2 Introduction to Stochastic Control

2.1 A motivating example from Merton's problem

In this part we give a motivating example to introduce the problem of dynamic asset allocation and stochastic optimization. We will not be particularly rigorous in these calculations.

The market Consider an investor can invest in a two asset Black-Scholes market: a risk-free asset ("bank" or "Bond") with rate of return $r > 0$ and a risky asset ("stock") with mean rate of return $\mu > r$ and constant volatility $\sigma > 0$. Suppose that the price of the risk-free asset at time t, B_t, satisfies

$$\frac{dB_t}{B_t} = r\,dt \quad \text{or} \quad B_t = B_0 e^{rt}, \qquad t \geq 0.$$

The price of the stock evolves according to the following SDE:

$$\frac{dS_t}{S_t} = \mu\,dt + \sigma\,dW_t,$$

where $(W_t)_{t\geq 0}$ is a standard one-dimensional Brownian motion one the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

The agent's wealth process and investments Let X_t^0 denote the investor's wealth in the bank at time $t \geq 0$. Let π_t denote the wealth in the risky asset. Let $X_t = X_t^0 + \pi_t$ be the investor's total wealth. The investor has some initial capital $X_0 = x > 0$ to invest. Moreover, we also assume that the investor saves / consumes wealth at rate C_t at time $t \geq 0$.

There are three popular possibilities to describe the investment in the risky asset:

(i) Let ξ_t denote the number of units stocks held at time t (allow to be fractional and negative),

(ii) the value in units of currency $\pi_t = \xi_t S_t$ invested in the risky asset at time t,

(iii) the fraction $\nu_t = \frac{\pi_t}{X_t}$ of current wealth invested in the risky asset at time t.

The investment in the bond is then determined by the accounting identity $X_t^0 = X_t - \pi_t$. The parametrizations are equivalent as long as we consider *only* positive wealth processes (which we shall do). The gains/losses from the investment in the stock are then given by

$$\xi_t\,dS_t, \qquad \frac{\pi_t}{S_t}\,dS_t, \qquad \frac{X_t \nu_t}{S_t}\,dS_t\,.$$

The last two ways to describe the investment are especially convenient when the model for S is of the exponential type, as is the Black-Scholes one. Using (ii),

$$\begin{aligned} X_t &= x + \int_0^t \frac{\pi_s}{S_s}\,dS_s + \int_0^t \big(X_s - \pi_s\big)\frac{dB_s}{B_s} - \int_0^t C_s\,ds \\ &= x + \int_0^t \big[\pi_s(\mu - r) + rX_s - C_s\big]\,ds + \int_0^t \pi_s \sigma\,dW_s \end{aligned}$$

or in differential form

$$dX_t = \big[\pi_t(\mu - r) + rX_t - C_t\big]\,dt + \pi_t\sigma\,dW_t, \qquad X_0 = x\,.$$

Alternatively, using (iii), the equation simplifies even further.[1] Recall $\pi = \nu X$.

$$\begin{aligned} dX_t &= X_t\nu_t\frac{dS_t}{S_t} + X_t\big(1-\nu_t\big)\frac{dB_t}{B_t} - C_t\,dt \\ &= \big[X_t\big(\nu_t(\mu - r) + r\big) - C_t\big]\,dt + X_t\nu_t\sigma\,dW_t. \end{aligned}$$

We can make a further simplification and obtain an SDE in "geometric Brownian motion" format if we assume that the consumption C_t can be written as a fraction of the total wealth, i.e. $C_t = \kappa_t X_t$. Then

$$dX_t = X_t\big[\nu_t(\mu - r) + r - \kappa_t\big]\,dt + X_t\nu_t\sigma\,dW_t\,. \tag{2.1}$$

Exercise 2.1. Assuming that all coefficients in SDE (2.1) are integrable, solve the SDE for X and hence show $X > 0$ when $X_0 = x > 0$.

The optimization problem The investment allocation/consumption problem is to choose the best investment possible in the stock, bond and at the same time consume the wealth optimally. How to translate the words "best investment" into a mathematical criteria?

Classical modeling for describing the behavior and preferences of agents and investors are: *expected utility* criterion and *mean-variance* criterion.

In the *first criterion* relying on the theory of choice in uncertainty, the agent compares random incomes for which he knows the probability distributions. Under some conditions on the preferences, Von Neumann and Morgenstern show that they can be represented through the expectation of some function, called *utility*. Denoting it by U, the utility function of the agent, the random income X is preferred to a random income X' if $\mathbb{E}[U(X)] \geq \mathbb{E}[U(X')]$. The deterministic utility function U is nondecreasing and concave, this last feature formulating the risk aversion of the agent.

Example 2.2 (Examples of utility functions)**.** The most common utility functions are

- Exponential utility: $U(x) = -e^{\alpha x}$, the parameter $\alpha > 0$ is the risk aversion.
- Log utility: $U(x) - \log(x)$
- Power utility: $U(x) = (x^\gamma - 1)/\gamma$ for $\gamma \in (-\infty, 0) \cup (0, 1)$.
- Iso-elastic: $U(x) = x^{1-\rho}/(1-\rho)$ for $\rho \in (-\infty, 0) \cup (0, 1)$.

In this portfolio allocation context, the criterion consists of maximizing the expected utility from consumption and from terminal wealth. In the **the finite time-horizon case**: $T < \infty$, this is

$$\sup_{\nu, C} \mathbb{E}\left[\int_0^T U\big(C_t\big)\,dt + U\big(X_t^{\nu,C}\big)\right]\ , \text{ where (2.1) gives } X_t^{\nu,C} = X_t. \tag{2.2}$$

[1] Note that, if ν_t expresses the fraction of the total wealth X invested in the stock, then the fraction of wealth invested in the bank account is simply $1 - \nu_t$.

Without consumption, i.e. $\forall\, t$ we have $C(t) = 0$, the optimization problem could be written as

$$\sup_{\nu} \mathbb{E}\left[U\big(X_t^{\nu}\big)\right]\ , \ \text{ where (2.1) gives } X_t^{\nu} = X_t. \tag{2.3}$$

Note that the maximization is done under the expectation.

In the **infinite time-horizon case**: $T = \infty$. In our context the optimization problem is written as (recall that $C_t = \kappa_t X_t^{\nu,\kappa}$)

$$\sup_{\kappa,\nu} \mathbb{E}\left[\int_0^{\infty} e^{-\gamma t} U\big(\kappa_t X_t^{\nu,\kappa}\big)\, dt\,,\ \text{ with (2.1) giving } X_t = X_t^{\nu,\kappa}.\right] \tag{2.4}$$

Let us go back to the **finite horizon case**: $T < \infty$. The *second criterion* for describing the behavior and preferences of agents and investors, the mean-variance criterion, relies on the assumption that the preferences of the agent depend only on the expectation and variance of his random incomes. To formulate the feature that the agent likes wealth and is risk-averse, the mean-variance criterion focuses on mean-variance-efficient portfolios, i.e. minimizing the variance given an expectation.

In our context and assuming that there is no consumption, i.e. $\forall\, t$ we have $C_t = 0$, then the optimization problem is written as

$$\inf_{\nu} \Big\{\mathrm{Var}\big(X_T^{\nu}\big)\ :\ \mathbb{E}[X_T^{\nu}] = m, \quad m \in (0,\infty)\Big\}.$$

We shall see that this problem may be reduced to the resolution of a problem in the form (2.2) for the quadratic utility function: $U(x) = \lambda - x^2$, $\lambda \in \mathbb{R}$.

2.1.1 Basic elements of a stochastic control problem

The above investment-consumption problem and its variants (is the so-called "Merton problem" and) is an example of a stochastic optimal control problem. Several key elements, which are common to many stochastic control problems, can be seen.

These include:

Time horizon. The time horizon in the investment-consumption problem may be finite or infinite, in the latter case we take the time index to be $t \in [0,\infty)$. We will also consider problems with finite horizon: $[0,T]$ for $T \in (0,\infty)$; and indefinite horizon: $[0,\tau]$ for some stopping time τ (for example, the first exit time from a certain set).

(Controlled) State process. The state process is a stochastic process which describes the state of the physical system of interest. The state process is often given by the solution of an SDE, and if the control process appears in the SDE's coefficients it is called a *controlled stochastic differential equation*. The evolution of the state process is influenced by a control. The state process takes values in a set called the state space, which is typically a subset of $\mathbb{R}^d$. In the investment-consumption problem, the state process is the wealth process $X^{\nu,C}$ in (2.1).

Control process. The control process is a stochastic process, chosen by the "controller" to influence the state of the system. For example, the controls in the investment-consumption problem are the processes $(\nu_t)_t$ and $(C_t)_t$ (see (2.1)).

We collect all the control parameters into one process denoted $\alpha = (\nu, C)$. The control process $(\alpha_t)_{t\in[0,T]}$ takes values in an action set A. The action set can be a complete separable metric space but most commonly $A \in \mathcal{B}(\mathbb{R}^m)$.

For the control problem to be meaningful, it is clear that the choice of control must allow for the state process to exist and be determined uniquely. More generally, the control may be forced satisfy further constraints like "no short-selling" (i.e. $\pi(t) \geq 0$) and or the control space varies with time. In the financial context, the control map at time t should be decided at time t based on the available information $\mathcal{F}_t$. This translates into requiring the control process to be adapted.

Admissible controls. Typically, only controls which satisfy certain "admissibility" conditions can be considered by the controller. These conditions can be both technical, for example, integrability or smoothness requirements, and physical, for example, constraints on the values of the state process or controls. For example, in the investment-consumption problem we will only consider processes $X^{\nu,C}$ for which a solution to (2.1) exists. We will also require $C_t \geq 0$ and that the investor have non-negative wealth at all times, which places further restrictions on the class of allowable controls.

Objective function. There is some cost/gain associated with the system, which may depend on the system state itself and on the control used. The objective function contains this information and is typically expressed as a function $J(x, \alpha)$ (or in finite-time horizon case $J(t, x, \alpha)$), representing the expected total cost/gain starting from system state x (at time t in finite-time horizon case) if control process α is implemented.

For example, in the setup of (2.3) the *objective functional* (or gain/cost map) is

$$J(0, x, \nu) = \mathbb{E}\left[U\big(X^\nu(T)\big)\right], \tag{2.5}$$

as it denotes the reward associated with initial wealth x and portfolio process ν. Note that in the case of no-consumption, and given the remaining parameters of the problem (i.e. μ and σ), both x and ν determine by themselves the value of the reward.

Value function. The value function describes the value of the maximum possible gain of the system (or minimal possible loss). It is usually denoted by v and is obtained, for initial state x (or (t, x) in finite-time horizon case), by optimizing the cost over all admissible controls. The goal of a stochastic control problem is to find the value function v and find a control α^* whose cost/gain attains the minimum/maximum value: $V(x) = J(x, \alpha^*)$ for starting state x. For completeness sake, from (2.3) and (2.5), if ν^* is the optimal control, then we have the *value function*

$$V(x) = \sup_\nu \mathbb{E}\left[U\big(X^\nu(T)\big)\right] = \sup_\nu J(x, \nu) = J(x, \nu^*). \tag{2.6}$$

Typical questions of interest Typical questions of interest in Stochastic control problems include:

- Is there an optimal control?
- Is there an optimal Markov control?
- How can we find an optimal control?
- How does the value function behave?
- Can we compute or approximate an optimal control numerically?

There are of course many more and, before we start, we need to review some concepts of stochastic analysis that will help in the rigorous discussion of the material in this section so far.

2.2 Controlled diffusions

We now introduce controlled SDEs with a finite time horizon $T > 0$; the infinite-horizon case is discussed later. Again, $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space with filtration $(\mathcal{F}_t)$ and a d'-dimensional Wiener process W compatible with this filtration.

We are given an action set A (in general separable complete metric space) and let $\mathcal{U}_0$ be the set of all A-valued progressively measurable processes, the controls. The controlled state is defined through an SDE as follows. Let

$$b : [0,T] \times \mathbb{R}^d \times A \to \mathbb{R}^d \quad \text{and} \quad \sigma : [0,T] \times \mathbb{R}^d \times A \to \mathbb{R}^{d\times d'}$$

be measurable functions satisfying the Lipschitz and linear growth conditions[2].

Assumption 2.3. There is a constant K and an integrable process $\kappa = \kappa_t$ such that for any t, x, y, a we have

$$|b(t,x,a) - b(t,y,a)| + |\sigma(t,x,a) - \sigma(t,y,a)| \le K|x-y|, \tag{2.7}$$
$$|b(t,x,a)| + |\sigma(t,x,a)| \le \kappa_t(1+|x|) \tag{2.8}$$

and $\mathbb{E}\int_0^t \kappa_s^2\,ds < \infty$ for any t.

Let $\mathcal{U} \subseteq \mathcal{U}_0$ be the subset of control processes for which we have Assumption 2.3. We will refer to this set as *admissible controls*. Note that in most of our examples $\alpha \in \mathcal{U}$ if and only if $\mathbb{E}\int_0^T \alpha_s^2\,ds < \infty$.

Given a fixed control $\alpha \in \mathcal{U}$, we consider the SDE for $0 \le t \le T \le \infty$ for $s \in [t,T]$

$$dX_s = b\big(s, X_s, \alpha_s\big)\,dt + \sigma\big(s, X_s, \alpha_s\big)\,dW_s, \quad X_t = \xi. \tag{2.9}$$

With Assumption 2.3 the SDE (2.9) is a special case of an SDE with random coefficients, see (1.5). As discussed in Section 1.3 and the results there, we have the following result.

Proposition 2.4 (Existence and uniqueness). *Let $t \in [0,T]$, $\xi \in L^2(\mathcal{F}_t)$ and $\nu \in \mathcal{U}_0$. Then SDE* (2.9) *has a unique (strong) Markov solution $X = X^{\alpha}_{t,\xi}$ on the interval $[t,T]$ such that*

$$\sup_{\alpha\in\mathcal{U}} \mathbb{E} \sup_{s\in[t,T]} |X_s|^2 \le c(1 + \mathbb{E}|\xi|^2)\,.$$

Moreover, the solution has the properties listed in Proposition 1.28.

2.3 Formulation of stochastic control problems

In this section we revisit the ideas of the opening one and give a stronger mathematical meaning to the general setup for optimal control problems. We distinguish the finite time horizon $T < \infty$ and the infinite time horizon $T = \infty$, the functional to optimize must differ.

[2] Note that the Lipschitz condition for a certain variable implies that the linear growth condition is satisfied for that variable. One is not able to conclude anything about the (possibly) other variables. From (2.7) one cannot conclude (2.8) as the latter assumes linear growth in the a variable. In mathematical terms, from (2.7) on can only conclude that $|b(t,x,a)| + |\sigma(t,x,a)| \le K_{t,a}(1+|x|)$ with the associated constant K depending on t and a.

In general, texts either discuss maximization or a minimization problems. Using analysis results, it is easy to jump between minimization and maximization problems: $\max_x f(x) = -\min_x -f(x)$ and the x^* that maximizes f is the same one that minimizes $-f$ (draw a picture to convince yourself).

Finite time horizon

Let

$$J(t,\xi,\alpha) := \mathbb{E}\left[\int_t^T f(s, X_s^{\alpha,t,\xi}, \alpha_s)\, ds + g(X_T^{\alpha,t,\xi})\right],$$

where $X_{t,\xi}$ solves (2.9) (with initial condition $X(t) = \xi$). The J here is called the *objective functional*. We refer to f as the *running gain* (or, if minimizing, *running cost*) and to g as the *terminal gain* (or *terminal cost*).

We will ensure the good behavior of J through the following assumption.

Assumption 2.5. There is $K > 0$ such that for all t, x, y, a we have

$$|g(x) - g(y)| + |f(t,x,a) - f(t,y,a)| \le K|x-y|,$$

$$|f(t,0,a)| \le K\,.$$

Note that this assumption is too restrictive for many of the problems we want to consider. Indeed a typical $g(x) = x^2$ is not covered by such assumption. However it makes the proofs much simpler. For bigger generality consult e.g. [Kry80].

Exercise 2.6 (The objective functional J is well-defined)**.** Let Assumptions 2.3 and 2.5 hold. Show that there is $c_T > 0$ such that $|J(\cdot,\cdot,\alpha)| < c_T(1+|x|)$ for any $\alpha \in \mathcal{U}[t,T]$.

The optimal control problem formulation We will focus on the following stochastic control problem. Let $t \in [0,T]$ and $x \in \mathbb{R}^d$. Let

$$(P)\begin{cases} v(t,x) := \sup\limits_{\alpha\in\mathcal{U}[t,T]} J(t,x,\alpha) = \sup\limits_{\alpha\in\mathcal{U}[t,T]} \mathbb{E}\left[\int_t^T f\big(s, X_s^{\alpha,t,x}, \alpha_s\big)\, ds + g\big(X_T^{\alpha,t,x}\big)\right] \\ \text{and } X^{\alpha,t,x} \text{ solves (2.9) with } X_t^{\alpha,t,x} = x. \end{cases}$$

The solution to the problem (P), is the *value function*, denoted by v. One of the mathematical difficulties in stochastic control theory is that we don't even know at this point whether v is measurable or not.

In many cases there is no optimal control process α^* for which we would have $v(t,x) = J(t,x,\alpha^*)$. Recall that v is the value function of the problem (P). However there is always an ε-optimal control (simply by definition of supremum).

Definition 2.7 (ε-optimal controls)**.** Take $t \in [0,T]$ and $x \in \mathbb{R}^m$. Let $\varepsilon \ge 0$. A control $\alpha^\varepsilon \in \mathcal{U}_{ad}[t,T]$ is said to be *ε-optimal* if

$$v(t,x) \le \varepsilon + J(t,x,\alpha^\varepsilon)\,. \tag{2.10}$$

Lemma 2.8 (Lipschitz continuity in x of the value function)**.** *If Assumptions 2.3 and 2.5 hold then there exists $C_T > 0$ such that for all $t \in [0,T]$ and $x, y \in \mathbb{R}^d$ we have*

$$|v(t,x) - v(t,y)| \le C_T|x-y|\,.$$

Proof. The first step is to show that there is $C_T > 0$ such that for any $\alpha \in \mathcal{U}$ we have

$$I := |J(t,x,\alpha) - J(t,y,\alpha)| \leq C_T|x-y|\,.$$

We note that due to Hölder's and Young's inequalities

$$\begin{aligned} I^2 &\leq \mathbb{E}\left[2\left(\int_t^T f(s,X_s^{t,x,\alpha},\alpha_s) - f(s,X_s^{t,y,\alpha},\alpha_s)\,ds\right)^2 + 2\big(g(X_s^{t,x,\alpha}) - g(X_s^{t,x,\alpha})\big)^2\right] \\ &\leq \mathbb{E}\left[2T\int_t^T |f(s,X_s^{t,x,\alpha},\alpha_s) - f(s,X_s^{t,y,\alpha},\alpha_s)|^2\,ds + 2\big|g(X_s^{t,x,\alpha}) - g(X_s^{t,x,\alpha})\big|^2\right]. \end{aligned}$$

Using Assumption 2.5 (Lipschitz continuity in x of f and g) we get

$$I^2 \leq 2TK^2\int_t^T \mathbb{E}|X_s^{t,x,\alpha} - X_s^{t,y,\alpha}|^2\,ds + 2K^2\mathbb{E}\big|X_s^{t,x,\alpha}X_s^{t,x,\alpha}\big|^2\,.$$

Then, using Proposition 1.28, we get

$$I^2 \leq 2(T^2+1)K^2\sup_{t\leq s\leq T}\mathbb{E}|X_s^{t,x,\alpha} - X_s^{t,y,\alpha}|^2 \leq C_T|x-y|^2\,.$$

We now need to apply this property of J to the value function v. Let $\varepsilon > 0$ be arbitrary and fixed. Then there is $\alpha^\varepsilon \in \mathcal{U}$ such that $v(t,x) \leq \varepsilon + J(t,x,\alpha^\varepsilon)$. Moreover $v(t,y) \geq J(t,y,\alpha^\epsilon)$. Thus

$$v(t,x) - v(t,y) \leq \varepsilon + J(t,x,\alpha^\varepsilon) - J(t,y,\alpha^\epsilon) \leq \varepsilon + C_T|x-y|\,.$$

With $\varepsilon > 0$ still the same and fixed there would be $\beta^\varepsilon \in \mathcal{U}$ such that $v(t,y) \leq \varepsilon + J(t,y,\beta^\varepsilon)$. Moreover $v(t,x) \geq J(t,x,\beta^\epsilon)$ and so

$$v(t,y) - v(t,x) \leq \varepsilon + J(t,y,\beta^\varepsilon) - J(t,x,\beta^\epsilon) \leq \varepsilon + C_T|x-y|\,.$$

Hence $-\varepsilon - C_T|x-y| \leq v(t,x) - v(t,y) \leq \varepsilon + C_T|x-y|$. Letting $\varepsilon \to 0$ concludes the proof. □

An important consequence of this is that if we fix t then $x \mapsto v(t,x)$ is measurable (as continuous functions are measurable).

2.4 Exercises

Exercise 2.9 (Further moment bounds)**.** Show that if b and σ satisfy Assumption 2.3 then for all $k > 0$ there is C_T such that

$$\sup_{s\in[0,T]}\mathbb{E}\big[\,|X_s|^{2k}\big] < C_T(1+|x|^k)\ \ \forall x \in \mathbb{R}^d\,.$$

Exercise 2.10 (Bound on J)**.** Show that, if there exist constants $M > 0$ and $k > 0$, such that for any $s \in [0,\ T]$, $x \in \mathbb{R}^d$ and $a \in A$,

$$|f(s,x,\nu)| + |g(x)| \leq M(1+|x|^k)$$

then there exists $C_T > 0$ such that $|J(t,x,\alpha)| < c_T(1+|x|^k)$ for any $\alpha \in \mathcal{U}$ and any $x \in \mathbb{R}^d$.