pi) \; (t - Tn) \quad for \; t \in [T_n, T_{n+1}] \tag{16}$$

for a sequence f_n with $f_n \in F$.

Solution of the Optimization problem

For $\pi \in \Pi$, $Z_n^\pi := (T_n, X_{T_n}^\pi)$ and $Z_n^\pi := \Delta$ when $T_n > T$. It was shown that the continuous time optimization problem can be solved by the discrete-time MDP. The following Theorem was proved in the paper. The proof is skipped here.

Theorem 1

a) For any Markovian portfolio strategy $\pi = (\pi_t)$, the expected reward of the Markov Policy (f_n) is equal to the expected utility when we start at time t with wealth x.

$$V_\pi(t,x) = J_{(f_n)}(t,x) \quad (t,x) \in E \tag{17}$$

b) The value function of the MDP $J(t,x)$ is equal to the maximal utility when we start at time t with wealth x.

$$V(t,x) = J(t,x) \;\; where \; (t,x) \in \; E \tag{18}$$

Existence of optimal strategies and characterization of the value function

Consider the following definition

$$\mathbb{M} := \{\nu \; : E \to \mathbb{R}_+ \mid \quad \nu \; is \; measurable \; \} \tag{19}$$

Operators L, T_f, T acting on $\mathbb{M}$ are defined as

$$L_\nu \; (t,x,\alpha) := \; e^{-\lambda(T-t)}U(\phi^\alpha_{T-t}(x)) + \int \nu(s,y)q(ds,dy \mid t,x,\alpha), \tag{20}$$

$$T_f \; \nu(t,x) := \; L_\nu \; (t,x,f(t,x)), \quad (t,x) \; \in E, \tag{21}$$

$$T \; \nu(t,x) := \; \sup_{\alpha \in A} \; L_\nu \; (t,x,\alpha), \tag{22}$$

From MDP Theory:

$$J_f = \lim_{n\to\infty} T_f \; ^n 0$$

and $J_f = T_f \;\; J_f$. It is shown that J_f is the unique fixed point of T_f . And it is possible to define a special norm on a subset of the function space $\mathbb{M}$ such that the operator T is contracting and has a unique fixed point and the fixed point is equal to the value function $V(t,x)$. The norm is a weighted supremum norm with a bounding function.

Bounding function for the MDP:

$$b(t,x) := e^{\beta(T-t)}(1+x) \quad (t,x) \; \in E \quad for \beta \geq 0 \tag{23}$$

Then there exist c, $c_\beta \in \mathbb{R}_+$ such that

1. $r(t,x,\alpha) \leq c \; b(t,x)$
2. For all $(t,x,\alpha) \in E \times A$

$$\int b(s,y)q(ds,dy \mid t,x,\alpha) \leq c_\beta \; b(t,x) \tag{24}$$

where

$$c_\beta := \frac{\lambda(1+\bar{y})}{\beta+\lambda-\bar{\mu}}(1-e^{-T(\beta+\lambda-\bar{\mu})}) \tag{25}$$

where $\overline{\mu}$ is the max appreciation rate of assets and $\overline{y}$ is the max expected relative jump height.

$b(t,x)$ is the bounding function of the MDP. For large β, $c_\beta < 1$. c_β is the contraction module of T and T_f

Definition of the norm:
The weighted supremum norm $||.||_b$ on $\mathbb{M}$ is defined as

$$||\nu||_b := \sup_{(t,x)\in E} \frac{\nu(t,x)}{b(t,x)} \tag{26}$$

also

$$\mathbb{B}_b := \{\nu \in \mathbb{M} \mid ||\nu||_b < \infty\} \tag{27}$$

Convergence in $||.||_b$ is equivalent to convergence on compact sets.
The set $(M)_c$ is defined as

$$\mathbb{M}_c := \{\nu \in \mathbb{B}_b \mid \nu \text{ is continuous, } \nu(t,x) \text{ is concave and increasing in } x, \text{ decreasing in } t\}$$

For $\nu \in \mathbb{M}_c,\ T_{\ \nu}(t,x) = \sup_{\alpha\in A} L_\nu\ (t,x,\alpha)$.
For $\nu,\ w\ \in \mathbb{B}_b$ and $f \in F$

$$||T_{f\ \ \nu} - T_{f\ \ w}||_b \leq c_\beta ||\nu - w||_b \tag{28}$$

$$||T_{\ \nu} - T_{\ w}||_b \leq c_\beta ||\nu - w||_b \tag{29}$$

T_f , T are contracting if $c_\beta < 1$. The proof is skipped here.
This means that J_f is the unique fixed point of T_f in $\mathbb{B}_b$. But $V(t,x) = J(t,x)$ for all $(t,x) \in E$. Hence V is the unique fixed point of the operator T in $\mathbb{M}_c$. It is proved that there exists an optimal portfolio strategy $\pi \in \Pi$ such that

$$\pi_t = f(Z_n^\pi)(t - Tn) \quad for \quad t \in [T_n, T_{n+1}) \tag{30}$$

Computational methods

The fixed point equation needs to be solved to find the value function. Let

$$l_\nu(t,x,u) := \lambda \int \nu(t, x(1 + u.y))Q(dy) \tag{31}$$

where Q is the relative jump height distribution and $\nu \in \mathbb{M}_c$.
Then the fixed point equation is written as

$$V(t,x) = \sup_{\alpha\in A} \left(e^{\lambda(T-t)} U(\phi_{T-1}^\alpha(x)) + \int_0^{T-1} e^{-\lambda s} l_\nu(t+s, \phi_s^\alpha(x), \alpha_s) ds \right) \tag{32}$$

This is a deterministic control problem. The problem was solved explicitly for the power utility function. The solution was found to be of the form

$$V(t,x) = \frac{1}{\gamma} x^\gamma e^{\theta(T-t)} \tag{33}$$

The optimal policy was found to be investing a constant fraction u^* of the wealth in the stock.

The problem has to be solved numerically in general. It was proved that Howard's policy improvement algorithm works for the portfolio problem.

It was argued that it is possible to use an approximate utility function which is simpler, to solve the problem. It was shown that the corresponding value functions would be close to each other.

The case when the iterates of the dynamic programming operator were concentrated on a grid of the state space was also discussed. Error bound on distance to true value function was obtained.