MDP Algorithms for Portfolio Optimization Problems in pure Jump Markets

Nicole Bäuerle
(joint work with U. Rieder)

Universität Karlsruhe (TH)
Research University • founded 1825

Linz, October 2008

Overview

- **Model and optimization problem**
- **Solution via discrete-time MDPs**
- **Solution method and computational aspects**
- **Model extensions**

Model and optimization problem

The market model

Suppose we have a financial market with one bond and d risky assets whose prices evolve as follows

- Price process (S_t^0) of the bond:

$$S_t^0 := e^{rt}, \quad r \geq 0.$$

- Price processes (S_t^k) of the risky assets $k = 1, \ldots, d$:

$$dS_t^k = S_{t-}^k(\mu_k dt + dC_t^k)$$

where $\mu_k \in \mathbb{R}$ are constants and $C_t := \sum_{n=1}^{N_t} Y_n$ with $N = (N_t)$ a Poisson process with rate $\lambda > 0$ and iid random vectors (Y_n) with values in $(-1, \infty)^d$ and distribution Q where $\mathbb{E}\|Y_n\| < \infty$.

The price processes of the risky assets

If we denote by $0 := T_0 < T_1 < T_2 < \dots$ the jump time points of the Poisson process and if $t \in [T_n, T_{n+1})$, then for $k = 1, \dots, d$

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

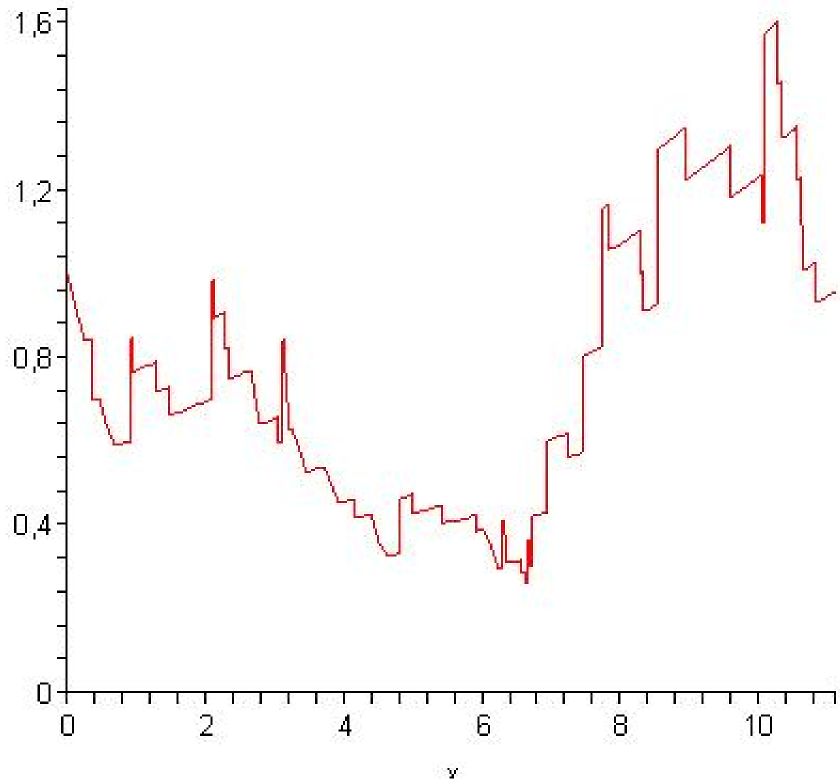
At the time of a jump we have

$$S_{T_n}^k - S_{T_n-}^k = S_{T_n-}^k Y_n^k.$$

In what follows we denote $S_t := (S_t^1, \dots, S_t^d)$.

(S_t) is a so-called *Piecewise Deterministic Markov Process* (PDMP).

A typical sample path



Simulated stock price in the PDMP model.

Portfolios and self-financing strategies

Portfolio strategy: $(\mathfrak{F}_t)$**-predictable stochastic process** (π_t) **with values in**

$$\mathcal{U} := \{u \in \mathbb{R}^d \mid u \geq 0, u \cdot e \leq 1\}$$

where $\pi_t = (\pi_t^1, \ldots, \pi_t^d)$ **gives the fractions of wealth invested in the risky assets at time** t**.** $1 - \pi_t \cdot e$ **is the fraction invested in the bond.**

The equation for the wealth process (X_t^π) **is then:**

$$dX_t^\pi = X_t^\pi \Big(r + \pi_t \cdot (\mu - re)dt + \pi_t dC_t \Big).$$

The optimization problem

Let $U : (0, \infty) \to \mathbb{R}_+$ be an increasing, concave utility function and define for a portfolio strategy π and $t \in [0, T], x > 0$:

$$\begin{aligned} V_\pi(t, x) &:= \mathbb{E}_{t,x} U(X_T^\pi). \\ V(t, x) &:= \sup_\pi V_\pi(t, x). \end{aligned}$$

Obviously

$$V_\pi(T, x) := U(x) = V(T, x).$$

Davis (1993), Norberg (2003), Schäl (2004,2005), Kirch and Runggaldier (2005), Jacobsen (2006), B. and Rieder (2008)

Solution via discrete-time MDPs

Solution via discrete-time MDPs: The model

- **State space**: $E = [0,T] \times (0,\infty)$. A state (t,x) gives the jump time point t and the wealth x of the process directly after the jump.

- **Action space**: $A := \{\alpha : [0,T] \to \mathcal{U} \text{ measurable}\}$. For $\alpha \in A$ we define the movement of the wealth between jumps by

$$\phi_t^{\alpha}(x) := x \exp\left(\int_0^t r + \alpha_s \cdot (\mu - re) ds\right).$$

- **Transition probability**:

$$q\big(B \mid t,x,\alpha\big) := \lambda \int_0^{T-t} e^{-\lambda s} \int 1_B\Big(t+s, \phi_s^{\alpha}(x)\big(1+\alpha_s \cdot y\big)\Big) Q(dy) ds$$

- **One-stage reward function**: $r\big(t,x,\alpha\big) := e^{-\lambda(T-t)} U\big(\phi_{T-t}^{\alpha}(x)\big)$.

Solution via discrete-time MDPs

A sequence (f_n) with $f_n \in F := \{f : E \to A$ measurable$\}$ is called $Markov\ policy$. The expected reward of such a Markov policy is given by

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

Define $J(t,x) := \sup_{(f_n)} J_{(f_n)}(t,x)$, for $(t,x) \in E$.

Theorem 1: We have $V(t,x) = J(t,x)$, for $(t,x) \in E$ and the optimal portfolio strategies "coincide".

Some important operators

We define the following operators $L, \mathcal{T}_f, \mathcal{T}$ which act on $\mathbb{M} := \{v : E \to \mathbb{R}_+ \mid v \text{ is measurable}\}$:

$$\begin{aligned}
(Lv)(t,x,\alpha) &:= e^{-\lambda(T-t)}U\left(\phi^{\alpha}_{T-t}(x)\right) + \int v(s,y)q(ds,dy \mid t,x,\alpha). \\
(\mathcal{T}_f v)(t,x) &:= Lv(t,x,f(t,x)),\ (t,x) \in E,\ f \in F. \\
(\mathcal{T}v)(t,x) &= \sup_{\alpha \in A} Lv(t,x,\alpha).
\end{aligned}$$

From MDP theory it follows that

$$\begin{aligned}
J_{(f_n)} &= \lim_{n\to\infty} \mathcal{T}_{f_0} \dots \mathcal{T}_{f_{n-1}} 0 \\
J_f := J_{(f)} &= \lim_{n\to\infty} \mathcal{T}_f^n 0 \\
J_f &= \mathcal{T}_f J_f.
\end{aligned}$$

A peculiar norm

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

Next we introduce the weighted supremum norm $\|\cdot\|_b$ on $I\!M$ by

$$\|v\|_b := \sup_{(t,x)\in E} \frac{v(t,x)}{b(t,x)} \quad \text{and} \quad I\!B_b := \{v \in I\!M \mid \|v\|_b < \infty\}.$$

Finally, we define the set

$$I\!M_c := \{v \in I\!B_b \mid v \text{ is continuous, } v(t,x) \text{ is concave and}$$

$$\text{increasing in } x, \text{ decreasing in } t\}.$$

Properties of $\mathcal{T}$ and $\mathcal{T}_f$

Theorem 2: It holds that

a) $\mathcal{T} : \mathbb{M}_c \to \mathbb{M}_c$.

b) For $v, w \in \mathbb{B}_b$ and $f \in F$ we have

$$\begin{aligned} \|\mathcal{T}_f v - \mathcal{T}_f w\|_b &\leq c_\beta \|v - w\|_b \\ \|\mathcal{T} v - \mathcal{T} w\|_b &\leq c_\beta \|v - w\|_b. \end{aligned}$$

with $c_\beta := \frac{\lambda(1+\bar{y})}{\beta+\lambda-\bar{\mu}}\left(1 - e^{-T(\beta+\lambda-\bar{\mu})}\right)$. Thus the operators $\mathcal{T}_f, \mathcal{T}$ are contracting if $c_\beta < 1$ which is the case if β is large enough.

Characterization of the value function

Theorem 3:

a) The value function V is the unique fixed point of $\mathcal{T}$ in $I\!M_c$.

b) For $J_0 \in I\!M_c$ it holds that

$$\|V - \mathcal{T}^n J_0\|_b \leq \frac{c_\beta^n}{1 - c_\beta} \|\mathcal{T} J_0 - J_0\|_b.$$

c) There exists an optimal stationary portfolio strategy π, i.e. there exists an $f \in F$ such that

$$\pi_t = f(T_n, X_n)(t - T_n) \text{ for } t \in [T_n, T_{n+1}).$$

Solution method and computational aspects

Howard's policy improvement algorithm

Theorem 4: Howard's policy improvement algorithm works for the portfolio problem. More precisely, let $f, g \in F$.

a) If for some subset $E_0 \subset E$

$$g(t,x) \in D(t,x,f) := \{\alpha \in A \mid LJ_f(t,x,\alpha) > J_f(t,x)\}, \qquad (t,x) \in E_0$$
$$g(t,x) = f(t,x), \qquad (t,x) \notin E_0$$

then $J_g \geq J_f$ and $J_g(t,x) > J_f(t,x)$ for $(t,x) \in E_0$.

b) If $D(t,x,f) = \emptyset$ for all $(t,x) \in E$ then $J_f = V$, i.e. the decision rule f defines the optimal stationary portfolio strategy.

Application: Suppose $r = \mu_i$ and U is continuously differentiable and $U'(x + u \cdot Y)Y$ is integrable for all $x > 0$ and $\|u\|$ small. Then "invest all the money in the bond" is optimal if and only if $\mathbb{E}Y \leq 0$.

Approximating the utility function

Let $U^{(n)}, U$ be utility functions and denote by $V^{(n)}, V$ the corresponding value functions and by

$$A_n^*(t,x) := \{\alpha \in A : \mathcal{T}^{(n)} V^{(n)}(t,x) = L^{(n)} V^{(n)}(t,x,\alpha)\}$$

$$A^*(t,x) := \{\alpha \in A : \mathcal{T} V(t,x) = LV(t,x,\alpha)\}.$$

Moreover, let us denote by

$$LsA_n^*(t,x) := \{\alpha \in A;\ \alpha \text{ is an accumulation point of a sequence } (\alpha_n) \text{ with } \alpha_n \in A_n^*(t,x)\ \forall n \in \mathbb{N}\}$$

the upper limit of the set sequence $\left(A_n^*(t,x)\right)$.

Jouini and Napp (2004), Schäl (1975)

Approximating the utility function

Theorem 5:

a) If U and $\tilde{U}$ are two utility functions with corresponding value functions V and $\tilde{V}$, then

$$\|V - \tilde{V}\|_b \leq \|U - \tilde{U}\|_b \frac{e^{T\bar{\mu}}}{1 - c_\beta}.$$

b) Let $\left(U^{(n)}\right)$ be a sequence of utility functions with $\lim_{n\to\infty} \|U^{(n)} - U\|_b = 0$. Then $\lim_{n\to\infty} \|V^{(n)} - V\|_b = 0$ and we get that

$$\emptyset \neq LsA_n^*(t,x) \subset A^*(t,x)$$

for all $(t,x) \in E$, i.e. in particular, the limit f^* of a sequence of decision rules (f_n^*) with $f_n^*(t,x) \in A_n^*(t,x)$ for all $(t,x) \in E$ defines an optimal portfolio strategy for the model with utility function U.

State space discretization

Choose a grid $G \subset E$ and define the grid operator T_G on $I\!M_c$ by

$$\mathcal{T}_G v(t,x) := \begin{cases} \mathcal{T} v(t,x), & \text{for } (t,x) \in G \\ \text{linear interpolation}, & \text{else} \end{cases}$$

and $b_G : E \to I\!R_+$ by a linear interpolation of b. For $v \in I\!M_c$:

$$\|v\|_G := \sup_{(t,x)\in E} \frac{v(t,x)}{b_G(t,x)} \leq \|v\|_b.$$

Theorem 6: Suppose that $c_G < 1$. Then it holds for $J_0 \in I\!M_c$ that

$$\|V - \mathcal{T}_G^n J_0\|_G \leq \frac{1}{1-c_G}\Big(c_G^n \|\mathcal{T}_G J_0 - J_0\|_G + m(h)\Big)$$

where $m(h) := \|V - \mathcal{T}_G V\|_G \to 0$ if the mesh size tends to zero.

Numerical example

One stock ($d = 1$): $\mu = r = 0$.

Relative jump distribution:

$$Q(dy) = \begin{cases} p\lambda_+ e^{-\lambda_+ y} dy & , y \geq 0 \\ (1-p)\lambda_-(y+1)^{\lambda_- - 1} & , -1 < y < 0 \end{cases}$$

$$p = 0.5, \quad \lambda_+ = \lambda_- = 1.$$

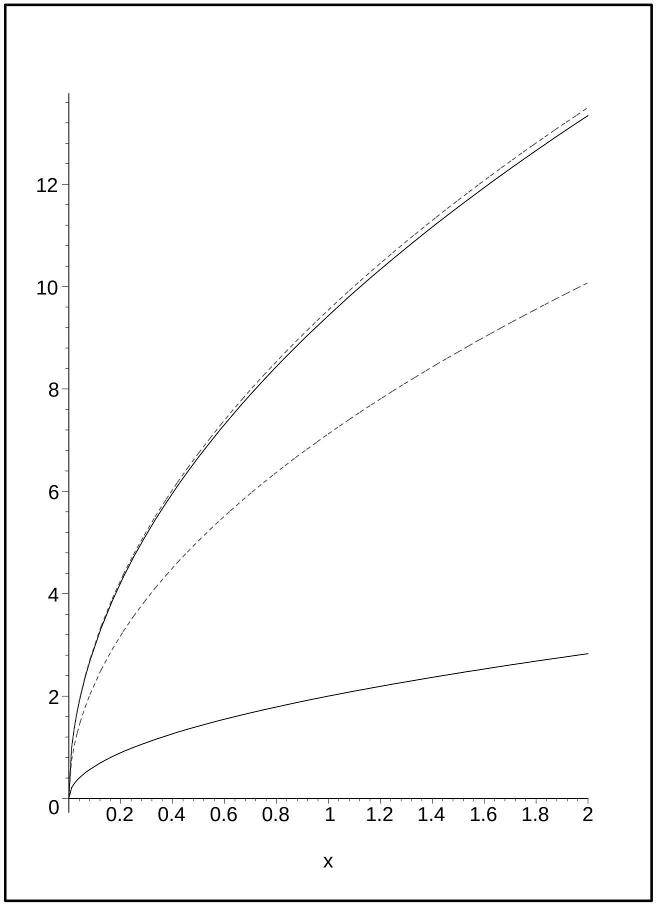
Utility function

$$U(x) = \frac{1}{\gamma} x^\gamma, \quad \gamma = 0.5.$$

Horizon $T = 1$

Mesh size: $h = 0.01$

Numerical example

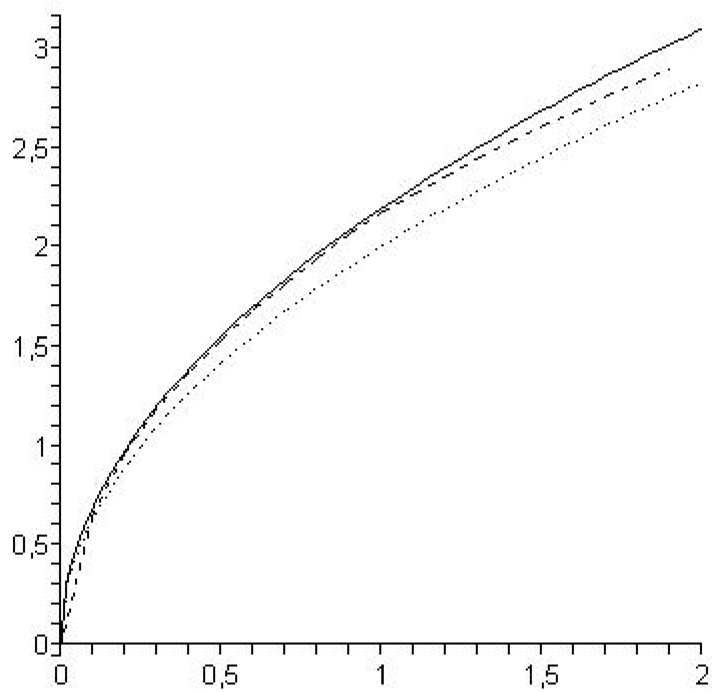


Upper solid line: $V(0, x)$.
Upper dotted line: $J_1(0, x)$.
Lower solid line: $V(1, x) = U(x)$.

Model extensions

Model extensions: Dynamic risk constraints

Next we impose an additional dynamic risk constraint of the form
$\mathcal{U}(x) = \{u \in \mathcal{U} \mid \rho\left(X^{\pi}_{T_1} - x\right) = ux\rho(Y) \leq 0.1\}$ where ρ is AVaR.



Solid line: Value function without constraints. Dotted line: Utility function. Dashed line: First iteration of the grid operator with risk constraints.

Model extensions: Partial information

Suppose we do not know the jump intensity $\Lambda \in \{\lambda_1, \ldots, \lambda_m\}$.

Define $p_k(t) := I\!\!P\Big(\Lambda = \lambda_k \mid \mathcal{F}_t^S\Big)$, $\quad \hat{\lambda}_t := \sum \lambda_k p_k(t) = I\!\!E\Big[\Lambda \mid \mathcal{F}_t^S\Big]$.

Filter equation:

$$p_k(t) \quad = \quad p_k(0) + \int_0^t p_k(s-)\Big(\frac{\lambda_k - \hat{\lambda}_{s-}}{\hat{\lambda}_{s-}}\Big) d\hat{\eta}_s$$

where

$$\hat{\eta}_t = N_t - \int_0^t \hat{\lambda}_s ds.$$

The filter becomes part of the state space, i.e. the value function is of the form $V(t, x, p)$. The optimization problem can be solved as before by reduction to a discrete time MDP.

Comparison: Full - Partial Information

Let $U(x) = \frac{1}{\gamma}x^\gamma$ with $0 < \gamma < 1$, $d = 1$ and $Q = \delta_{y_0}$ with $y_0 < 0$.

Here, the optimal fractions of wealth invested in the stock $u^*_{\boldsymbol{\lambda}}$ (full information), $u^*(t,p)$ (partial information) do not depend on the wealth itself. Let $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \ldots \boldsymbol{\lambda}_m)$.

Theorem 7:

The optimal fraction $u^*(t,p)$ invested in the stock has the following property for all $(t,p) \in [0,T] \times \mathcal{P}$:

$$u^*_{\boldsymbol{\lambda}' p} \leq u^*(t,p).$$

Bäuerle and Rieder (2007)

References

[1] Bäuerle, N., Rieder, U. (2008) : MDP algorithms for portfolio optimization problems in pure jump markets. Preprint.

[2] Davis, M.H.A. (1993) : Markov models and optimization. Chapman & Hall.

[3] Jacobsen, M. (2006) : Point process theory and applications. Birkhäuser.

[4] Kirch, M., Runggaldier, W. (2005) : Efficient hedging when asset prices follow a geometric Poisson process with unknown intensity . SIAM J. Control Optim., 43:1174-1195.

[5] Norberg, R. (2003) : The Markov chain market. Astin Bulletin, 33:265-287.

[6] Schäl, M. (2005) : Control of ruin probabilities by discrete-time investments, Math. Methods Oper. Res., 62:141–158.

Thank you for your attention!