Chapter 6
Partially Observable Markov Decision Problems in Finance

All the models which have been considered in Chapter 4 may also be treated in the case of partial information. Indeed models of this type occur somehow natural in mathematical finance because there are underlying economic factors influencing asset prices which are not specified and cannot be observed. Moreover, for example the drift of a stock is notoriously difficult to estimate. In this chapter we assume that the relative risk return distribution of the stocks is determined up to an unknown parameter which may change. This concept can also be interpreted as one way of dealing with model ambiguity. We choose two of the models from Chapter 4 and extend them to partial observation. The first is the general terminal wealth problem of Section 4.2 and the second is the dynamic mean-variance problem of Section 4.6.

We consider a financial market with one riskless bond (with interest rate $i_n = i$) and d risky assets with relative risk process (R_n). Here we assume that the distribution of R_{n+1} depends on an underlying stationary Markov process (Y_n) which cannot be observed. In Section 4.4 (Regime-switching model) the process (Y_n) is observable. The state space of (Y_n) is E_Y, a Borel subset of a Polish space. We assume that (R_n, Y_n) is a Markov process and moreover

$$\begin{aligned} &\mathbb{P}\big(R_{n+1} \in B, Y_{n+1} \in C | Y_n = y, R_n\big) \\ &= \mathbb{P}\big(R_{n+1} \in B, Y_{n+1} \in C | Y_n = y\big) \qquad (6.1) \\ &= \mathbb{P}\big(R_{n+1} \in B | Y_n = y\big) \cdot \mathbb{P}\big(Y_{n+1} \in C | Y_n = y\big) =: Q^R(B|y) Q^Y(C|y) \end{aligned}$$

for $B \in \mathcal{B}(\mathbb{R}^d), C \in \mathcal{B}(E_Y)$. Q^Y is the transition kernel of the 'hidden' Markov process (Y_n) and $Q^R(\cdot|y)$ is the (conditional) distribution of R_{n+1} given $Y_n = y$ (independent of n). In the following let $R(y)$ be a random variable with distribution $Q^R(\cdot|y)$, i.e. $\mathbb{P}(R(y) \in B) = Q^R(B|y) = \mathbb{P}(R_{n+1} \in B | Y_n = y)$. Given (Y_n), the random vectors $R_1, R_2, \dots$ are independent, and given Y_n, the random variables R_{n+1} and Y_{n+1} are independent.

N. Bäuerle and U. Rieder, *Markov Decision Processes with Applications to Finance*, Universitext, DOI 10.1007/978-3-642-18324-9_6,

6.1 Terminal Wealth Problems

We start with problems of terminal wealth maximization under partial observation. Suppose we have an investor with utility function $U : dom\, U \to \mathbb{R}$ with $dom\, U = [0, \infty)$ or $dom\, U = (0, \infty)$ and initial wealth $x > 0$. Our investor can only observe the stock price and not the driving Markov process (Y_n), i.e. the filtration $(\mathcal{F}_n)$ to which portfolio strategies have to be adapted is given by $\mathcal{F}_n := \mathcal{F}_n^S = \sigma(S_0, S_1, \dots, S_n) = \sigma(R_1, \dots, R_n)$. The aim is to maximize the expected utility of her terminal wealth.
The following assumption on the financial market is used throughout this section.

Assumption (FM):

(i) *There are no arbitrage opportunities in the market, i.e. for all $y \in E_Y$ and $\phi \in \mathbb{R}^d$ it holds:*

$$\phi \cdot R(y) \geq 0 \;\; \mathbb{P}\text{-}a.s. \quad \Rightarrow \quad \phi \cdot R(y) = 0 \;\; \mathbb{P}\text{-}a.s.$$

(ii) *The support of $R(y)$ is independent of $y \in E_Y$.*
(iii) $\sup_y \mathbb{E}\, \|R(y)\| < \infty$.

The second assumption guarantees that the support of R_{n+1} is independent of Y_n and n. There are a lot of examples where this assumption is satisfied. According to (3.1) the wealth process (X_n) evolves as follows

$$X_{n+1} = \big(1 + i\big)\Big(X_n + \phi_n \cdot R_{n+1}\Big)$$

where $\phi = (\phi_n)$ is a portfolio strategy such that ϕ is $(\mathcal{F}_n)$-adapted. The partially observable terminal wealth problem is then given by

$$\begin{cases} \mathbb{E}_x\, U(X_N^\phi) \to \max \\ \\ \phi \text{ is a portfolio strategy and } X_N^\phi \in dom\, U \;\; \mathbb{P}\text{-a.s.}. \end{cases} \tag{6.2}$$

Recall that $\mathcal{F}_n = \mathcal{F}_n^S$ i.e. the admissible portfolio strategies depend only on the observable stock prices (S_n). Problem (6.2) can be solved by the following stationary Partially Observable Markov Decision Model:

- $E_X := dom\, U$ where $x \in E_X$ denotes the wealth,
- E_Y is the state space of (Y_n), where $y \in E_Y$ is the unobservable state,
- $A := \mathbb{R}^d$ where $a \in A$ is the amount of money invested in the risky assets,
- $D(x) := \Big\{a \in \mathbb{R}^d \mid \big(1+i\big)\big(x + a \cdot R(y)\big) \in dom\, U \;\; \mathbb{P}\text{-a.s.}\Big\}$,
- $\mathcal{Z} := [-1, \infty)^d$ where $z \in \mathcal{Z}$ denotes the relative risk,
- $T_X\big(x, a, z\big) := (1+i)\big(x + a \cdot z\big)$,
- $Q^{Z,Y}(B \times C | x, y, a) := Q^R(B|y) Q^Y(C|y)$ for $B \in \mathcal{B}(\mathcal{Z}), C \in \mathcal{B}(E_Y)$,