Stochastic Control and Dynamic Asset Allocation*

Gonçalo dos Reis[†] David Šiška[‡] Jiawei Li[§]

13th January 2026

Contents

*Lecture notes for academic year 2025/26, School of Mathematics, University of Edinburgh.

[†] G.dosReis@ed.ac.uk
[‡] D.Siska@ed.ac.uk
[§] jli12@ed.ac.uk

Reading these notes

The reader is expected to know basic stochastic analysis and ideally a little bit of financial mathematics. The notation and basic results used throughout the notes are in Appendix A.

Section 1 is introduction. Section 2 is a brief introduction to the controlled Markov chains: a discrete space and time setting for stochastic control problems. Section 3 covers basics of stochastic differential equations and is essential reading for what follows. Section 4 introduces the Bellman principle / Dynamic Programming Principle and the Bellman PDE / Hamilton–Jacobi–Bellman PDE for controlled diffusions. This is the first set of tools that can be used for solving control problems involving controlled diffusions. Section 5 extends Section 4 to the case of jump diffusions (without giving proofs) and then focuses on some applications in algorithmic trading and market making. Section 6 takes a "calculus of variations" approach to solving control problems by establishing a "first order condition" by calculating the derivative of the objective functional w.r.t. a perturbation in control. This is known as the Stochastic Maximum Principle or Pontryagin's maximum or optimality principle. Sections 4 and 6 are basically independent of each other as provide two independent ways of solving control problems.

Exercises

You will find a number of exercises throughout these notes. You must make an effort to solve them (individually or with friends).

Solutions to some (most) of the exercises are available but remember: no one ever learned swimming solely by watching other people swim (and similarly no-one ever learned mathematics solely by reading others' solutions).

Other reading

It is recommended that you read the relevant chapters of Pham [15, at least Chapters 1-3 and 6] as well as Touzi [19, at least Chapters 1-4 and 9].

Additionally one recommends Krylov [14] for those wishing to see everything done in great generality and with proofs that do not contain any vague arguments but it is not an easy book to read. Chapter 1 however, is very readable and much recommended. Those interested in applications in algorithmic trading should read Cartea, Jaimungal and Penalva [4] and those who would like to learn about mean field games there is Carmona and Delarue [3].

1 Introduction to stochastic control through examples

In this section, we introduce the basic elements in a stochastic control problem. To start with, let us look at some motivating examples.

1.1 Optimal stopping with discrete time

We start with an optimal stopping example with discrete time and space.

Example 1.1. A very simple example of an optimal stopping problem is the following: given a fair die we are told that we're allowed to roll the die for up to three times. After each roll we can either choose to stop the game and our gain is equal to the number currently apearing on the die, or to carry on. If we choose to carry on then we get nothing for this roll and we hope to get more next time. Of course if this is the 3rd time we rolled the die then we have to accept whichever number it is we got in this last roll.

In this case solving the problem is a matter of simple calculation, working backward in time. If we're in the third round then we stop, because we have no choice.

In the second round reason as follows: our expected winning in round three is

$$\frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = 3.5\,.$$

So we stop in the second round if we rolled 4, 5 or 6 as that's more then our expected outcome from continuing.

In the first round we reason as follows: our expected winning from continuing into round two are

$$\frac{1}{6}(4+5+6) + \frac{1}{2}\frac{21}{6} = 2.5 + 1.75 = 4.25$$

The first part corresponds to the decision to stop in round two. The second part corresponds to the decision to continue, weighted by the respective probabilities. So in the first round it is optimal to stop if we got 5 or 6. The optimal expected "payoff" for this optimal stopping problem is $4 + \frac{2}{3}$.

Example 1.2. There is a biased coin with $p \in (0,1)$, $p \neq \frac{1}{2}$, probability of getting heads and $q = 1-p$ probability of getting tails.

We will start with an initial wealth $x = i\delta$, $i \in \mathbb{N}$ with $i < m$, with some $m \in \mathbb{N}$ fixed reasonably large.

At each turn we choose an action $a \in \{-1, 1\}$. By choosing $a = 1$ we bet that the coin comes up heads and our wealth is increased by δ if we are correct, decreased by δ otherwise. By choosing $a = -1$ we bet on tails and our wealth is updated accordingly. That is, given that $X_{n-1} = x$ and our action $a \in \{-1, 1\}$ we have

$$\mathbb{P}(X_n = x + a\delta | X_{n-1} = x, a) = p\,, \quad \mathbb{P}(X_n = x - a\delta | X_{n-1} = x, a) = q\,.$$

The game terminates when either $x = 0$ or $x = m\delta$. Let $N = \min\{n \in \mathbb{N} : X_n = 0 \text{ or } X_n = m\delta\}$. Our aim is to maximize

$$J^\alpha(x) = \mathbb{E}\Big[X_N^\alpha | X_0 = x\Big]$$

over functions $\alpha = \alpha(x)$ telling what action to choose in each given state.

1.2 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

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

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 consumes/saves wealth at rate C_t at time $t \geq 0$ ($C_t > 0$ corresponds to consumption).

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),

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

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

$$dX_t = X_t\big[\nu_t(\mu - r) + r - \kappa_t\big]\,dt + X_t\nu_t\sigma\,dW_t\,. \tag{1.1}$$

Exercise 1.3. Assuming that all coefficients in SDE (1.1) are integrable, solve the SDE for X and hence show $X > 0$ when $X_0 = x > 0$. See Exercise 1.14 for a hint.

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 1.4 (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 $\gamma \in (-\infty, 0) \cup (0, 1)$.

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

$$\sup_{\nu,\kappa} \mathbb{E}\left[\int_0^T U\big(\kappa_t X_t^{\nu,\kappa}\big)\,dt + U\big(X_T^{\nu,\kappa}\big)\right], \text{ where } X_t^{\nu,\kappa} = X_t \text{ must satisfy (1.1).} \tag{1.2}$$

[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 $1 - \nu_t$.

Note that we could also consider the problem including a constant discount factor $\gamma \neq 0$. This will lead to a problem that different to (1.2) above:

$$\sup_{\nu,\kappa} \mathbb{E}\left[\int_0^T e^{-\gamma t} U\big(\kappa_t X_t^{\nu,\kappa}\big)\,dt + e^{-\gamma T} U\big(X_T^{\nu,\kappa}\big)\right]\,, \text{ where } X_t^{\nu,\kappa} = X_t \text{ must satisfy (1.1).}$$

Without consumption, i.e. $\forall\, t$ we have $\kappa(t) = 0$, the optimization problem (1.2) can be written as

$$\sup_{\nu} \mathbb{E}\left[U\big(X_T^{\nu}\big)\right]\,, \text{ where } X_t^{\nu} = X_t \text{ must satisfy (1.1).} \tag{1.3}$$

Again, we could have included discounting but it's not essential.

In the **infinite time-horizon case**: $T = \infty$ and we must include a discount factor $\gamma > 0$ so that the integrals coverge. The optimization problem is then (recall that $C_t = \kappa_t X_t^{\nu,\kappa}$)

$$\sup_{\nu,\kappa} \mathbb{E}\left[\int_0^\infty e^{-\gamma t} U\big(\kappa_t X_t^{\nu,\kappa}\big)\,dt\,, \text{ where } X_t^{\nu,\kappa} = X_t \text{ must satisfy (1.1).}\right] \tag{1.4}$$

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 (1.2) for the quadratic utility function: $U(x) = \lambda - x^2$, $\lambda \in \mathbb{R}$. See Example 6.12.

Exercise 1.5. Consider the problem (1.3) with $U(x) = x^\gamma$, $\gamma \in (0,1]$. Assume further that you are not allowed to change your investments as time goes by: i.e. you must choose the allocation ν_0 and $\nu_t = \nu_0$ for all $t \in [0,T]$. You should solve the problem (1.3) in the following steps:

i) Use Exercise 1.3 to obtain the solution to (1.1).

ii) Substitute this into (1.3) and use the fact that $W_T \sim \sqrt{T}N(0,1)$ to express (1.3) as a function of ν_0.

iii) Use calculus to maximize the above mentioned function.

Exercise 1.6. List some of the ways in which the model above may be considered to simplify reality too much. If you can, propose a different model or changes to the model that would rectify the issue you identified.

1.3 Optimal liquidation problem

The optimal liquidation problem is faced by a trader wishing to sell (large) amount of a certain asset within a limited time while trying to achieve the best execution price despite the fact that their trading in the asset has temporary and permanent price impacts. Trader's inventory, an $\mathbb{R}$-valued process:

$$dQ_u = -\alpha_u \, du \text{ with } Q_t = q > 0 \text{ initial inventory.}$$

Here α will typically be mostly positive as the trader should sell all the assets. We will denote this process $Q_u = Q_u^{t,q,\alpha}$ because clearly it depends on the starting point q at time t and on the trading strategy α. Asset price, an $\mathbb{R}$-valued process:

$$dS_u = -\lambda \, \alpha_u \, du + \sigma \, dW_u \, , \quad S_t = S \, .$$

We will denote this process $S_u = S_u^{t,S,\alpha}$ because clearly it depends on the starting point S at time t and on the trading strategy. Here the constant λ controls how much permanent impact the trader's own trades have on its price. Trader's execution price (for $\kappa > 0$):

$$\hat{S}_t = S_t - \kappa \alpha_t \, .$$

This means that there is a temporary price impact of the trader's trading: she doesn't receive the full price S_t but less, in proportion to her selling intensity (for instance, we assume that the trader submits a combination of limit and market orders to optimize the price when trading on a limit order book).

Quite reasonably we wish to maximize (over trading strategies α), up to to some finite time $T > 0$, the expected amount gained in sales, whilst penalising the terminal inventory (with $\theta > 0$):

$$J(t,q,S,\alpha) := \mathbb{E}\Bigg[\underbrace{\int_t^T \hat{S}_u^{t,S,\alpha} \, \alpha_u \, du}_{\text{gains from sale}} + \underbrace{Q_T^{t,q,\alpha} \, S_T^{t,S,\alpha}}_{\text{val. of inventory}} - \underbrace{\theta \, |Q_T^{t,q,\alpha}|^2}_{\text{penalty for unsold}} \Bigg] .$$

The goal is to find

$$V(t,q,S) := \sup_\alpha J(t,q,S,\alpha) \, .$$

In Section 4.2 we will show that V satisfies a nonlinear partial differential equation, called the HJB equation which will allow us to solve this optimal control problem and we will see that, in the case $\lambda = 0$, the value function (see also Figure 1.1) is

$$V(t,q,S) = qS + \gamma(t) q^2 \, ,$$

whilst the optimal control (see also Figure 1.2) is

$$a^*(t,q,S) = -\frac{1}{\kappa}\gamma(t) q \, ,$$

where

$$\gamma(t) = -\left(\frac{1}{\theta} + \frac{1}{\kappa}(T-t) \right)^{-1} .$$

It is possible to solve this with either the Bellman principle (see Exercise 4.16) or with Pontryagin maximum principle (see Example 6.11). Problems of this type arise in algorithmic trading. More can be found e.g. in Cartea, Jaimungal and Penalva [4].

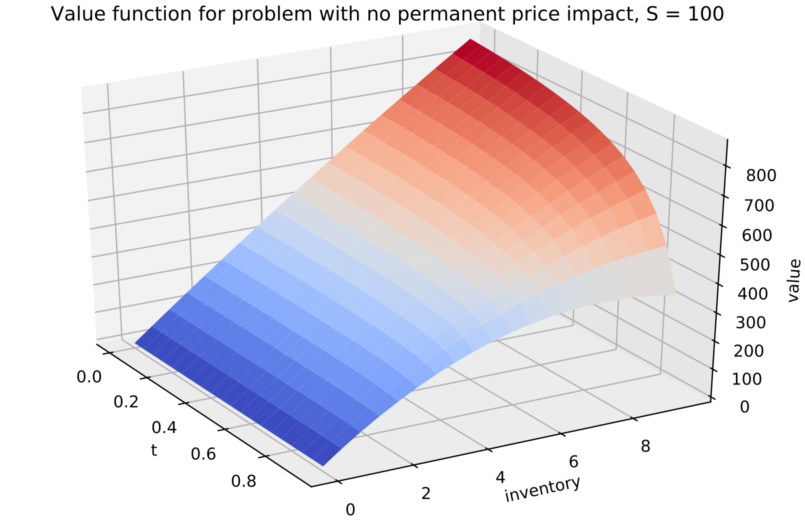


Figure 1.1: Value function for the Optimal Liquidation problem, Section 1.3, as function of time and inventory, in the case $\lambda = 0$, $T = 1$, $\theta = 10$, $\kappa = 1$ and $S = 100$.

Figure 1.2: Optimal control for the Optimal Liquidation problem, Section 1.3, as function of time and inventory, in the case $\lambda = 0$, $T = 1$, $\theta = 10$ and $\kappa = 1$.

1.4 Systemic risk - toy model

The model describes a network of N banks. We will use X_t^i to denote the logarithm of cash reserves of bank $i \in \{1, \dots, N\}$ at time $t \in [0, T]$. Let us assume that there are $N+1$ independent Wiener processes $W^0, W^1, \dots W^N$. Let us fix $\rho \in [-1, 1]$. Each bank's reserves are impacted by B_t^i where

$$B_t^i := \sqrt{1-\rho^2} W_t^i + \rho W_t^0 \,.$$

We will have bank i's reserves influenced by "its own" i.e. "idiosyncratic" source of randomness W^i and also by a source of uncertainty common to all the banks, namely W^0 (the "common noise"). Let $\bar{X}_t := \frac{1}{N} \sum_{i=1}^N X_t^i$ i.e. the mean level of log-reserves. We model the reserves as

$$dX_u^i = \left[a(\bar{X}_u - X_u^i) + \alpha_u^i\right] du + \sigma dB_u^i \,, \quad u \in [t, T] \,, \quad X_t^i = x^i \,.$$

Let us look at the terms involved:

i) The term $a(\bar{X}_u - X_u^i)$ models inter-bank lending and borrowing; if bank i is below the average then it borrows money (the log reserves increase) whilst if bank i's level is above the average then it lends out money (the log reserves decrease). This happens at rate $a > 0$.

ii) The term α_t^i is the "control" of bank i and the interpretation is that it represents lending / borrowing outside the network of the N banks (e.g. taking deposits from / lending to individual borrowers).

iii) The stochastic term (with $\sigma > 0$) models unpredictable gains / losses to the bank's reserves with the idiosyncratic and common noises as explained above.

iv) The initial reserve (at time t) of bank i is x_i.

v) Note that we should be really writing $X_u^{i,t,x,\alpha}$ for X_u^i since each bank's reserves depend on the starting point $x = (x^1, \dots, x^N)$ of all the banks and also on the controls $\alpha_u = (\alpha_u^1, \dots, \alpha_u^N)$ of all the individual banks. The equations are thus fully coupled.

We will say that in this model each bank tries to *minimize*

$$J^i(t, x, \alpha) := \mathbb{E}\bigg[\int_t^T \left(\frac{1}{2} |\alpha_u^i|^2 - q\, \alpha_u^i (\bar{X}_u^{i,t,x,\alpha} - X_u^{i,t,x,\alpha}) + \frac{\varepsilon}{2} |\bar{X}_u^{i,t,x,\alpha} - X_u^{i,t,x,\alpha}|^2 \right) du \\ + \frac{c}{2} |\bar{X}_T^{i,t,x,\alpha} - X_T^{i,t,x,\alpha}|^2 \bigg] \,.$$

Let's again look at the terms involved:

i) The term $\frac{1}{2}|\alpha_u^i|^2$ indicates that lending / borrowing outside the bank network carries a cost.

ii) With $-q\, \alpha_u^i (\bar{X}_u^{i,t,x,\alpha} - X_u^{i,t,x,\alpha})$ for some constant $q > 0$ we insist that bank i will want to borrow if it's below the mean ($\alpha_u^i > 0$) and vice versa.

iii) The final two terms provide a running penalty and terminal penalty for being too different from the average (think of this as the additional cost imposed on the bank if it's "too big to fail" versus the inefficiency of a bank that is much smaller than competitors).

Amazingly, under the assumption that $q^2 \leq \varepsilon$ it is possible to solve this problem explicitly, using either techniques we will develop in Sections 4 or 6. This is an example from the field of N-player games, much more can be found in Carmona and Delarue [3].

1.5 Optimal stopping

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which we have a d'-dimensional Wiener process $W = (W_u)_{u\in[0,T]}$ generating $\mathcal{F}_u := \sigma(W_s : s \leq u)$. Let $\mathcal{T}_{t,T}$ be the set of all $(\mathcal{F}_t)$-stopping times taking values in $[t,T]$.

Given some $\mathbb{R}^d$-valued stochastic process $(X_u^{t,x})_{u\in[t,T]}$, such that $X_t^{t,x} = x$, adapted to the filtration $(\mathcal{F}_u)_{u\in[t,T]}$ and a reward function $g : \mathbb{R}^d \to \mathbb{R}$ the optimal stopping problem is to find

$$w(t,x) = \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}\big[g(X_\tau^{t,x})\big] \,. \tag{1.5}$$

Example 1.7. A typical example is the American put option. In the Black–Scholes model for one risky asset the process $(X_u^{t,x})_{u\in[t,T]}$ is geometric Brownian motion, W is $\mathbb{R}$-valued Wiener process (and $\mathbb{P}$ denotes the risk-neutral measure in our notation here) so that

$$dX_u = rX_u\,du + \sigma X_u\,dW_u\,,\quad u\in[t,T]\,,\quad X_t = x$$

where $r \in \mathbb{R}$ and $\sigma \in [0,\infty)$ are given constants. For the American put option $g(x) := [K-x]_+$. In this case w given by (1.6) gives the no-arbitrage price of the American put option for current asset price x at time t.

It has been shown (see Krylov [14] or Gyöngy and Šiška [8]) that the optimal stopping problem (1.6) is a special case of optimal control problem given by

$$v(t,x) = \sup_{\rho\in\mathfrak{R}} \mathbb{E}\left[\int_t^T g(X_u^{t,x})\,\rho_u\,e^{-\int_t^u \rho_r\,dr} + g(X_T^{t,x})\,e^{-\int_t^T \rho_r\,dr}\right] \tag{1.6}$$

so that $w(t,x) = v(t,x)$. Here the control processes ρ_u must be adapted and such that for a given $\rho = (\rho_u)_{u\in[t,T]}$ there exists $N\in\mathbb{N}$ such that $\rho_u \in [0,N]$ for all $u\in[t,T]$.

1.6 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 (1.1).

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 (1.1)).

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. 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 (1.1) exists. We will also require the consumption process C_t such that the investor has non-negative wealth at all times.

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 (1.3) the *objective functional* (or gain/cost map) is

$$J(0, x, \nu) = \mathbb{E}\left[U\big(X_T^\nu\big)\right], \tag{1.7}$$

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: $v(t, x) = J(t, x, \alpha^*)$ for starting time t and state x. For completeness sake, from (1.3) and (1.7), if ν^* is the optimal control, then we have the *value function*

$$v(t, x) = \sup_\nu \mathbb{E}\left[U\big(X_T^\nu\big)|X_t = x\right] = \sup_\nu J(t, x, \nu) = J(t, x, \nu^*). \tag{1.8}$$

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.

1.7 Exercises

The aim of the exercises in this section is to build some confidence in manipulating the basic objects that we will be using later. It may help to browse through Section A before attempting the exercises.

Exercise 1.8. Read Definition A.18. Show that $\mathcal{H} \subset \mathcal{S}$.

Exercise 1.9 (On Gronwall's lemma)**.** Prove Gronwall's Lemma (see Lemma A.6) by following these steps:

i) Let
$$z(t) = \left(e^{-\int_0^t \lambda(r)dr}\right) \int_0^t \lambda(s)y(s)\,ds.$$
and show that
$$z'(t) \le \lambda(t)e^{-\int_0^t \lambda(r)dr}\left(b(t) - a(t)\right).$$

ii) Integrate from 0 to t to obtain the first conclusion Lemma A.6.

iii) Obtain the second conclusion of Lemma A.6.

Exercise 1.10 (On liminf)**.** Let $(a_n)_{n\in\mathbb{N}}$ be a bounded sequence. Then the number
$$\lim_{n\to\infty}\left(\inf\{a_k : k \ge n\}\right)$$
is called *limit inferior* and is denoted by $\liminf_{n\to\infty} a_n$.

1. Show that the limit inferior is well defined, that is, the limit $\lim_{n\to\infty}\left(\inf\{a_k : k \ge n\}\right)$ exists and is finite for any bounded sequence (a_n).
2. Show that the sequence $(a_n)_{n\in\mathbb{N}}$ has a subsequence that converges to $\lim_{n\to\infty}\inf a_n$.

   Hint: Argue that for any $n \in \mathbb{N}$ one can find $i \ge n$ such that
   $$\inf\{a_k : k \ge n\} \le a_i < \inf\{a_k : k \ge n\} + \frac{1}{n}.$$
   Use this to construct the subsequence we are looking for.

Exercise 1.11 (Property of the supremum/infimum)**.** Let $a, b \in \mathbb{R}$. Prove that
$$\text{if } b > 0, \text{ then} \quad \sup_{x\in X}\big\{a + bf(x)\big\} = a + b\sup_{x\in X} f(x),$$
$$\text{if } b < 0, \text{ then} \quad \sup_{x\in X}\big\{a + bf(x)\big\} = a + b\inf_{x\in X} f(x).$$

Exercise 1.12. Assume that $X = (X_t)_{t\geq 0}$ is a martingale with respect to a filtration $\mathbb{F} := (\mathcal{F}_t)_{t\geq 0}$. Show that:

1. if for all $t \geq 0$ it holds that $\mathbb{E}|X_t|^2 < \infty$ then the process given by $|X_t|^2$ is a submartingale and
2. the process given by $|X_t|$ is a submartingale.

Exercise 1.13 (ODEs)**.** Assume that (r_t) is an adapted stochastic process such that for any $t \geq 0$ $\int_0^t |r_s|\, ds < \infty$ holds $\mathbb{P}$-almost surely (in other words $r \in \mathcal{A}$).

1. Solve
$$dB_t = B_t r_t dt, \quad B_0 = 1. \tag{1.9}$$
2. Is the function $t \mapsto B_t$ continuous? Why?
3. Calculate $d(1/B_t)$.

Exercise 1.14 (Geometric Brownian motion)**.** Assume that $\mu \in \mathcal{A}$ and $\sigma \in \mathcal{S}$. Let W be a real-valued Wiener martingale.

1. Solve
$$dS_t = S_t \left[\mu_t\, dt + \sigma_t\, dW_t\right], \quad S(0) = s. \tag{1.10}$$
Hint: Solve this first in the case that μ and σ are real constants. Apply Itô's formula to the process S and the function $x \mapsto \ln x$.
2. Is the function $t \mapsto S_t$ continuous? Why?
3. Calculate $d(1/S_t)$, assuming $s \neq 0$.
4. With B given by (1.9) calculate $d(S_t/B_t)$.

Exercise 1.15 (Multi-dimensional gBm)**.** Let W be an $\mathbb{R}^d$-valued Wiener martingale. Let $\mu \in \mathcal{A}^m$ and $\sigma \in \mathcal{S}^{m\times d}$. Consider the stochastic processes $S_i = (S_i(t))_{t\in[0,T]}$ given by
$$dS^i_t = S^i_t \mu^i_t\, dt + S^i_t \sum_{j=1}^m \sigma^{ij}_t\, dW^j_t,\ S^i_0 = s_i,\ i = 1, \ldots, m. \tag{1.11}$$

1. Solve (1.11) for $i = 1, \ldots, m$.
Hint: Proceed as when solving (1.10). Start by assuming that μ and σ are constants. Apply the multi-dimensional Itô formula to the process S_i and the function $x \mapsto \ln(x)$. Note that the process S_i is just $\mathbb{R}$-valued so the multi-dimensionality only comes from W being $\mathbb{R}^d$ valued.
2. Is the function $t \mapsto S^i_t$ continuous? Why?

Exercise 1.16 (Ornstein–Uhlenbeck process)**.** Let $a, b, \sigma \in \mathbb{R}$ be constants such that $b > 0, \sigma > 0$. Let W be a real-valued Wiener martingale.

1. Solve
$$dr_t = (b - ar_t)\, dt + \sigma_t\, dW_t, \quad r(0) = r_0. \tag{1.12}$$
Hint: Apply Itô's formula to the process r and the function $(t, x) \mapsto e^{at}x$.

2. Is the function $t \mapsto r_t$ continuous? Why?

3. Calculate $\mathbb{E}[r_t]$ and $\mathbb{E}[r_t^2]$.

4. What is the distribution of r_t?

Exercise 1.17. If X is a Gaussian random variable with $\mathbb{E}[X] = \mu$ and $\mathrm{Var}(X) = \mathbb{E}[X^2 - (\mathbb{E}[X])^2] = \sigma^2$ then we write $X \sim N(\mu, \sigma^2)$. Show that if $X \sim N(\mu, \sigma^2)$ then $\mathbb{E}[e^X] = e^{\mu + \frac{\sigma^2}{2}}$.

1.8 Solutions to Exercises

Solution (Solution to Exercise 1.9)**.** Let

$$z(t) = \left(e^{-\int_0^t \lambda(r)dr}\right) \int_0^t \lambda(s) y(s)\, ds.$$

Then, almost everywhere in I,

$$z'(t) = \lambda(t) e^{-\int_0^t \lambda(r)dr} \underbrace{\left(y(t) - \int_0^t \lambda(s) y(s)\, ds\right)}_{\leq b(t) - a(t)},$$

by the inequality in our hypothesis. Hence for a.a. $s \in I$

$$z'(s) \leq \lambda(s) e^{-\int_0^s \lambda(r)dr} \left(b(s) - a(s)\right).$$

Integrating from 0 to t and using the fundamental theorem of calculus (which gives us $\int_0^t z'(s)\, ds = z(t) - z(0) = z(t)$) we obtain

$$\begin{aligned} \int_0^t \lambda(s) y(s)\, ds &\leq e^{\int_0^t \lambda(r)dr} \int_0^t \lambda(s) e^{-\int_0^s \lambda(r)dr} \left(b(s) - a(s)\right)\, ds \\ &= \int_0^t \lambda(t) e^{\int_s^t \lambda(r)dr} \left(b(s) - a(s)\right)\, ds. \end{aligned}$$

Using the left hand side of above inequality as the right hand side in the inequality in our hypothesis we get

$$y(t) + a(t) \leq b(t) + \int_0^t \lambda(s) e^{\int_s^t \lambda(r)dr} \left(b(s) - a(s)\right)\, ds,$$

which is the first conclusion of the lemma. Assume now further that b is monotone increasing and a nonnegative. Then

$$\begin{aligned} y(t) + a(t) &\leq b(t) + b(t) \int_0^t \lambda(s) e^{\int_s^t \lambda(r)dr}\, ds \\ &= b(t) + b(t) \int_0^t -d e^{\int_s^t \lambda(r)dr} = b(t) + b(t) \left(-1 + e^{\int_0^t \lambda(r)dr}\right) \\ &= b(t) e^{\int_0^t \lambda(r)\, dr}. \end{aligned}$$

Solution (Solution to Exercise 1.10)**.** Let $n \in \mathbb{N}$.

1. The sequence $b_n := \inf\{a_k : k \geq n\}$ is monotone increasing as $\{a_k : k \geq n+1\}$ is a subset of $\{a_k : k \geq n\}$, hence $b_n \leq b_{n+1}$. Additionally, the sequence is also bounded by the same bounds as the initial sequence (a_n). A monotone and bounded sequence of real numbers must converge and hence we can conclude that $\liminf_{n\to\infty} a_n$ exists.

2. It follows from the definition of infimum that there exists a sequence $i = i(n) \geq n$ such that

$$b_n = \inf\{a_k : k \geq n\} \leq a_i < \inf\{a_k : k \geq n\} + \frac{1}{n} = b_n + \frac{1}{n}.$$

The sequence of indices $\big(i(n)\big)_{n\in\mathbb{N}}$ might not be monotone, but since $i(n) \geq n$ it is always possible to select its subsequence, say $\big(j(n)\big)_{n\in\mathbb{N}}$, that is monotone.

Since $|a_{i(n)} - b_n| \to 0$ and $(b_n)_{n\in\mathbb{N}}$ converges to $\liminf_{n\to\infty} a_n$, then so does $(a_{i(n)})_n$. As $(a_{j(n)})_n$ is a subsequence of $(a_{i(n)})_n$ the same is true for $(a_{j(n)})_n$. Hence the claim follows.

Solution (Solution to Exercise 1.11)**.** We will show the result when $b > 0$, assuming that the sup takes a finite value. Let $f^* := \sup_{x\in X} f(x)$, and $V^* := \sup_{x\in X} \big\{a + bf(x)\big\}$.

To show that $V^* = a + bf^*$, we start by showing that $V^* \leq a + bf^*$.

Note that for all $x \in X$ we have $a + bf^* \geq a + bf(x)$, that is, $a + bf^*$ is an upper bound for the set $\{y : y = a + bf(x) \text{ for some } x \in X\}$. As a consequence, its least upper bound V^* must be such that $a + bf^* \geq V^* = sup_{x\in X}\{a + bf(x)\}$.

To show the converse, note that from the definition of f^* as a supremum (see Definition A.1), we have that for any $\varepsilon > 0$ there must exist a $\overline{x}^\varepsilon \in X$ such that $f(\overline{x}^\varepsilon) > f^* - \varepsilon$.

Hence $a + bf(\overline{x}^\varepsilon) > a + bf^* - b\varepsilon$. Since $\overline{x}^\varepsilon \in X$, it is obvious that $V^* \geq a + bf(\overline{x}^\varepsilon)$. Hence $V^* \geq a + bf^* - b\varepsilon$. Since ε was arbitrarily chosen, we have our result: $V^* \geq a + bf^*$.

Solution (to Exercise 1.12)**.**

1. Since X_t is $\mathcal{F}_t$-measurable it follows that $|X_t|^2$ is also $\mathcal{F}_t$-measurable. Integrability holds by assumption. We further note that the conditional expectation of a non-negative random variable is non-negative and hence for $t \geq s \geq 0$ we have

$$\begin{aligned} 0 &\leq \mathbb{E}[|X_t - X_s|^2|\mathcal{F}_s] = \mathbb{E}[|X_t|^2|\mathcal{F}_s] - 2\mathbb{E}[X_tX_s|\mathcal{F}_s] + \mathbb{E}[|X_s|^2|\mathcal{F}_s] \\ &= \mathbb{E}[|X_t|^2|\mathcal{F}_s] - 2X_s\mathbb{E}[X_t|\mathcal{F}_s] + |X_s|^2 = \mathbb{E}[|X_t|^2|\mathcal{F}_s] - |X_s|^2\,, \end{aligned}$$

since X_s is $\mathcal{F}_s$-measurable and since X is a martingale. Hence $\mathbb{E}[|X_t|^2|\mathcal{F}_s] \geq |X_s|^2$ for all $t \geq s \geq 0$.

2. First note that the adaptedness and integrability properties hold. Next note that $\big|\mathbb{E}[X_t|\mathcal{F}_s]\big| \leq \mathbb{E}\Big[|X_t|\Big|\mathcal{F}_s\Big]$ by standard properties of conditional expectations. Since X is a martingale we have

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s$$

and taking absolute value on both sides we see that

$$|X_s| = \big|\mathbb{E}[X_t|\mathcal{F}_s]\big| \leq \mathbb{E}\Big[|X_t|\Big|\mathcal{F}_s\Big]\,.$$

Solution (Solution to Exercise 1.13)**.** Let $t \in [0, \infty)$.

1. We are looking to solve:

$$B(t) = 1 + \int_0^t B(s)r(s)\,ds,$$

which is equivalent to

$$\frac{dB(t)}{dt} = r(t)B(t) \text{ for almost all } t,\ B(0) = 1.$$

Let us calculate (using chain rule and the above equation)

$$\frac{d}{dt}\left[\ln B(t)\right] = \frac{dB(t)}{dt} \cdot \frac{1}{B(t)} = r(t).$$

Integrating both sides and using the fundamental theorem of calculus

$$\ln B(t) - \ln B(0) = \int_0^t r(s)\,ds$$

and hence

$$B(t) = \exp\left(\int_0^t r(s)\,ds\right).$$

2. First we note that for any function f integrable on $[0,\infty)$ we have that the map $t \mapsto \int_0^t f(x)\,dx$ is absolutely continuous in t and hence it is continuous. The function $x \mapsto e^x$ is continuous and composition of continuous functions is continuous. Hence $t \mapsto B(t)$ must be continuous.

3. There are many ways to do this. We can start with (1.9) and use chain rule:

$$\frac{d}{dt}\left[\frac{1}{B(t)}\right] = \frac{dB(t)}{dt}\cdot\left(-\frac{1}{B^2(t)}\right) = -r(t)\left(-\frac{1}{B(t)}\right)$$

and so

$$d\left(\frac{1}{B(t)}\right) = -r(t)\frac{1}{B(t)}dt.$$

Or we can start with the solution that we have calculated write

$$\begin{aligned}\frac{d}{dt}\left[\frac{1}{B(t)}\right] &= \frac{d}{dt}\exp\left(-\int_0^t r(s)ds\right)\\ &= -r(t)\exp\left(-\int_0^t r(s)ds\right) = -r(t)\left(-\frac{1}{B(t)}\right)\end{aligned}$$

which leads to the same conclusion again.

Solution (Solution to Exercise 1.14)**.** 1. We follow the hint (but skip directly to the general μ and σ). From Itô's formula:

$$d\big(\ln S(t)\big) = \frac{1}{S(t)}dS(t) - \frac{1}{2}\frac{1}{S^2(t)}dS(t)\cdot dS(t) = \left(\mu(t) - \frac{1}{2}\sigma^2(t)\right)dt + \sigma(t)dW(t).$$

Now we write this in the full integral notation:

$$\ln S(t) = \ln S(0) + \int_0^t \left[\mu(s) - \frac{1}{2}\sigma^2(s)\right]ds + \int_0^t \sigma(s)dW(s).$$

Hence

$$S(t) = s\exp\left(\int_0^t \left[\mu(s) - \frac{1}{2}\sigma^2(s)\right]ds + \int_0^t \sigma(s)dW(s)\right). \tag{1.13}$$

Now this is the correct result but using invalid application of Itô's formula. If we want a full proof we call (1.13) a guess and we will now check that it satisfies (1.10). To that end we apply Itô's formula to $x \mapsto s\exp(x)$ and the Itô process

$$X(t) = \int_0^t \left[\mu(s) - \frac{1}{2}\sigma^2(s)\right]ds + \int_0^t \sigma(s)dW(s).$$

Thus

$$\begin{aligned}dS(t) = d(f(X(t)) &= se^{X(t)}dX(t) + \frac{1}{2}se^{X(t)}dX(t)dX(t)\\ &= S(t)\left[\left(\mu(t) - \frac{1}{2}\sigma^2(t)\right)dt + \sigma(t)dW(t)\right] + \frac{1}{2}S(t)\sigma^2(t)dt.\end{aligned}$$

Hence we see that the process given by (1.13) satisfies (1.10).

2. The continuity question is now more intricate than in the previous exercise due to the presence of the stochastic integral. From stochastic analysis in finance you know that Z given by
$$Z(t) := \int_0^t \sigma(s)dW(s)$$
is a continuous stochastic process. Thus there is a set $\Omega' \in \mathcal{F}$ such that $\mathbb{P}(\Omega') = 1$ and for each $\omega \in \Omega'$ the function $t \mapsto S(\omega, t)$ is continuous since it's a composition of continuous functions.

3. If $s \neq 0$ then $S(t) \neq 0$ for all t. We can thus use Itô's formula
$$\begin{aligned} d\left(\frac{1}{S(t)}\right) &= -\frac{1}{S^2(t)}dS(t) + \frac{1}{S^3(t)}dS(t)dS(t) \\ &= -\frac{1}{S(t)}\left[\mu(t)dt + \sigma(t)dW(t)\right] + \frac{1}{S(t)}\sigma^2(t)dt \\ &= \frac{1}{S(t)}\left[\left(-\mu(t) + \sigma^2(t)\right)dt - \sigma(t)dW(t)\right]. \end{aligned}$$

4. We calculate this with Itô's product rule:
$$\begin{aligned} d\left(\frac{S(t)}{B(t)}\right) &= S(t)d\left(\frac{1}{B(t)}\right) + \frac{1}{B(t)}dS(t) + dS(t)d\left(\frac{1}{B(t)}\right) \\ &= -r(t)\frac{S(t)}{B(t)}dt + \mu(t)\frac{S(t)}{B(t)}dt + \sigma(t)\frac{S(t)}{B(t)}dW(t) \\ &= \frac{S(t)}{B(t)}\left[\left(\mu(t) - r(t)\right)dt + \sigma(t)dW(t)\right]. \end{aligned}$$

Solution (Solution to Exercise 1.15)**.** 1. We Itô's formula to the function $x \mapsto \ln(x)$ and the process S_i. We thus obtain, for $X_i(t) := \ln(S_i(t))$, that
$$\begin{aligned} dX_i(t) = d\ln(S_i(t)) &= \frac{1}{S_i(t)}dS_i(t) - \frac{1}{2}\frac{1}{S_i^2(t)}dS_i(t)dS_i(t) \\ &= \mu_i(t)dt + \sum_{j=1}^n \sigma_{ij}(t)dW_j(t) - \frac{1}{2}\sum_{j=1}^n \sigma_{ij}^2(t)dt \\ &= \left[\mu_i(t) - \frac{1}{2}\sum_{j=1}^n \sigma_{ij}^2(t)\right]dt + \sum_{j=1}^n \sigma_{ij}(t)dW_j(t). \end{aligned}$$

Hence
$$\begin{aligned} X_i(t) - X_i(0) &= \ln S_i(t) - \ln S_i(t) \\ &= \int_0^t \left[\mu_i(s) - \frac{1}{2}\sum_{j=1}^n \sigma_{ij}^2(s)\right]ds + \sum_{j=1}^n \int_0^t \sigma_{ij}(s)dW_j(s). \end{aligned}$$

And so
$$S_i(t) = S_i(0)\exp\left\{\int_0^t \left[\mu_i(s) - \frac{1}{2}\sum_{j=1}^n \sigma_{ij}^2(s)\right]ds + \sum_{j=1}^n \int_0^t \sigma_{ij}(s)\,dW_j(s)\right\}.$$

2. Using the same argument as before and in particular noticing that for each j the function $t \mapsto \int_0^t \sigma_{ij}(s)dW_j(s)$ is continuous for almost all $\omega \in \Omega$ we get that $t \mapsto S_i(t)$ is almost surely continuous.

Solution (to Exercise 1.16)**.** 1. What the hint suggests is sometimes referred to as the "integrating factor technique." We see that

$$d(e^{at}r(t)) = e^{at}dr(t) + ae^{at}r(t)dt = e^{at}\left[bdt + \sigma dW(t)\right].$$

Integrating we get

$$e^{at}r(t) = r(0) + \int_0^t e^{as}b\,ds + \int_0^t e^{as}\sigma dW(s)$$

and hence

$$r(t) = e^{-at}r(0) + \int_0^t e^{-a(t-s)}b\,ds + \int_0^t e^{-a(t-s)}\sigma dW(s).$$

2. Yes. The arguments are the same as in previous exercises.
3. We know that stochastic integral of a deterministic integrand is a normally distributed random variable with mean zero and variance given via the Itô isometry. Hence

$$\mathbb{E}r(t) = e^{-at}r(0) + \frac{b}{a}\left(1 - e^{-at}\right)$$

and

$$\begin{aligned}\mathbb{E}r^2(t) &= (\mathbb{E}r(t))^2 + e^{-2at}\sigma^2\mathbb{E}\left[\left(\int_0^t e^{as}dW(s)\right)^2\right]\\ &= (\mathbb{E}r(t))^2 + e^{-2at}\sigma^2\int_0^t e^{2as}ds = (\mathbb{E}r(t))^2 + \frac{\sigma^2}{2a}\left(1 - e^{-2at}\right).\end{aligned}$$

Hence

$$\mathrm{Var}\left[r(t)\right] = \frac{\sigma^2}{2a}\left(1 - e^{-2at}\right).$$

4. Stochastic integral of a deterministic integrand is a normally distributed random variable. Hence for each t we know that $r(t)$ is normally distributed with mean and variance calculated above.

Solution (to Exercise 1.17)**.** Let $Y \sim N(0,1)$. Then

$$\begin{aligned}\mathbb{E}e^X &= \mathbb{E}e^{\mu+\sigma Y} = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{\mu+\sigma z}e^{-\frac{1}{2}z^2}dz = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{1}{2}\left[(z-\sigma)^2-\sigma^2-2\mu\right]}dz\\ &= e^{\frac{1}{2}\sigma^2+\mu}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{1}{2}(z-\sigma)^2}dz = e^{\frac{1}{2}\sigma^2+\mu},\end{aligned}$$

since $z \mapsto \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(z-\sigma)^2}$ is a density of normal random variable with mean σ and variance 1 and thus its integral over the whole of real numbers must be 1.

2 Controlled Markov chains

In this section we consider the control problem in the setting of discrete space and time. This will allow us to introduce the dynamic programming principle and the Bellman equation in a simpler setting that that required for controlled diffusions: in particular most of the tricky problems around measurability won't arise here.

2.1 Problem setting

Let S be a discrete state space. Let $\partial S \subset S$ this is the set of absorbing states (often we'll think of this set as the boundary). Let A be an action set which will be assumed to be finite in this section. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For $a \in A$, $y, y' \in S$ assume we are given $p^a(y, y')$ which are the transition probabilities of a discrete time Markov chain $(X_n^\alpha)_{n=0,1,\ldots}$ so that

$$\mathbb{P}(X_{n+1}^\alpha = y' \,|\, X_n^\alpha = y) = p^{\alpha_n}(y, y')\,.$$

Here α is control process. We require that α_n is measurable with respect to $\sigma(X_k^\alpha : k \leq n)$. We will label the collection of all such controlled processes by $\mathcal{A}$. This covers the description of the controlled Markov chain.

Let $\gamma \in (0,1)$ be a fixed discount factor. Let $f : A \times S \to \mathbb{R}$ and $g : S \to \mathbb{R}$ be the given "running reward" and "terminal reward". Assume that $f(a, x) = 0$ for all $a \in A$ and $x \in \partial S$ (in other words in the stopping states the running reward is zero). Let

$$N^{\alpha,x} := \min\{n = 0, 1, \ldots : X_n^\alpha \in \partial S \;\text{ and }\; X_0^\alpha = x\} \tag{2.1}$$

be the first hitting time of one of the absorbing states. Let $\mathbb{E}^x[\cdot] := \mathbb{E}[\cdot | X_0 = x]$ and let

$$J^\alpha(x) := \mathbb{E}^x\left[\sum_{n=0}^{N} \gamma^n f(\alpha_n, X_n^\alpha) + \gamma^N g(X_N^\alpha)\right]. \tag{2.2}$$

Our aim is to maximise J over all control processes α which are adapted as described above (i.e. over all $\alpha \in \mathcal{A}$). Finally, for all $x \in S$, let

$$v(x) := \max_{\alpha \in \mathcal{A}} J^\alpha(x)\,. \tag{2.3}$$

Note that from (2.1) we have $N^{x,\alpha} = 0$ for $x \in \partial S$. So for $x \in \partial S$ we have $v(x) = g(x)$, due to (2.2)-(2.3) and since we are assuming that in the stopping states the running reward is 0.

2.2 Dynamic programming for controlled Markov chains

Theorem 2.1. *Let f and g be bounded. Then for all $x \in S \setminus \partial S$ we have*

$$v(x) = \max_{a \in \mathcal{A}} \mathbb{E}^x\left[f(a, x) + \gamma v(X_1^a)\right]. \tag{2.4}$$

Proof. Fix the control process α. From (2.2) and the tower property we get

$$J^\alpha(x) = f(\alpha_0, x) + \mathbb{E}^x\left[\mathbb{E}\Big[\sum_{n=1}^{N} \gamma^n f(\alpha_n, X_n^\alpha) + \gamma^N g(X_N^\alpha) | X_1^\alpha\Big]\right]. \tag{2.5}$$

Consider now a control process β given by $\beta_n = \alpha_{n+1}$ and the associated controlled Markov chain Y^β s.t. $Y_0^\beta = X_1^\alpha$ with the transition probabilities given by $p^{\beta_n}(y,y')$. This means that Y_{n-1}^β has the same law as X_n^α and so

$$\mathbb{E}\Big[\sum_{n=1}^{N}\gamma^n f(\alpha_n, X_n^\alpha) + \gamma^N g(X_N^\alpha)|X_1^\alpha\Big] = \gamma\mathbb{E}\Big[\sum_{n=1}^{N}\gamma^{n-1} f(\beta_{n-1}, Y_{n-1}^\beta) + \gamma^{N-1} g(Y_{N-1}^\beta)|Y_0^\beta = X_1^\alpha\Big]$$
$$= \gamma\mathbb{E}\Big[\sum_{n=0}^{N-1}\gamma^n f(\beta_n, Y_n^\beta) + \gamma^{N-1} g(Y_{N-1}^\beta)|Y_0^\beta = X_1^\alpha\Big] = \gamma J^\beta(X_1^\alpha)\,,$$

since $N-1$ is the first hitting time of ∂S for the process Y^β. Hence from (2.5) we get

$$\begin{aligned} J^\alpha(x) &= f(\alpha_0, x) + \mathbb{E}^x\big[\gamma J^\beta(X_1^\alpha)\big] \\ &\le f(\alpha_0, x) + \mathbb{E}^x\big[\gamma v(X_1^\alpha)\big] \le \max_{a\in A}\Big[f(a,x) + \mathbb{E}^x\big[\gamma v(X_1^a)\big]\Big]\,. \end{aligned}$$

Taking supremum on the left hand side over all control processes α leads to

$$v(x) \le \max_{a\in A}\mathbb{E}^x\Big[f^a(x) + \gamma v(X_1^a)\Big]\,.$$

It remains to prove the inequality in the other direction. Let $a^* \in A$ be the action which achieves the maximum in $\max_{a\in A}\mathbb{E}^x\Big[f^a(x) + \gamma v(X_1^a)\Big]$. Let $\alpha^{*,y}$ be the control process which, for a given $y \in S$, achieves the maximum in (2.3) (with x replaced by y). Define a new control process

$$\beta_n = \begin{cases} a^* & \text{if} \quad n = 0 \\ \alpha_{n-1}^{*,y} & \text{if} \quad n > 0 \;\text{ and }\; X_1^{a^*} = y\,. \end{cases}$$

We see that the processes $(x, Y_0^{\alpha^*, X_1^{a^*}}, Y_1^{\alpha^*, X_1^{a^*}}, \ldots)$ and $(X_0^\beta, X_1^\beta, X_2^\beta, \ldots)$ are identically distributed. Then, since $N-1$ is the first hitting time of ∂S by the process $Y^{\alpha^*, X^{a^*}}$, we have

$$\begin{aligned}
&\mathbb{E}^x\big[f(a^*, x) + \gamma v(X_1^{a^*})\big] \\
&= \mathbb{E}^x\Bigg[f(a^*, x) + \gamma\mathbb{E}^x\Big[\sum_{n=0}^{N-1}\gamma^n f(\alpha_n^{*,X_1^{a^*}}, Y_n^{\alpha^*, X_1^{a^*}}) + \gamma^{N-1} g(Y_{N-1}^{\alpha^*, X_1^{a^*}})|Y_0 = X_1^{a^*}\Big]\Bigg] \\
&= \mathbb{E}^x\Bigg[f(a^*, x) + \gamma\mathbb{E}^x\Big[\sum_{k=1}^{N}\gamma^{k-1} f(\alpha_{k-1}^{*,X_1^{a^*}}, Y_{k-1}^{\alpha^*, X_1^{a^*}}) + \gamma^{N-1} g(Y_{N-1}^{\alpha^*, X_1^{a^*}})|Y_0 = X_1^{a^*}\Big]\Bigg] \\
&= \mathbb{E}^x\Bigg[f(\beta_0, X_0^\beta) + \gamma\Big[\sum_{k=1}^{N}\gamma^{k-1} f(\beta_k, X_k^\beta) + \gamma^{N-1} g(X_N^\beta)\Big]\Bigg] \\
&= \mathbb{E}^x\Big[\sum_{n=0}^{N}\gamma^n f(\beta, X_n^\beta) + \gamma^N g(X_N^\beta)\Big] \le v(x)\,.
\end{aligned}$$

This completes the proof. $\square$

An immediate consequence of the Bellman principle is that among all control processes α it is enough to consider the ones that depend only on the current state. Indeed, define a function[2] $a^* = a^*(x)$ as

$$a^*(x) \in \arg\max_{b \in A} \mathbb{E}^x \left[f^b(x) + \gamma v(X_1^b) \right].$$

Define X_n^* as $X_0^* = x$ and for $n \in \mathbb{N}$ define X_n^* by $\mathbb{P}(X_n^* = y') = p^{a^*(X_{n-1}^*)}(X_{n-1}^*, y')$. Define the control process $\alpha_n^* := a^*(X_n^*)$. Then

$$v(x) = \mathbb{E}^x \left[f^{a^*}(x) + \gamma v(X_1^*) \right]$$

i.e. the optimal payoff is achieved with this Markov control.

2.3 Bellman equation for controlled Markov chain

Note that

$$\mathbb{E}^x \left[v(X_1^a) \right] = \sum_{y \in S} v(y) p^a(x, y).$$

Define the following two vectors and the matrix:

$$\begin{aligned} V_i &:= v(x_i), \quad F_i^a := f(a, x_i), \quad i = 1, \dots, |S|, \quad a \in A, \\ P_{ij}^a &:= p^a(x_i, x_j), \quad i, j = 1, \dots, |S|, \quad a \in A. \end{aligned}$$

Then (2.4) can be stated as the following nonlinear system which we will call the Bellman equation:

$$\begin{aligned} V_i &= \max_{a \in A} \left[F_i^a + \gamma (P^a V)_i \right] \text{ for } i \text{ such that } x_i \in S \setminus \partial S, \\ V_i &= g(x_i) \text{ for } i \text{ such that } x_i \in \partial S. \end{aligned} \tag{2.6}$$

Note that if we have managed to solve (2.6) then we can very easily obtain the optimal control for each state $i = 1, \dots, |S|$. Indeed we just have to solve the (static) maximization:

$$a_i^* \in \arg\max_{a \in A} \left[F_i^a + \gamma (P^a V)_i \right].$$

On the other hand if we somehow figure out the optimal decision a_i^* to be taken in each state $x_i \in S$ then the Bellman equation reduces to a linear problem:

$$V_i = F_i^{a_i^*} + \gamma (P^{a_i^*} V)_i, \quad i = 1, \dots, |S|.$$

There are two basic numerical algorithms that can be used to solve the Bellman equation.

Value iteration Start with an initial guess $V^{(0)} \in \mathbb{R}^{|S|}$. For $k >$ define $V^{(k)} \in \mathbb{R}^{|S|}$ recursively as

$$\begin{aligned} V_i^{(k)} &= \max_{a \in A} \left[F_i^a + \gamma (P^a V^{(k-1)})_i \right], \quad i = 1, \dots, |S|, \quad x_i \in S \setminus \partial S. \\ V_i^{(k)} &= g(x_i), \quad x_i \in \partial S. \end{aligned}$$

It can be shown that (often) $\lim_{k \to \infty} V^{(k)} = V$ and moreover this convergence is fast (e.g. exponential). See Puterman [16, Ch. 6, Sec. 3].

[2]Since there may be multiple b which maximize the expression $\mathbb{E}^x \left[f^b(x) + \gamma v(X_1^b) \right]$ we need to make a choice of a specific b to make $a^* = a^*(x)$ into a function taking values in A. Since A is countable this is easy. We may, for example, index the elements of A and then always choose the one with the lowest index.

Policy iteration Start with an initial guess of $a_i^{(0)}$ for each $i = 1, \ldots, S$. Let $V_i^{(k)}$ and $a_i^{(k)}$, for all $i = 1, \ldots, S$ be defined through the iterative procedure for $k \geq 0$:

i) Solve the linear system

$$U_i = F_i^{a_i^{(k)}} + \gamma(P^{a_i^{(k)}} U)_i \,, \;\; i = 1, \ldots, |S|$$

and set $V^{(k+1)} = U$ on $S \setminus \partial S$ and set $V_i^{(k+1)} = g(x_i)$ for all $x_i \in \partial S$.

ii) Solve the static maximization problem

$$a_i^{(k+1)} \in \arg\max_{a \in A} \left[F_i^a + \gamma(P^a V^{(k+1)})_i \right] .$$

Again, it can be shown that (often) $\lim_{k \to \infty} V^{(k)} = V$ and moreover this convergence is fast (e.g. exponential). See Puterman [16, Ch. 6, Sec. 4].

2.4 Q-learning for unknown environments

So far we assumed that $p = p^a(y, y')$, $f = f^a(y)$, $g = g(y)$ are known. If they are unknown then Q-learning provides an iterative method for learning the value function and hence for obtaining the optimal policy. There are other numerical methods, see Barto and Sutton [18] for a comprehensive overview of Reinforcement Learning.

Define the Q-values (or action values) as:

$$Q(x, a) := f^a(x) + \gamma \mathbb{E}^x \Big[v(X_1^a) \Big] . \tag{2.7}$$

We see that this is the (discounted) expected reward for executing action a in state x and then following the optimal policy thereafter. Let us now take the maximum over all possible actions $a \in A$. Then

$$\max_{a \in A} Q(x, a) = \max_{a \in A} \mathbb{E}^x \Big[f^a(x) + \gamma v(X_1^a) \Big] .$$

From Theorem 2.4 we then see that

$$\max_{a \in A} Q(x, a) = v(x) .$$

From this and the definition of Q-function (2.7) we have

$$Q(x, a) = f^a(x) + \gamma \mathbb{E}^x \Big[\max_{b \in A} Q(X_1^a, b) \Big]$$

which we can re-arrange as

$$0 = \mathbb{E}^x \Big[f^a(x) + \gamma \max_{b \in A} Q(X_1^a, b) - Q(x, a) \Big] .$$

At this point we are close to formulating learning the Q-function as a "stochastic approximation" algorithm. If our state space S and action space A are finite (as we assumed earlier e.g. in Section 2.3) then the Q function can be thought of as a matrix $\mathbf{Q}$ with $|S| \times |A|$ entries[3] i.e. $\mathbf{Q} \in \mathbb{R}^{|S| \times |A|}$. Then $\mathbf{Q}_{ik} = Q(x_i, a_k)$ with $S = \{x_1, \ldots, x_{|S|}\}$ and $A = \{a_1, \ldots, a_{|A|}\}$.

[3] For a finite set S we write $|S|$ to denote the number of elements in S.

Define

$$\mathcal{F}(\mathbf{Q}, X, x_i, a_k) := f^{a_k}(x_i) + \gamma \max_{k'=1,\ldots,|A|} Q(X, a_{k'}) - Q(x_i, a_k)\,.$$

To find $\mathbf{Q}$ we need to solve

$$0 = \mathbb{E}\Big[\mathcal{F}(\mathbf{Q}, X^{x_i,a_k}, x_i, a_k)\Big] \text{ for all } \; i = 1, \ldots, |S|\,, \quad k = 1, \ldots, |A|\,,$$

where X^{x_i,a_k} is the r.v. representing the state the system will be when action a_k is taken in state x_i.

2.5 Robbins–Monro algorithm

Forget for a moment everything about control and Q-learning and consider the problem of finding $\theta \in \mathbb{R}^p$ such that

$$0 = \mathbb{E}\big[c(X^\theta, \theta)\big]\,,$$

where X^θ is an $\mathbb{R}^d$-valued random variable which depends on the parameters θ and $c = c(x, \theta)$ is function of x and the parameters θ. I.e. we can say $X^\theta \sim \pi_\theta$ for some family of distributions $(\pi_\theta)_{\theta \in \mathbb{R}^p}$ and write the problem equivalently as

$$0 = \int_{\mathbb{R}^d} c(x, \theta)\, \pi_\theta(dx)\,.$$

The Robbins–Monro algorithm starts with a guess θ_0 and then updates the estimate as

$$\theta_{k+1} = \theta_k + \delta_k c(x^{\theta_k})\,,$$

where (x^{θ_k}) are samples from π_{θ_k}. It is possible to prove that under fairly weak assumptions this sequence converges to the true solution as long as $(\delta_k)_{k\in\mathbb{N}}$ are such that $\sum_k \delta_k = \infty$ and $\sum_k \delta_k^2 < \infty$.

Example 2.2. You can think about the problem if finding implied volatility in the Black–Scholes model.[4] Let and $X^\theta = S\exp((r-(1/2)\theta^2)T + \theta\sqrt{T}Z)$ with $Z \sim N(0,1)$. Let $c(X^\theta) = e^{-rT}[X^\theta - K]_+ - C$. Here S, the current stock price, T, the maturity, r, the risk-free rate, K, the strike and C the market price of the call are fixed and known. The implied volatility is θ such that

$$C = e^{-rT}\mathbb{E}\Big[[X^\theta - K]_+\Big] \text{ equivalently } \; 0 = \mathbb{E}\big[c(X^\theta)\big]\,.$$

Example 2.3 (Ordinary least squares)**.** Consider some $\mathbb{R}^d$-valued r.v. X. We wish to find a linear approximation of $f : \mathbb{R}^d \to \mathbb{R}^{d'}$ so that $f(X) \approx WX + b$ for a matrix $W \in \mathbb{R}^{d' \times d}$ and vector $b \in \mathbb{R}^d$ which minimizes the mean-square error $\mathbb{E}|f(X) - WX - b|^2$.

If we had a finite number N of samples from X to estimate this we will solve this as by finding the minimum of $\frac{1}{N}\sum_{k=1}^N |f(x_k) - Wx_k - b|^2$. The minimum is achieved when the gradient is 0 i.e. when

$$\sum_{k=1}^N 2x_k(f(x_k) - Wx_k - b) = 0 \text{ and } \sum_{k=1}^N (f(x_k) - Wx_k) - Nb = 0\,.$$

[4] In practice you would of course use the Black–Scholes formula and a nonlinear solver like the bisection method or Newton method.

Now we could solve this for the optimal (W, b).

If, however, our observations arrive in time as a sequence $(x^{(k)})_{k\in\mathbb{N}}$, we don't know how many there will be and we want to keep a running estimate of the optimal (W, b) then we would start with a guess $(W^{(0)}, b^{(0)})$ and then update these using Robins–Monro:

$$W^{(k+1)} = W^{(k)} - 2\delta_k x^{(k)}\Big(f(x^{(k)}) - Wx^{(k)} - b^{(k)}\Big) \text{ and } b^{(k+1)} = b^{(k)} - 2\delta_k\Big(f(x^{(k)}) - Wx^{(k)} - b^{(k)}\Big)\,.$$

The reason for this is that we want

$$\nabla_W\Big[\mathbb{E}|f(X) - WX - b|^2\Big] = 0 \text{ and } \nabla_b\Big[\mathbb{E}|f(X) - WX - b|^2\Big] = 0\,.$$

This is

$$-\mathbb{E}\Big[2X(f(X) - WX - b)\Big] = 0 \text{ and } -\mathbb{E}\Big[f(X) - WX - b\Big] = 0\,.$$

So in the notation above our $\theta = (W, b)$ and the nonlinear function c is

$$c(X, \theta) = \begin{pmatrix} -2X\big(f(X) - WX - b\big) \\ -\big(f(X) - WX - b\big) \end{pmatrix}\,.$$

2.6 The Q-learning algorithm

Coming back to Q-learning we will formulate the algorithm as a version of the Robins–Monro method. We need to fix the learning rate at each step: $(\delta_k)_{k\in\mathbb{N}}$ such that $\delta_k \in (0, 1)$, $\sum_k \delta_k^2 < \infty$ and $\sum_k \delta_k = \infty$. We can take $\delta_k = \frac{1}{k}$ if we wish to.

We start by making an initial guess for $\mathbf{Q}$, call it $\mathbf{Q}^{(0)} = Q^{(0)}(x, a)$. The learning proceeds in episodes, where at the k-th episode:

i) We are in the state $x^{(k)}$ (this can be either the resulting state of a previous episode or one chosen at random).

ii) We select and perform action $a^{(k)}$ (randomly, or cycle through all possible actions systematically, or using some ε-greedy heuristic[5]).

iii) We observe the state we landed in, denoting it $y^{(k)}$. This is our sample from $X^{x^{(k)},a^{(k)}}$. If $y^{(k)} \in \partial S$ then set $R^{(k)} = f(a^{(k)}, x^{(k)}) + g(y^{(k)})$ and we will re-start. If $y^{(k)} \in S \setminus \partial S$ then set $R^{(k)} = f(a^{(k)}, x^{(k)}) + \gamma V^{(k-1)}(y^{(k)})$, where

$$V^{(k-1)}(y^{(k)}) := \max_{b\in A} Q^{(k-1)}(y^{(k)}, b)\,.$$

iv) We adjust the $Q^{(k-1)}(x^{(k)}, a^{(k)})$ to $Q^{(k)}(x^{(k)}, a^{(k)})$ using δ_k as

$$Q^{(k)}(x_k, a_k) = Q^{(k)}(x_k, a_k) + \delta_k \mathcal{F}(\mathbf{Q}, y^{(k)}, x^{(k)}, a^{(k)}) = (1 - \delta_k)Q_{k-1}(x_k, a_k) + \delta_k R_k$$

and we leave the remaining values of $Q^{(k)}$ set as $Q^{(k-1)}$ in this episode. Finally, if $y^{(k)} \in \partial S$ then $x^{(k+1)}$ is chosen at random from S, while if $y^{(k)} \in S \setminus \partial S$ then $x^{(k+1)} = y^{(k)}$.

[5]With probability $\varepsilon \in (0, 1)$ we choose a random action (exploration), with probability $1 - \varepsilon$ we choose the optimal action according to our best knowledge at episode k (exploitation):

$$a^{(k)} \in \arg\max_{b\in A} Q^{(k-1)}(x, b)\,.$$

That's all. If the algorithm converges so that $Q(x,a) = \lim_{k\to\infty} Q^{(k)}(x,a)$ we then have $v(x) = \max_a Q(x,a)$ for all x and a, as we have seen above.

An optimal behaviour in an unknown environment would then balance *exploration* and *exploitation*. You can find more on this in Sutton and Barto [18] and the proof of convergence of the Q-learning algorithm is in Watkins and Dayan [20].

2.7 Exercises

Exercise 2.4 (Simplified version of Example 1.2)**.** There is a biased coin with $p \in (0,1)$, $p \neq \frac{1}{2}$, probability of getting heads and $q = 1-p$ probability of getting tails.

We will start with an initial wealth $x = i$, $i \in \mathbb{N}$ with $i < m$, with some $m = 2$.

At each turn we choose an action $a \in \{-1, 1\}$. By choosing $a = 1$ we bet that the coin comes up heads and our wealth is increased by 1 if we are correct, decreased by 1 otherwise. By choosing $a = -1$ we bet on tails and our wealth is updated accordingly.

That is, given that $X_{n-1} = x$ and our action $a \in \{-1, 1\}$ we have

$$\mathbb{P}(X_n = x + a \,|\, X_{n-1} = x, a) = p\,, \quad \mathbb{P}(X_n = x - a \,|\, X_{n-1} = x, a) = q\,.$$

The game terminates when either $x = 0$ or $x = m = 2$. Let $N = \min\{n \in \mathbb{N} : X_n = 0 \text{ or } X_n = m\}$. Our aim is to maximize

$$J^\alpha(x) = \mathbb{E}\Big[X_N^\alpha | X_0 = x\Big]$$

over functions $\alpha = \alpha(x)$ telling what action to choose in each given state.

1. Write down what the state space S and the stopping set ∂S are and write down the transition probability matrix for P^a for $a = 1$ and for $a = -1$.
2. Write down the Bellman equation for the problem.
3. Assume that $p > 1/2$. Guess the optimal strategy. With your guess the Bellman equation is linear. Solve it for V.

2.8 Solutions to Exercises

Solution (to Exercise 2.4)**.**

1. First, $S = \{0, 1, 2\}$ and $\partial S = \{0, 2\}$. The transition probability matrices for $a = 1$ and $a = -1$ are, respectively,

$$P^{(a=1)} = \begin{pmatrix} 1 & 0 & 0 \\ q & 0 & p \\ 0 & 0 & 1 \end{pmatrix}\,, \quad P^{(a=-1)} = \begin{pmatrix} 1 & 0 & 0 \\ p & 0 & q \\ 0 & 0 & 1 \end{pmatrix}\,.$$

2. There is no running reward and $\gamma = 1$ so the Bellman equation is

$$V_i = \max_{a \in \{-1,1\}} \Big[(P^a V)_i\Big]\,.$$

3. If $p > \frac{1}{2}$ then we want to bet on heads i.e. $a = 1$. To solve the Bellman equation: we know that $V_0 = 0$, $V_2 = 2$ so we only need to find V_1. From the Bellman equation with $a = 1$ we have

$$V_1 = qV_0 + 0\,V_1 + pV_2 = 2p\,.$$

3 Stochastic control of diffusion processes

In this section we introduce existence and uniqueness theory for controlled diffusion processes and building on that formulate properly the stochastic control problem we want to solve. Finally we explore some properties of the value function associated to the control problem.

3.1 Stochastic differential equations

Let a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ be given. Let W be a d'-dimensional Wiener process and let ξ be a $\mathbb{R}^d$-valued random variable independent of W. Let $\mathcal{F}_t := \sigma(\xi, W_s : s \leq t)$. We consider a stochastic differential equation (SDE) of the form,

$$dX_t = b_t(X_t)\,dt + \sigma_t(X_t)\,dW_t\,, \qquad t \in [0,T]\,, \qquad X_0 = \xi\,. \tag{3.1}$$

Equivalently, we can write this in the integral form as

$$X_t = \xi + \int_0^t b_s(X_s)\,ds + \int_0^t \sigma_s(X_s)\,dW_s\,, \qquad t \in [0,T]\,. \tag{3.2}$$

Here $\sigma : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d'}$ and $b : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$. Written component-wise, the SDE is

$$dX_t^i = b^i(t, X_t)\,dt + \sum_{j=1}^{d'} \sigma^{ij}(t, X_t)\,dW_t^j\,, \quad t \in [0,T]\,, \quad X_0^i = \xi^i, \quad i \in \{1, \cdots, m\}.$$

The drift and volatility coefficients

$$(t, \omega, x) \mapsto \big(b_t(\omega, x), \sigma_t(\omega, x)\big)$$

are progressively measurable with respect to $\mathcal{F}_t \otimes \mathcal{B}(\mathbb{R}^d)$; as usual, we suppress ω in the notation and will typically write $b_t(x)$ instead of $b_t(\omega, x)$ etc. Note that $t = 0$ plays no special role in this; we may as well start the SDE at some time $t \geq 0$ (even a stopping time), and we shall write $X^{t,x} = (X_s^{t,x})_{s\in[t,T]}$ for the solution of the SDE started at time t with initial value x (assuming it exists and is unique).

Definition 3.1 (Solution of an SDE)**.** We say that a process X is a (strong) solution to the SDE (3.2) if

i) The process X is continuous on $[0,T]$ and adapted to $(\mathcal{F}_t)_{t\in[0,T]}$,

ii) we have

$$\mathbb{P}\left[\int_0^T |b_s(X_s)|\,ds < \infty\right] = 1 \ \text{ and } \ \mathbb{P}\left[\int_0^T |\sigma_s(X_s)|^2\,ds < \infty\right] = 1\,,$$

iii) The process X satisfies (3.1) almost surely for all $t \in [0,T]$ i.e. there is $\bar{\Omega} \in \mathcal{F}$ such that $\mathbb{P}(\bar{\Omega}) = 1$ and for all $\omega \in \bar{\Omega}$ it holds that

$$X_t(\omega) = \xi(\omega) + \int_0^t b_s(\omega, X_s(\omega))\,ds + \int_0^t \sigma_s(\omega, X_s(\omega))\,dW_s(\omega)\,,\ \forall t \in [0,T]\,. \tag{3.3}$$

Given $T \geq 0$, and $m \in \mathbb{N}$, we write $\mathbb{H}_T^m$ for the set of progressively measurable processes ϕ such that

$$\|\phi\|_{\mathbb{H}_T^m} := \left(\mathbb{E}\int_0^T |\phi_t|^m \, dt\right)^{\frac{1}{m}} < \infty.$$

Proposition 3.2 (Existence and uniqueness of solutions)**.** *Assume that for all $x \in \mathbb{R}^d$ the processes $(b_t(x))_{t\in[0,T]}$ and $(\sigma_t(x))_{t\in[0,T]}$ are progressively measurable, that $\mathbb{E}|\xi|^2 < \infty$ and that there exists a constant K such that a.s. for all $t \in [0,T]$ and $x, y \in \mathbb{R}^d$ it holds that*

$$\begin{aligned} \|b(0)\|_{\mathbb{H}_T^2} + \|\sigma(0)\|_{\mathbb{H}_T^2} &\leq K, \\ |b_t(x) - b_t(y)| + |\sigma_t(x) - \sigma_t(y)| &\leq K|x-y| \,. \end{aligned} \tag{3.4}$$

Then the SDE has a unique (strong) solution X on the interval $[0,T]$. Moreover, there exists a constant $C = C(K,T)$ such that

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X_t|^2\right] \leq C\big(1 + \mathbb{E}[\,|\xi|^2]\big).$$

We give an iterative scheme which we will show converges to the solution. To that end let $X_t^0 = \xi$ for all $t \in [0,T]$. For $n \in \mathbb{N}$ let, for $t \in [0,T]$, the process X^n be given by

$$X_t^n = \xi + \int_0^t b_s(X_s^{n-1})\, ds + \int_0^t \sigma_s(X_s^{n-1})\, dW_s \,. \tag{3.5}$$

Note that here the superscript on X indicates the iteration index.[6] We can see that X^0 is $(\mathcal{F}_t)_{t\in[0,T]}$-adapted and hence (due to progressive measurability of b and σ) X^1 is $(\mathcal{F}_t)_{t\in[0,T]}$-adapted and repeating this argument we see that each X^n is $(\mathcal{F}_t)_{t\in[0,T]}$-adapted.

Before we prove Proposition 3.2 by taking the limit in the above iteration we will need the following result.

Lemma 3.3. *Under the conditions of Proposition 3.2 there is a constant C, depending on K and T (but independent of n) such that for all $n \in \mathbb{N}$ and $t \in [0,T]$ it holds that*

$$\mathbb{E}|X_t^n|^2 < C(1 + \mathbb{E}|\xi|^2)e^{Ct} \,.$$

Proof. We see that

$$\mathbb{E}|X_t^n|^2 \leq 4\mathbb{E}|\xi|^2 + 4\mathbb{E}\left(\int_0^t |b_s(X_s^{n-1})|\, ds\right)^2 + 4\mathbb{E}\left(\int_0^t |\sigma_s(X_s^{n-1})|\, dW_s\right)^2 .$$

Using Hölder's inequality and Itô's isometry we can see that

$$\mathbb{E}|X_t^n|^2 \leq 4\mathbb{E}|\xi|^2 + 4t\mathbb{E}\int_0^t |b_s(X_s^{n-1})|^2\, ds + 4\mathbb{E}\int_0^t |\sigma_s(X_s^{n-1})|^2\, ds \,.$$

Using the Lipschitz continuity and growth assumption (3.4) we thus obtain that

$$\mathbb{E}\int_0^t |b_s(X_s^{n-1})|^2\, ds \leq 2\mathbb{E}\int_0^t |b_s(0)|^2\, ds + 2K^2\mathbb{E}\int_0^t |X_s^{n-1}|^2\, ds \leq 2K^2\left(1 + \mathbb{E}\int_0^t |X_s^{n-1}|^2\, ds\right)$$

[6] Instead of a power or index in a vector.

and similarly

$$\mathbb{E}\int_0^t |\sigma_s(X_s^{n-1})|^2\,ds \le 2K^2\left(1+\mathbb{E}\int_0^t |X_s^{n-1}|^2\,ds\right).$$

Thus for all $t\in[0,T]$ we have, with $L := 16K^2(t\vee 1)$, that

$$\mathbb{E}|X_t^n|^2 \le L\left(1+\mathbb{E}|\xi|^2\right) + L\int_0^t \mathbb{E}|X_s^{n-1}|^2\,ds\,.$$

Let us iterate this. For $n=1$ we have

$$\mathbb{E}|X_t^1|^2 \le L\left(1+\mathbb{E}|\xi|^2\right)+Lt\mathbb{E}|\xi|^2 \le L\left(1+\mathbb{E}|\xi|^2\right)+LtL(1+\mathbb{E}|\xi|^2) = L\left(1+\mathbb{E}|\xi|^2\right)[1+Lt]\,.$$

For $n=2$ we have

$$\begin{aligned}\mathbb{E}|X_t^2|^2 &\le L\left(1+\mathbb{E}|\xi|^2\right) + L\int_0^t \mathbb{E}|X_s^1|^2\,ds \le L\left(1+\mathbb{E}|\xi|^2\right) + L\cdot L(1+\mathbb{E}|\xi|^2)t + L\cdot\frac{(Lt)^2}{2}\\ &\le L(1+\mathbb{E}|\xi|^2)\left[1+Lt+\frac{(Lt)^2}{2}\right].\end{aligned}$$

If we carry on we see that

$$\mathbb{E}|X_t^n|^2 \le L(1+\mathbb{E}|\xi|^2)\left[1+Lt+\frac{(Lt)^2}{2!}+\cdots+\frac{(Lt)^n}{n!}\right] \le L(1+\mathbb{E}|\xi|^2)\left[\sum_{j=0}^\infty \frac{(Lt)^j}{j!}\right]$$

and hence for all $t\in[0,T]$ we have that

$$\mathbb{E}|X_t^n|^2 \le L(1+\mathbb{E}|\xi|^2)e^{Lt}\,.$$

$\square$

Proof of Proposition 3.2. We start with (3.5), take the difference between iteration $n+1$ and n, take the square of the $\mathbb{R}^d$ norm, take supremum and take the expectation. Then we see that

$$\begin{aligned}&\mathbb{E}\sup_{s\le t}|X_s^{n+1}-X_s^n|^2\\ &\le 2\mathbb{E}\sup_{s\le t}\left|\int_0^s [b_r(X_r^n)-b_r(X_r^{n-1})]\,dr\right|^2 + 2\mathbb{E}\sup_{s\le t}\left|\int_0^s [\sigma_r(X_r^n)-\sigma_r(X_r^{n-1})]\,dW_r\right|^2\\ &=: 2I_1(t)+2I_2(t)\,.\end{aligned}$$

We note that for all $t\in[0,T]$, having used Hölder's inequality in the penultimate step and assumption (3.4) in the final one, it holds that

$$\begin{aligned}I_1(t) &= \mathbb{E}\sup_{s\le t}\left|\int_0^s [b_r(X_r^n)-b_r(X_r^{n-1})]\,dr\right|^2 \le \mathbb{E}\sup_{s\le t}\left(\int_0^s |b_r(X_r^n)-b_r(X_r^{n-1})|\,dr\right)^2\\ &\le \mathbb{E}\left(\int_0^t |b_r(X_r^n)-b_r(X_r^{n-1})|\,dr\right)^2 \le t\mathbb{E}\int_0^t |b_r(X_r^n)-b_r(X_r^{n-1})|^2\,dr\\ &\le K^2 t\mathbb{E}\int_0^t |X_r^n-X_r^{n-1}|^2\,dr\,.\end{aligned}$$

Moreover $M_t = \int_0^t [\sigma_r(X_r^n) - \sigma_r(X_r^{n-1})]\,dW_r$ is a martingale and so $(|M_t|)_{t\in[0,T]}$ is a non-negative sub-martingale. Then Doob's maximal inequality, see Theorem A.14 with $p=2$, followed by Itô's isometry implies that for all $t \in [0,T]$ it holds that

$$
\begin{aligned}
I_2(t) &= \mathbb{E}\sup_{s\le t}\left|\int_0^s [\sigma_r(X_r^n) - \sigma_r(X_r^{n-1})]\,dW_r\right|^2 \le 4\mathbb{E}\left|\int_0^t [\sigma_r(X_r^n) - \sigma_r(X_r^{n-1})]\,dW_r\right|^2 \\
&= 4\mathbb{E}\int_0^t |\sigma_r(X_r^n) - \sigma_r(X_r^{n-1})|^2\,dr \le 4K^2\mathbb{E}\int_0^t |X_r^n - X_r^{n-1}|^2\,dr\,.
\end{aligned}
$$

Hence, with $L := 2K^2(T+4)$ we have for all $t\in[0,T]$ that

$$
\mathbb{E}\sup_{s\le t}|X_s^{n+1} - X_s^n|^2 \le L\int_0^t \mathbb{E}|X_r^n - X_r^{n-1}|^2\,dr\,. \tag{3.6}
$$

Let

$$
C^* := \sup_{t\in T}\mathbb{E}|X_t^1 - \xi|^2
$$

and note that Lemma 3.3 implies that $C^* < \infty$. Using this and iterating the estimate (3.6) we see that for all $t\in[0,T]$ we have that

$$
\mathbb{E}\sup_{s\le t}|X_s^{n+1} - X_s^n|^2 \le C^*\frac{L^n t^n}{n!}\,. \tag{3.7}
$$

For $f \in C([0,T];\mathbb{R}^d)$ let us define the norm $\|f\|_\infty := \sup_{s\in[0,T]}|f_s|$. Due to Chebychev–Markov's inequality we thus have

$$
\begin{aligned}
\mathbb{P}\left[\|X^{n+1} - X^n\|_\infty > \frac{1}{2^{n+1}}\right] &= \mathbb{P}\left[\|X^{n+1} - X^n\|_\infty^2 > \frac{1}{2^{2(n+1)}}\right] \\
&\le 4^{n+1}C^*\frac{L^n t^n}{n!} = 4C^*\frac{4^n L^n t^n}{n!}\,.
\end{aligned}
$$

Let $E_n := \{\omega\in\Omega : \|X^{n+1}(\omega) - X^n(\omega)\|_\infty > \frac{1}{2^{n+1}}\}$. Note that clearly[7] it holds that

$$
\sum_{n=0}^\infty \mathbb{P}E_n < \infty\,.
$$

By the Borel–Cantelli Lemma it thus holds that there is $\bar\Omega \in \mathcal{F}$ and a random variable $N : \Omega \to \mathbb{N}$ such that $\mathbb{P}(\bar\Omega) = 1$ and for all $\omega \in \bar\Omega$ we have that

$$
\|X^{n+1}(\omega) - X^n(\omega)\|_\infty \le 2^{-(n+1)} \quad \forall n \ge N(\omega)\,.
$$

For any $\omega\in\bar\Omega$, any $m\in\mathbb{N}$ and $n \ge N(\omega)$ we then have, due to the triangle inequality, that

$$
\begin{aligned}
\|X^{n+m}(\omega) - X^n(\omega)\|_\infty &\le \sum_{j=0}^{m-1}\|X^{n+j+1}(\omega) - X^{n+j}(\omega)\|_\infty \\
&\le \sum_{j=0}^{m-1} 2^{-(n+j+1)} = 2^{-(n+1)}\frac{1-\left(\frac12\right)^m}{1-\frac12} \le 2^{-n}\,.
\end{aligned} \tag{3.8}
$$

This means that the sequence $X^n(\omega)$ is a Cauchy sequence in the Banach space $C([0,T];\mathbb{R}^d)$ and thus a limit $X(\omega)$ such that $X^n(\omega) \to X(\omega)$ in $C([0,T];\mathbb{R}^d)$ as

[7] Indeed for any $x\in\mathbb{R}$ we have $\sum_{n=0}^\infty \frac{x^n}{n!} = e^x < \infty$.

$n \to \infty$. Moreover for each $n \in \mathbb{N}$ and each $t \in [0, T]$ the random variable X_t^n is $\mathcal{F}_t$ measurable which means that $X_t = \lim_{n\to\infty} X_t^n$ is $\mathcal{F}_t$ measurable.

Finally we have to show that the limit X satisfies the SDE. On the left hand side the convergence is trivial. To take the limit in the bounded variation integral we can use simply that for all $\omega \in \Omega$ we have that $\|X(\omega) - X^n(\omega)\|_\infty < 2^{-n}$ for $n \geq N(\omega)$. This follows by taking $m \to \infty$ in (3.8) with $n \in \mathbb{N}$ fixed. Then

$$\left| \int_0^t b_s(\omega, X_s^n(\omega))\, ds - \int_0^t b_s(\omega, X_s(\omega))\, ds \right| \leq K \int_0^t |X_s^n(\omega) - X_s(\omega)|\, ds \to 0$$

as $n \to \infty$ due to Lebesgue's theorem on dominated convergence.

To deal with the stochastic integral we need to do a bit more work. We see that for any $t \in [0, T]$ it holds that

$$\mathbb{E}|X_t^{n+m} - X_t^n|^2 = \mathbb{E}\left| \sum_{j=0}^{m-1} (X_t^{n+j+1} - X_t^{n+j}) 2^{-(n+j)} 2^{n+j} \right|^2 .$$

Using Hölder's inequality we get that for any $t \in [0, T]$ it holds that

$$\mathbb{E}|X_t^{n+m} - X_t^n|^2 \leq \left(\sum_{j=0}^{m-1} 4^{-(n+j)} \right) \left(\sum_{j=0}^{m-1} \mathbb{E}|X_t^{n+j+1} - X_t^{n+j}|^2\, 4^{n+j} \right) .$$

We note that

$$\sum_{j=0}^{m-1} 4^{-j} = \frac{1 - \left(\frac{1}{4}\right)^m}{1 - \frac{1}{4}} \leq \frac{4}{3}$$

From (3.7) we thus get that for all $m \in \mathbb{N}$ and for all $t \in [0, T]$ it holds that

$$\mathbb{E}|X_t^{n+m} - X_t^n|^2 \leq \frac{4}{3} C^* 4^{-n} \sum_{j=0}^{m-1} \frac{(4Lt)^{n+j}}{(n+j)!} \leq \frac{4}{3} C^* e^{4Lt} 4^{-n} .$$

Hence for any $t \in [0, T]$ the sequence $(X_t^n)_{n\in\mathbb{N}}$ is Cauchy in $L^2(\Omega)$ and so $X_t^n \to X_t$ in $L^2(\Omega)$ as $n \to \infty$ for all $t \in [0, T]$. Finally $\mathbb{E}|X_t|^2 \leq \liminf_{n\to\infty} \mathbb{E}|X_t^n|^2 \leq C(1+|\xi|^2)e^{Lt}$ due to Lemma 3.3. Thus for each $n \in \mathbb{N}$ we have

$$\mathbb{E}|X_t^n - X_t|^2 \leq 2\mathbb{E}|X_t^n|^2 + 2\mathbb{E}|X_t|^2 \leq 4C(1 + |\xi|^2)e^{Lt} =: g(t) .$$

Noting that $g \in L^1(0, T)$ we can conclude, using Lebesgue's theorem on dominated convergence that

$$\lim_{n\to\infty} \int_0^T \mathbb{E}|X_t^n - X_t|^2\, dt = \int_0^T \lim_{n\to\infty} \mathbb{E}|X_t^n - X_t|^2 = 0 .$$

This, together with Itô's isometry and assumption (3.4) allows us to take the limit in the stochastic integral term arising in (3.5). □

Remark 3.4. In the setup above the coefficients b and σ are random. In applications we will deal essentially with two settings for b and σ.

i) b and σ are deterministic, measurable, functions, i.e. $(t, x) \mapsto b_t(x)$ and $(t, x) \mapsto \sigma_t(x)$ are not random.

ii) b and σ are effectively random maps, but the randomness has a specific form. Namely, the random coefficients $b(t,\omega,x)$ and $\sigma(t,\omega,x)$ are of the form

$$b_t(\omega,x) := \bar{b}_t^{\alpha_t(\omega)}(x) \quad \text{and} \quad \sigma_t(\omega,x) := \bar{\sigma}_t^{\alpha_t(\omega)}(x)$$

where $\bar{b}$, $\bar{\sigma}$ are deterministic measurable functions on $[0,T]\times\mathbb{R}^d\times A$, A is a complete separable metric space and $(\alpha_t)_{t\in[0,T]}$ is a progressively measurable process valued in A.

This case arises in stochastic control problems that we will study later on, an example of which can already be seen with SDE (1.1).

Some properties of SDEs

In the remainder, we always assume that the coefficients b and σ satisfy the above conditions.

Proposition 3.5 (Further moment bounds)**.** *Let $m\in\mathbb{N}$, $m\geq 2$. Assume that for all $x\in\mathbb{R}^d$ the processes $(b_t(x))_{t\in[0,T]}$ and $(\sigma_t(x))_{t\in[0,T]}$ are progressively measurable, that $\mathbb{E}|\xi|^m<\infty$ and that there exists a constant K such that a.s. for all $t\in[0,T]$ and $x,y\in\mathbb{R}^d$ it holds that*

$$\|b(0)\|_{\mathbb{H}_T^m}+\|\sigma(0)\|_{\mathbb{H}_T^m}\leq K,$$
$$|b_t(x)-b_t(y)|+|\sigma_t(x)-\sigma_t(y)|\leq K|x-y|\,.$$

Assume that X is a solution of (3.1)*. Then there exists a constant $C=C(K,T,m,d,d')$ such that*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|X_t|^m\right]\leq C\big(1+\mathbb{E}[\,|\xi|^m]\big).$$

This can be proved using similar steps to those used in the proof of Lemma 3.3 but employing the Burkholder–Davis–Gundy inequality when estimating the expectation of the supremum of the stochastic integral term.

Proposition 3.6 (Stability)**.** *Let $m\in\mathbb{N}$, $m\geq 2$. Assume that for all $x\in\mathbb{R}^d$ the processes $(b_t(x))_{t\in[0,T]}$ and $(\sigma_t(x))_{t\in[0,T]}$ are progressively measurable, that $\mathbb{E}|\xi|^m<\infty$ and that there exists a constant K such that a.s. for all $t\in[0,T]$ and $x,y\in\mathbb{R}^d$ it holds that*

$$\|b(0)\|_{\mathbb{H}_T^m}+\|\sigma(0)\|_{\mathbb{H}_T^m}\leq K,$$
$$|b_t(x)-b_t(y)|+|\sigma_t(x)-\sigma_t(y)|\leq K|x-y|\,.$$

Let $x,x'\in\mathbb{R}^d$ and $0\leq t\leq t'\leq T$.

i) There exists a constant $C=C(K,T,m)$ such that

$$\mathbb{E}\left[\sup_{t\leq s\leq T}|X_s^{t,x}-X_s^{t,x'}|^m\right]\leq C|x-x'|^m.$$

ii) Suppose in addition that there is a constant K' such that

$$\mathbb{E}\left[\int_s^{s'} |b_r(0)|^2 + |\sigma_r(0)|^2 \, dr\right] \leq K'|s - s'|$$

for all $0 \leq s \leq s' \leq T$. Then there exists $C = C(K, T)$ such that

$$\mathbb{E}\left[\sup_{t' \leq s \leq T} |X_s^{t,x} - X_s^{t',x}|^2\right] \leq C(K + |x|^2)|t - t'|.$$

To prove the above two propositions one uses often the following inequalities: Cauchy-Schwartz, Hölder and Young's inequality; Gronwall's inequality (see Lemma A.6); Doob's maximal inequality (see Theorem A.14) and Burkholder–Davis–Gundy inequality.

Proposition 3.7 (Flow property). *Let $x \in \mathbb{R}^m$ and $0 \leq t \leq t' \leq T$. Then*

$$X_s^{t,x} = X_s^{t', X_{t'}^{t,x}}, \qquad s \in [t', T].$$

(This property holds even if t, t' are stopping times.)

See Exercise 3.17 for proof.

Proposition 3.8 (Markov property). *Let $x \in \mathbb{R}^d$ and $0 \leq t \leq t' \leq s \leq T$. If b and σ are* deterministic functions*, then*

$$X_s^{t,x} \quad \text{is a function of } t,\ x,\ s,\ \text{and } \big(W_r - W_t\big)_{r \in [t,s]}.$$

Moreover,

$$\mathbb{E}\left[\Phi\big(X_r^{t,x},\ t' \leq r \leq s\big)|\mathcal{F}_{t'}\right] = \mathbb{E}\left[\Phi\big(X_r^{t,x},\ t' \leq r \leq s\big)|X_{t'}^{t,x}\right]$$

for all bounded and measurable functions $\Phi : C^0([t', s]; \mathbb{R}^m) \to \mathbb{R}$.

On the left hand side (LHS), the conditional expectation is on $\mathcal{F}_{t'}$ that contains all the information from time $t = 0$ up to time $t = t'$. On the right hand side (RHS), that information is replaced by the process $X_{t'}^{t,x}$ at time $t = t'$. In words, for Markovian processes the best prediction of the future, given all knowledge of the present and past (what you see on the LHS), is the present (what you see on the RHS; all information on the past can be ignored).

3.2 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 complete separable metric space) and let $\mathcal{A}_0$ 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.

Assumption 3.9. Assume that for each $t \in [0,T]$ that $(x,a) \mapsto b(t,x,a)$ and $(x,a) \mapsto \sigma(t,x,a)$ are continuous, assume that for each $t \in [0,T]$ we have $x \mapsto b(t,x,a)$ and $x \mapsto \sigma(t,x,a)$ continuous in x uniformly in $a \in A$ and that there is a constant K such that for any t, x, y, a we have

$$|b(t,x,a) - b(t,y,a)| + |\sigma(t,x,a) - \sigma(t,y,a)| \leq K|x-y|. \tag{3.9}$$

Moreover for all t, x, a it holds that

$$|b(t,x,a)| + |\sigma(t,x,a)| \leq K(1 + |x| + |a|)\,. \tag{3.10}$$

We will fix $m \geq 2$ and refer to the set

$$\mathcal{A} := \{\alpha \in \mathbb{H}^m_T : \forall \omega \in \Omega, t \in [0,T]\ \alpha_t(\omega) \in A \ \text{ and } \ \alpha \ \text{ is progressively measurable}\,\}$$

set as *admissible controls.*

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

$$dX_s = b\big(s, X_s, \alpha_s\big)\,dt + \sigma\big(s, X_s, \alpha_s\big)\,dW_s, \quad X_t = \xi. \tag{3.11}$$

With Assumption 3.9 the SDE (3.11) is a special case of an SDE with random coefficients, see (3.1). In particular, if we fix $\alpha \in \mathcal{A}$ then taking $\tilde{b}_t(x) := b(t,x,\alpha_t)$ and $\tilde{\sigma}_t(x) := \sigma(t,x,\alpha_t)$ we have the progressive measurability of $\tilde{b}$ and $\tilde{\sigma}$ (since b, σ are assumed to be measurable and α is progressively measurable. Moreover

$$\|\tilde{b}(0)\|^2_{\mathbb{H}^2_T} = \mathbb{E}\int_0^T |b(t,0,\alpha_t)|^2\,dt \leq \mathbb{E}\int_0^T K^2(1+|\alpha_t)|)^2\,dt \leq 2K^2T + 2K^2\|\alpha\|^2_{\mathbb{H}^2_T} < \infty$$

and similarly $\|\tilde{\sigma}(0)\|^2_{\mathbb{H}^2_T} < \infty$. Finally the Lipschitz continuity of the coefficients in space clearly holds and so due to Proposition 3.2 we have the following result.

Proposition 3.10 (Existence and uniqueness)**.** *Let $t \in [0,T]$, $\xi \in L^2(\mathcal{F}_t)$ and $\alpha \in \mathcal{A}_0$. Then SDE* (3.11) *has a unique (strong) Markov solution $X = X^{t,\xi,\alpha}$ on the interval $[t,T]$ such that*

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

Moreover, the solution has the properties listed in Propositions 3.5 and 3.6.

3.3 Stochastic control problem with finite time horizon

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.

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

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

where $X^{t,\xi,\alpha}$ solves (3.11) (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 3.11. There is $K > 0$, $m \in \{0, 1, \ldots\}$ such that for all t, x, y, a we have

$$|g(x) - g(y)| + |f(t,x,a) - f(t,y,a)| \le K(1 + |x|^m + |y|^m)|x - y|,$$

$$|f(t,0,a)| \le K(1 + |a|^2)\,.$$

Note that this assumption is not the most general. For bigger generality consult e.g. [14].

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

$$(P)\begin{cases} v(t,x) := \sup\limits_{\alpha\in\mathcal{A}[t,T]} J(t,x,\alpha) = \sup\limits_{\alpha\in\mathcal{A}[t,T]} \mathbb{E}\left[\int_t^T f\big(s, X_s^{\alpha,t,x}, \alpha_s\big)\, ds + g\big(X_T^{\alpha,t,x}\big)\right] \\ \text{and } X^{\alpha,t,x} \text{ solves (3.11) with } X_t^{\alpha,t,x} = x. \end{cases}$$

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

$$v(t,x) \le \varepsilon + J(t,x,\alpha^\varepsilon)\,. \tag{3.12}$$

Lemma 3.13 (Lipschitz continuity in x of the value function)**.** *If Assumptions 3.9 and 3.11 hold and if $\mathcal{A} \subset \mathbb{H}_T^{m+1}$ then there exists $C = C_{T,K,m} > 0$ such that for all $t \in [0,T]$ and $x, y \in \mathbb{R}^d$ we have*

$$|J(t,x,\alpha) - J(t,y,\alpha)| \le C(1 + |x|^m + |y|^m)|x - y|\,.$$

and

$$|v(t,x) - v(t,y)| \le C(1 + |x|^m + |y|^m)|x - y|\,.$$

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

$$I := |J(t,x,\alpha) - J(t,y,\alpha)| \le C_T(1 + |x|^m + |y|^m)|x - y|\,.$$

Using Assumption 3.11 (local Lipschitz continuity in x of f and g) we get

$$
\begin{aligned}
I &\leq \mathbb{E}\left[\int_t^T |f(s, X_s^{t,x,\alpha}, \alpha_s) - f(s, X_s^{t,y,\alpha}, \alpha_s)|\, ds + \left|g(X_T^{t,x,\alpha}) - g(X_T^{t,x,\alpha})\right|\right] \\
&\leq K\mathbb{E}\bigg[\int_t^T (1 + |X_s^{t,x,\alpha}|^m + |X_s^{t,y,\alpha}|^m)|X_s^{t,x,\alpha} - X_s^{t,y,\alpha}|\, ds \\
&\qquad\qquad + (1 + |X_T^{t,x,\alpha}|^m + |X_T^{t,y,\alpha}|^m)|X_T^{t,x,\alpha} - X_T^{t,y,\alpha}|\bigg].
\end{aligned}
$$

We note that due to Hölder's inequality

$$
\begin{aligned}
I \leq & C_{K,m}\left(\mathbb{E}\int_t^T (1 + |X_s^{t,x,\alpha}|^{m+1} + |X_s^{t,y,\alpha}|^{m+1})\, ds\right)^{\frac{m}{m+1}} \left(\mathbb{E}\int_t^T |X_s^{t,x,\alpha} - X_s^{t,y,\alpha}|^{m+1}\, ds\right)^{\frac{1}{m+1}} \\
&+ C_{K,m}\left(\mathbb{E}(1 + |X_T^{t,x,\alpha}|^{m+1} + |X_T^{t,y,\alpha}|^{m+1})\right)^{\frac{m}{m+1}} \left(\mathbb{E}|X_T^{t,x,\alpha} - X_T^{t,y,\alpha}|^{m+1}\right)^{\frac{1}{m+1}}.
\end{aligned}
$$

Then, using Propositions 3.5 and 3.6, we get

$$
\begin{aligned}
I &\leq C_{T,K,m}\left(\sup_{t\leq s\leq T} \mathbb{E}\left[|X_s^{t,x,\alpha}|^{m+1} + |X_s^{t,y,\alpha}|^{m+1}\right]\right)^{\frac{m}{m+1}} \left(\sup_{t\leq s\leq T} \mathbb{E}|X_s^{t,x,\alpha} - X_s^{t,y,\alpha}|^{m+1}\right)^{\frac{1}{m+1}} \\
&\leq C_{T,K,m}(1 + |x|^m + |y|^m)|x - y|\,.
\end{aligned}
$$

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

$$
v(t,x) - v(t,y) \leq \varepsilon + J(t,x,\alpha^\varepsilon) - J(t,y,\alpha^\epsilon) \leq \varepsilon + C(1 + |x|^m + |y|^m)|x - y|\,.
$$

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

$$
v(t,y) - v(t,x) \leq \varepsilon + J(t,y,\beta^\varepsilon) - J(t,x,\beta^\epsilon) \leq \varepsilon + C(1 + |x|^m + |y|^m)|x - y|\,.
$$

Hence

$$
-\varepsilon - C(1 + |x|^m + |y|^m)|x - y| \leq v(t,x) - v(t,y) \leq \varepsilon + C(1 + |x|^m + |y|^m)|x - y|\,.
$$

Letting $\varepsilon \to 0$ concludes the proof. $\square$

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

3.4 Exercises

Exercise 3.14 (Non-existence of solution)**.**

1. Let $I = \left[0, \frac{1}{2}\right]$. Find a solution X for

$$
\frac{dX_t}{dt} = X_t^2\,,\;\; t \in I\,,\;\; X_0 = 1\,.
$$

2. Does a solution to the above equation exist on $I = [0, 1]$? If yes, show that it satisfies Definition 3.1. If not, which property is violated?

Exercise 3.15 (Non-uniqueness of solution)**.** Fix $T > 0$. Consider

$$\frac{dX_t}{dt} = 2\sqrt{|X_t|}\,, \quad t \in [0,T]\,, \quad X_0 = 0\,.$$

1. Show that $\bar{X}_t := 0$ for all $t \in [0,T]$ is a solution to the above ODE.

2. Show that $X_t := t^2$ for all $t \in [0,T]$ is also a solution.

3. Find at least two more solutions different from $\bar{X}$ and X.

Exercise 3.16. Consider the SDE

$$X_t = \xi + \int_0^t b_s(X_s)\,ds + \int_0^t \sigma_s(X_s)\,dW_s\,, \qquad t \in [0,T]\,. \tag{3.13}$$

and assume that the conditions of Proposition 3.2 hold. Show that the solution to the SDE is unique in the sense that if X and Y are two solutions with $X_0 = \xi = Y_0$ then

$$\mathbb{P}\left[\sup_{0\le t\le T} |X_t - Y_t| > 0\right] = 0\,.$$

Exercise 3.17. Consider the SDE

$$dX_s^{t,x} = b(X_s^{t,x})\,ds + \sigma(X_s^{t,x})\,dW_s, \;\; t \le s \le T, \;\; X_t^{t,x} = x\,.$$

Assume it has a pathwise unique solution i.e. if $Y_s^{t,x}$ is another process that satisfies the SDE then

$$\mathbb{P}\left[\sup_{t\le s\le T} |X_s^{t,x} - Y_s^{t,x}| > 0\right] = 0\,.$$

Show that then the *flow property* holds i.e. for $0 \le t \le t' \le T$ we have almost surely that

$$X_s^{t,x} = X_s^{t',X_{t'}^{t,x}}\,, \qquad \forall s \in [t',T].$$

3.5 Solutions to Exercises

Solution (to Exercise 3.14)**.**

1. We can use the following method to get a guess: from the ODE we get $X^{-2}dX = dt$ which means that, after integrating, we get $-X^{-1} = t + C$. So $X_t = -(t+C)^{-1}$. Since $X_0 = 1$ we get $C = -1$. Thus

$$X_t = \frac{1}{1-t}\,, \quad t \in \left[0, \tfrac{1}{2}\right]\,.$$

   We check by calculating that $\frac{dX_t}{dt} = (1-t)^{-2} = X_t^2$ so the equation holds in $[0, 1/2]$.

2. We can see that $\lim_{t\nearrow 1} X_t = \infty$ and so the $t \mapsto X_t$ is not continuous on $[0,1]$.

Solution (to Exercise 3.15)**.**

1. Clearly $\bar{X}_0 = 0$ and for $t \in [0,T]$ we have $\frac{d\bar{X}_t}{dt} = 0 = 2\sqrt{|\bar{X}_t|}$.

2. Clearly $X_0 = 0$ and for $t \in [0,T]$ we have $\frac{dX_t}{dt} = 2t = 2\sqrt{t^2} = 2\sqrt{|X_t|}$.
3. Fix any $\tau \in (0,T)$ and define

$$X_t^{(\tau)} := \begin{cases} 0 & \text{for} \quad t \in [0,\tau)\,, \\ (t-\tau)^2 & \text{for} \quad t \in [\tau,T]\,. \end{cases}$$

Then, clearly, $dX_0^{(\tau)} = 0$ and moreover if $t \in [0,\tau)$ then we have

$$\frac{dX_t^{(\tau)}}{dt} = 0 = 2\sqrt{|X_t^{(\tau)}|}\,,$$

while if $t \in [\tau,T]$ then we have

$$\frac{dX_t^{(\tau)}}{dt} = 2(t-\tau) = 2\sqrt{|(t-\tau)^2|} = 2\sqrt{X_t^{(\tau)}}\,.$$

So, in fact, there are uncountably many different solutions.

Solution (to Exercise 3.16)**.** We will use exactly the same estimates as in the proof of Proposition 3.2, up to (3.6). We start with (3.13) written once for X and once for Y and take the difference, take the square of the $\mathbb{R}^d$ norm, take supremum and take the expectation. Then we see that

$$\begin{aligned}&\mathbb{E}\sup_{s\le t}|X_s - Y_s|^2 \\ &\le 2\mathbb{E}\sup_{s\le t}\left|\int_0^s [b_r(X_r) - b_r(Y_r)]\,dr\right|^2 + 2\mathbb{E}\sup_{s\le t}\left|\int_0^s [\sigma_r(X_r) - \sigma_r(Y_r)]\,dW_r\right|^2 \\ &=: 2I_1(t) + 2I_2(t)\,.\end{aligned}$$

We note that for all $t \in [0,T]$, having used Hölder's inequality in the penultimate step and assumption (3.4) in the final one, it holds that

$$\begin{aligned} I_1(t) &= \mathbb{E}\sup_{s\le t}\left|\int_0^s [b_r(X_r) - b_r(Y_r)]\,dr\right|^2 \le \mathbb{E}\sup_{s\le t}\left(\int_0^s |b_r(X_r) - b_r(Y_r)|\,dr\right)^2 \\ &\le \mathbb{E}\left(\int_0^t |b_r(X_r) - b_r(Y_r)|\,dr\right)^2 \le t\mathbb{E}\int_0^t |b_r(X_r) - b_r(Y_r)|^2\,dr \\ &\le K^2 t\mathbb{E}\int_0^t |X_r - Y_r|^2\,dr\,. \end{aligned}$$

Moreover $M_t = \int_0^t [\sigma_r(X_r) - \sigma_r(Y_r)]\,dW_r$ is a martingale and so $(|M_t|)_{t\in[0,T]}$ is a non-negative sub-martingale. Then Doob's maximal inequality, see Theorem A.14 with $p = 2$, followed by Itô's isometry implies that for all $t \in [0,T]$ it holds that

$$\begin{aligned} I_2(t) &= \mathbb{E}\sup_{s\le t}\left|\int_0^s [\sigma_r(X_r) - \sigma_r(Y_r)]\,dW_r\right|^2 \le 4\mathbb{E}\left|\int_0^t [\sigma_r(X_r) - \sigma_r(Y_r)]\,dW_r\right|^2 \\ &= 4\mathbb{E}\int_0^t |\sigma_r(X_r) - \sigma_r(Y_r)|^2\,dr \le 4K^2\mathbb{E}\int_0^t |X_r - Y_r|^2\,dr\,. \end{aligned}$$

Hence, with $L := 2K^2(T+4)$ we have for all $t \in [0,T]$ that

$$\mathbb{E}\sup_{s\le t}|X_s - Y_s|^2 \le L\int_0^t \mathbb{E}|X_r - Y_r|^2\,dr\,.$$

Hence

$$\mathbb{E}\sup_{s\le t}|X_s - Y_s|^2 \le L\int_0^t \mathbb{E}\sup_{s\le r}|X_s - Y_s|^2\,dr\,.$$

From Gronwall's lemma (applied with $y(t) := \mathbb{E}\sup_{s\leq t}|X_s - Y_s|^2$, $a(t) = 0$, $b(t) = 0$ and $\lambda(t) = L$) we get that for all $t \in [0,T]$ we have

$$\mathbb{E}\sup_{s\leq t}|X_s - Y_s|^2 \leq 0\,.$$

But this means that

$$\mathbb{P}\left[\sup_{t\leq T}|X_s - Y_s|^2 = 0\right] = 1\,.$$

Solution (to Exercise 3.17)**.** Let $Y_s := X_s^{t',X_{t'}^{t,x}}$ for $s \in [t',T]$ and note that the process Y solves the SDE for $s \in [t',T]$ with $Y_{t'} = X_{t'}^{t',X_{t'}^{t,x}} = X_{t'}^{t,x}$. Let $X_s := X_s^{t,x}$ for $s \in [t',T]$ and note that this also solves the SDE for $s \in [t',T]$ with

$$X_{t'} = X_{t'}^{t,x} = Y_{t'}\,.$$

Hence both Y and X solve the same SDE with the same starting point. By the pathwise uniqueness property of the solutions of this SDE we then have

$$\mathbb{P}\left[\sup_{t\leq s\leq T}|X_s - Y_s| = 0\right] = 1$$

but this means that almost surely it holds that for all $s \in [t',T]$ it holds that

$$X_s^{t',X_{t'}^{t,x}} = Y_s = X_s = X_s^{t,x}\,.$$

4 Dynamic programming and the Hamilton–Jacobi–Bellman equation

4.1 Dynamic programming principle

Dynamic programming (DP) is one of the most popular approaches to study the stochastic control problem (P). The main idea was originated from the so-called Bellman's principle, which states

> *An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.*

The following is the statement of Bellman's principle / dynamic programming.

Theorem 4.1 (Bellman's principle / Dynamic programming)**.** *For any* $0 \le t \le \hat{t} \le T$*, for any* $x \in \mathbb{R}^d$*, we have*

$$v(t,x) = \sup_{\alpha \in \mathcal{A}[t,\hat{t}]} \mathbb{E}\left[\int_t^{\hat{t}} f\big(s, X_s^{\alpha,t,x}, \alpha_s\big)\, ds + v\big(\hat{t}, X_{\hat{t}}^{\alpha,t,x}\big)\Big| X_t^{\alpha,t,x} = x\right]. \tag{4.1}$$

The idea behind the dynamic programming principle is as follows. The expectation on the RHS of (4.1) represents the gain if we implement the time t until time $\hat{t}$ optimal strategy and then implement the time $\hat{t}$ until T optimal strategy. Clearly, this gain will be no larger than the gain associated with using the overall optimal strategy from the start (since we can apply the overall optimal control in both scenarios and obtain the LHS).

What equation (4.1) says is that if we determine the optimal strategy separately on each of the time intervals $[t,\hat{t}]$ and $[\hat{t},T]$ we get the same answer as when we consider the whole time interval $[t,T]$ at once. Underlying this statement, hides a deeper one: that if one puts the optimal stategy over $[t,\hat{t}\,)$ together with the optimal stategy over $[\hat{t},T]$ this is still an optimal strategy.

Note that without Lemma 3.13 we would not even be allowed to write (4.1) since we need $v(\hat{t}, X_{\hat{t}}^{\alpha,t,x})$ to be a random variable (so that we are allowed to take the expectation).

Let us now prove the Bellman principle.

Proof of Theorem 4.1. We will start by showing that $v(t,x) \le$ RHS of (4.1). We note that with $\alpha \in \mathcal{A}[t,T]$ we have

$$J(t,x,\alpha) = \mathbb{E}\left[\int_t^{\hat{t}} f(s,X_s^\alpha,\alpha_s)ds + \int_{\hat{t}}^T f(s,X_s^\alpha,\alpha_s)ds + g(X_T^\alpha)\Big| X_t^\alpha = x\right].$$

We will use the tower property of conditional expectation and use the Markov property of the process. Let $\mathcal{F}_{\hat{t}}^{X^\alpha} := \sigma(X_s^\alpha : t \le s \le \hat{t})$. Then

$$\begin{aligned}
&J(t,x,\alpha)\\
&= \mathbb{E}\left[\int_t^{\hat{t}} f(s,X_s^\alpha,\alpha_s)ds + \mathbb{E}\left[\int_{\hat{t}}^T f(s,X_s^\alpha,\alpha_s)ds + g(X_T^\alpha)\Big|\mathcal{F}_{\hat{t}}^{X^\alpha}\right]\Big| X_t^\alpha = x\right]\\
&= \mathbb{E}\left[\int_t^{\hat{t}} f(s,X_s^\alpha,\alpha_s)ds + \mathbb{E}\left[\int_{\hat{t}}^T f(s,X_s^\alpha,\alpha_s)ds + g(X_T^\alpha)\Big| X_{\hat{t}}^\alpha\right]\Big| X_t^\alpha = x\right].
\end{aligned}$$

Now, because of the flow property of SDEs,

$$\mathbb{E}\left[\int_{\hat{t}}^{T} f(s, X_s^{\alpha,t,x}, \alpha_s)ds + g(X_T^{\alpha,t,x})\middle| X_{\hat{t}}^{\alpha,t,x}\right] = J\left(\hat{t}, X_{\hat{t}}^{\alpha,t,x}, (\alpha_s)_{s\in[\hat{t},T]}\right) \leq v\left(\hat{t}, X_{\hat{t}}^{\alpha,t,x}\right).$$

Hence

$$J(t,x,\alpha) \leq \sup_{\alpha\in\mathcal{A}} \mathbb{E}\left[\int_t^{\hat{t}} f(s, X_s^{\alpha}, \alpha_s)ds + v\left(\hat{t}, X_{\hat{t}}^{\alpha,t,x}\right)\middle| X_t^{\alpha} = x\right].$$

Taking supremum over control processes α on the left shows that $v(t,x) \leq$ RHS of (4.1).

We now need to show that RHS of (4.1) $\leq v(t,x)$. Fix $\varepsilon > 0$. Then there is $\alpha^\varepsilon \in \mathcal{A}[t,\hat{t}]$ such that

$$\text{RHS of (4.1)} \leq \varepsilon + \mathbb{E}\left[\int_t^{\hat{t}} f\big(s, X_s^{\alpha^\varepsilon,t,x}, \alpha_s^\varepsilon\big)\, ds + v\big(\hat{t}, X_{\hat{t}}^{\alpha^\varepsilon,t,x}\big)\middle| X_t^{\alpha^\varepsilon,t,x} = x\right].$$

Let us write $X_s := X_s^{\alpha^\varepsilon,t,x}$ for brevity from now on. We now have to be careful so that we can construct an ε-optimal control which is progressively measurable on the whole $[t,T]$. To that end let $\delta = \delta(\omega) > 0$ be such that

$$2^m C(1 + |X_{\hat{t}}(\omega)|^m)\delta(\omega) < \varepsilon \;\text{ and }\; 2^{m-1}\delta(\omega)^m < 1.$$

where C is the constant from Lemma 3.13. Take $(x_i)_{i\in\mathbb{N}}$ dense in $\mathbb{R}^d$. By density of $(x_i)_i$ we know that for each $\delta(\omega)$ there exists $i(\omega)$ such that $|x_{i(\omega)} - X_{\hat{t}}(\omega)| \leq \delta(\omega)$. Moreover

$$C(1+|x_i|^m)\delta \leq C(1 + 2^{m-1}|x_i - X_{\hat{t}}|^m + 2^{m-1}|X_{\hat{t}}|^m)\delta \leq 2^m C(1+|X_{\hat{t}}|^m)\delta < \varepsilon.$$

The open covering of $\mathbb{R}^d$ given by $\bigcup_{\omega\in\Omega} B_{\delta(\omega)}(x_{i(\omega)})$ has a countable sub-cover $\bigcup_{k\in\mathbb{N}} B_{\delta_k}(x_k)$. Let (Q_k) be constructed as follows:

$$Q_1 = B_{\delta_1}(x_1) \;\text{ and }\; Q_k = B_{\delta_k}(x_k) \setminus \bigcup_{k'=1}^{k-1} Q_{k'}.$$

Then for each x_i there is $\alpha^{\varepsilon,i} \in \mathcal{A}(\hat{t},T]$ such that $v(\hat{t},x_i) \leq \varepsilon + J(\hat{t},x_i,\alpha^{\varepsilon,i})$. Moreover if $X_{\hat{t}} \in Q_i$ then $|X_{\hat{t}}|^m \leq 2^{m-1}|X_{\hat{t}} - x_i|^m + 2^{m-1}|x_i|^m$ and due to Lemma 3.13 we have,

$$\begin{aligned}
|v(\hat{t},X_{\hat{t}}) - v(\hat{t},x_i)| &\leq C(1+|X_{\hat{t}}|^m + |x_i|^m)|X_{\hat{t}} - x_i| \\
&\leq C(1 + 2^{m-1}|X_{\hat{t}} - x_i|^m + 2^{m-1}|x_i|^m)|X_{\hat{t}} - x_i| \\
&\leq C(1 + 2^{m-1}\delta^m + 2^{m-1}|x_i|^m)\delta \\
&\leq C(2 + 2^{m-1}|x_i|^m)\delta \leq 2^m C(1+|x_i|^m)\delta < \varepsilon.
\end{aligned}$$

Similarly we have

$$|J(\hat{t},x_i,\alpha^{\varepsilon,i}) - J(\hat{t},X_{\hat{t}},\alpha^{\varepsilon,i})| \leq \varepsilon.$$

Hence we get

$$v(\hat{t},X_{\hat{t}}) \leq v(\hat{t},x_i) + \varepsilon \leq \varepsilon + J(\hat{t},x_i,\alpha^{\varepsilon,i}) + \varepsilon \leq \varepsilon + J(\hat{t},X_{\hat{t}},\alpha^{\varepsilon,i}) + 2\varepsilon.$$

And so

$$v(\hat{t},X_{\hat{t}}) \leq 3\varepsilon + J(\hat{t},X_{\hat{t}},\alpha^{\varepsilon,i}).$$

Therefore RHS of (4.1)

$$\le 3\varepsilon + \mathbb{E}\Bigg[\int_t^{\hat t} f\big(s, X_s^{\alpha^\varepsilon,t,x}, \alpha_s^\varepsilon\big)\,ds$$

$$+\,\mathbb{E}\left[\int_{\hat t}^T f(s, Y_s^{\alpha^{\varepsilon,i}}, \alpha_s^{\varepsilon,i})\,ds + g\left(Y_T^{\alpha^{\varepsilon,i}}\right)\middle|Y_{\hat t}^{\alpha^{\varepsilon,i}} = X_{\hat t}^{\alpha^\varepsilon,t,x}\right]\Bigg|X_s^{\alpha^\varepsilon,t,x} = x\Bigg].$$

Regarding controls we now have the following: $\alpha^\varepsilon \in \mathcal{A}[t,\hat t]$ and for each i we have $\alpha^{\varepsilon,i} \in \mathcal{A}(\hat t, T]$. From these we build one control process β^ε as follows:

$$\beta_s^\varepsilon := \begin{cases} \alpha_s^\varepsilon & s \in [t,\hat t] \\ \alpha_s^{\varepsilon,i} & s \in (\hat t, T] \text{ and } X_{\hat t}^{\alpha^\varepsilon,t,x} \in Q_i. \end{cases}$$

This process is progressively measurable with values in A and so $\beta^\varepsilon \in \mathcal{A}[t,T]$. Due to the flow property we may write that RHS of (4.1)

$$\le 3\varepsilon + \mathbb{E}\left[\int_t^{\hat t} f\big(s, X_s^{\beta^\varepsilon,t,x}, \beta_s^\varepsilon\big)\,ds + \int_{\hat t}^T f(s, X_s^{\beta^\varepsilon}, \beta_s^\varepsilon)\,ds + g\left(X_T^{\beta^\varepsilon}\right)\middle|X_s^{\beta^\varepsilon,t,x} = x\right].$$

Finally taking supremum over all possible control strategies we see that RHS of (4.1) $\le 3\varepsilon + v(t,x)$. Letting $\varepsilon \to 0$ completes the proof. □

Lemma 4.2 ($\frac{1}{2}$-Hölder continuity of value function in time)**.** *Let Assumptions 3.9 and 3.11 hold. Let $\mathcal{A} \subseteq \mathbb{H}_T^{m+1}$. Then there is a constant $C = C_{T,K,m} > 0$ such that for any $x \in \mathbb{R}^d$, $0 \le t, \hat t \le T$ we have*

$$|v(t,x) - v(\hat t, x)| \le C(1 + |x|^m)|t - \hat t|^{1/2}.$$

Proof. Still needs to be written down.

□

Corollary 4.3. *Let Assumptions 3.9 and 3.11 hold. Let $\mathcal{A} \subseteq \mathbb{H}_T^{m+1}$. Then there is a constant $C = C_{T,K,m} > 0$ such that for any $x, y \in \mathbb{R}^d$, $0 \le s, t \le T$ we have*

$$|v(s,x) - v(t,y)| \le C(1 + |x|^m + |y|^m)\left(|t - \hat t|^{1/2} + |x - y|\right).$$

This means that the value function v is jointly measurable in (t,x). With this we get the following.

Theorem 4.4 (Bellman's principle / Dynamic programming with stopping time)**.** *For any stopping times $t, \hat t$ such that $0 \le t \le \hat t \le T$, for any $x \in \mathbb{R}^d$, we have* (4.1).

The proof uses the same arguments as before except that now have to cover the whole $[0,T] \times \mathbb{R}^d$ and we need to use the $\frac{1}{2}$-Hölder continuity in time as well.

Corollary 4.5 (Global optimality implies optimality from any time)**.** *Take $x \in \mathbb{R}^d$. A control $\beta \in \mathcal{A}[0,T]$ is optimal for (P) with the state process $X_s = X_s^{\beta,0,x}$ for $s \in [0,T]$ if and only if for any $\hat t \in [0,T]$ we have*

$$v(\hat t, X_{\hat t}) = J\Big(\hat t, X_{\hat t}, (\beta_r)_{r \in [\hat t, T]}\Big).$$

Proof. To ease the notation we will take $f = 0$. The reader is encouraged to prove the general case.

Due to the Bellman principle, Theorem 4.4, we have

$$v(0,x) = \sup_{\alpha \in \mathcal{A}[0,\hat{t}]} \mathbb{E}\left[v\left(\hat{t}, X_{\hat{t}}^{\alpha,0,x}\right)\right] \geq \mathbb{E}\left[v\left(\hat{t}, X_{\hat{t}}^{\beta,0,x}\right)\right]. \tag{4.2}$$

If β is an optimal control

$$v(0,x) = J(0,x,\beta) = \mathbb{E}\left[g\left(X_T\right)\right],$$

where the first equality follows from β being assumed to be an optimal control and second equality is definition of J. From this, using the tower property of conditional expectation, we see

$$v(0,x) = \mathbb{E}\left[\mathbb{E}\left[g\left(X_T\right)\middle|\mathcal{F}_{\hat{t}}^X\right]\right] = \mathbb{E}\left[J\left(\hat{t}, X_{\hat{t}}, \beta\right)\right] \leq \mathbb{E}\left[v\left(\hat{t}, X_{\hat{t}}\right)\right] \leq v(0,x),$$

where the last inequality is (4.2) again. Since the very left and very right of these inequalities are equal we get that

$$\mathbb{E}\left[J\left(\hat{t}, X_{\hat{t}}, \beta\right)\right] = \mathbb{E}\left[v\left(\hat{t}, X_{\hat{t}}\right)\right]$$

Moreover $v \geq J$ and so we can conclude that $v\left(\hat{t}, X_{\hat{t}}\right) = J\left(\hat{t}, X_{\hat{t}}, \beta\right)$ a.s. The completes the first part of the proof. The "only if" part of the proof is clear because we can take $\hat{t} = 0$ and get $v(0,x) = J(0,x,\beta)$ which means that β is an optimal control. $\square$

From this observation we can prove the following description of optimality.

Theorem 4.6 (Martingale optimality)**.** *Let the assumptions required for Bellman's principle hold. For any (t,x) and $\alpha \in \mathcal{A}$ let*

$$M_s^{t,x,\alpha} := \int_t^s f_r^{\alpha_r}\left(X_r^{t,x,\alpha}\right) dr + v\big(s, X_s^{t,x,\alpha}\big). \tag{4.3}$$

Then the process $(M_s^{t,x,\alpha})_{s\in[t,T]}$ is an $\mathcal{F}_s^X := \sigma(X_r^{\alpha,t,x}; t \leq r \leq s)$ super-martingale. Moreover α is optimal if and only if it is a martingale.

Proof. We have by, Theorem 4.1 (the Bellman principle) that for any $0 \leq t \leq s \leq \hat{s} \leq T$ that

$$v\big(s, X_s^{t,x,\alpha}\big) = \sup_{\hat{\alpha}\in\mathcal{A}} \mathbb{E}\left[\int_s^{\hat{s}} f_r^{\hat{\alpha}_r}\left(X_r^{s,X_s^{t,x,\alpha},\hat{\alpha}}\right) dr + v\big(\hat{s}, X_{\hat{s}}^{s,X_s^{t,x,\alpha},\hat{\alpha}}\big)\middle| X_s^{\alpha,t,x}\right].$$

From the Markov property we get that

$$v\big(s, X_s^{t,x,\alpha}\big) = \sup_{\hat{\alpha}\in\mathcal{A}} \mathbb{E}\left[\int_s^{\hat{s}} f_r^{\hat{\alpha}_r}\left(X_r^{s,X_s^{t,x,\alpha},\hat{\alpha}}\right) dr + v\big(\hat{s}, X_{\hat{s}}^{s,X_s^{t,x,\alpha},\hat{\alpha}}\big)\middle| \mathcal{F}_s^X\right].$$

Hence

$$v\big(s, X_s^{t,x,\alpha}\big) \geq \mathbb{E}\left[\int_s^{\hat{s}} f_r^{\alpha_r}\left(X_r^{s,X_s^{t,x,\alpha},\alpha}\right) dr + v\big(\hat{s}, X_{\hat{s}}^{s,X_s^{t,x,\alpha},\alpha}\big)\middle| \mathcal{F}_s^X\right]$$

and so, using the flow property of the SDE,

$$M_s^{t,x,\alpha} \geq \int_t^s f_r^{\alpha_r}\left(X_r^{t,x,\alpha}\right) dr + \mathbb{E}\left[\int_s^{\hat{s}} f_r^{\alpha_r}\left(X_r^{t,x,\alpha}\right) dr + v\big(\hat{s}, X_{\hat{s}}^{t,x,\alpha}\big)\Big|\mathcal{F}_s^X\right]$$
$$= \mathbb{E}\left[M_{\hat{s}}^{t,x,\alpha}|\mathcal{F}_s^X\right] .$$

This means that $M^{t,x,\alpha}$ is a super-martingale. Moreover we see that if α is optimal then the inequalities above are equalities and hence $M^{t,x,\alpha}$ is a martingale.

Now assume that $M_s^{t,x,\alpha} = \mathbb{E}[M_{\hat{s}}^{t,x,\alpha}|\mathcal{F}_s^X]$. We want to ascertain that the control α driving $M^{t,x,\alpha}$ is an optimal one. But the martingale property implies that $J(t,x,\alpha) = \mathbb{E}[M_T^{t,x,\alpha}] = \mathbb{E}[M_t^{t,x,\alpha}] = v(t,x)$ and so α is indeed an optimal control. $\square$

One question you may ask yourself is: How can we use the dynamic programming principle to compute an optimal control? Remember that the idea behind the DPP is that it is not necessary to optimize the control α over the entire time interval $[0,T]$ at once; we can partition the time interval into smaller sub-intervals and optimize over each individually. We will see below that this idea becomes particularly powerful if we let the partition size go to zero: the calculation of the optimal control then becomes a pointwise minimization linked to certain PDEs (see Theorem A.29). That is, for each fixed state x we compute the optimal value of control, say $a \in A$, to apply whenever $X(t) = x$.

4.2 Hamilton-Jacobi-Bellman (HJB) and verification

If the value function $v = v(t,x)$ is smooth enough, then we can apply Itô's formula to v and X in (4.3). Thus we get the *Hamilton-Jacobi-Bellman* (HJB) equation (also know and the *Dynamic Programming equation* or *Bellman* equation).

For notational convenience we will write $\sigma^a(t,x) := \sigma(t,x,a)$, $b^a(t,x) := b(t,x,a)$ and $f^a(t,x) := f(t,x,a)$. We then define

$$L^a v := \frac{1}{2}\mathrm{tr}\left[\sigma^a(\sigma^a)^\top \partial_{xx} v\right] + b^a \partial_x v \, .$$

Recall that trace is the sum of all the elements on the diagonal of a square matrix i.e. for a matrix $(a^{ij})_{i,j=1}^d$ we get $\mathrm{tr}[a] = \sum_{i=1}^d a^{ii}$, that $\partial_{xx} v$ denotes the Jacobian matrix i.e. $(\partial_{xx} v)_{ij} = \partial_{x^i}\partial_{x^j} v$ whilst $\partial_x v$ denotes the gradient vector i.e. $(\partial_x v)_i = \partial_{x^i} v$. This means that

$$\mathrm{tr}\left[\sigma^a(\sigma^a)^\top \partial_{xx} v\right] = \sum_{i,j=1}^d [\sigma^a(\sigma^a)^\top]^{ij} \partial_{x^i x^j} v \;\text{ and }\; b^a \partial_x v = \sum_{i=1}^d (b^a)^i \partial_{x^i} v \, .$$

Theorem 4.7 (Hamilton-Jacobi-Bellman (HJB))**.** *Assume that the Bellman principle holds. Assume that b and σ are bounded and continuous in (t,x). If the value function v for (P) is $C^{1,2}([0,T) \times \mathbb{R}^d)$, then it satisfies*

$$\begin{aligned} \partial_t v + \sup_{a \in A}\Big(L^a v + f^a\Big) &= 0 \quad \text{on } [0,T) \times \mathbb{R}^d \\ v(T,x) &= g(x) \quad \forall x \in \mathbb{R}^d \, . \end{aligned} \tag{4.4}$$

Proof. Let $x \in \mathbb{R}$ and $t \in [0,T]$. Then the condition $v(T,x) = g(x)$ follows directly from the definition of v. Fix $\alpha \in \mathcal{A}[t,T]$ and let M be given by (4.3) i.e.

$$M_s := \int_t^s f_r^{\alpha_r}\left(X_r^{t,x,\alpha}\right) dr + v\big(s, X_s^{t,x,\alpha}\big)\,.$$

Then, Itô's formula applied to v and $X = (X_s^{t,x,\alpha})_{s\in[t,T]}$ yields

$$dM_s = \Big[\Big(\partial_t v + L^{\alpha_s} v + f^{\alpha_s}\Big)\big(s, X_s^{\alpha,t,x}\big)\Big]\,ds + \Big[\big(\partial_x v\, \sigma^{\alpha_s}\big)\big(s, X_s^{\alpha,t,x})\big)\Big]\,dW_s.$$

For any $(t,x) \in [0,T]\times\mathbb{R}$ and $R > 0$ take the stopping time $\tau_R = \tau_R^{\alpha,t,x}$ given by

$$\tau := \inf\Big\{t' \geq t : \int_t^{t'} (\partial_x v\, \sigma^{\alpha_s})\big(s, X_s^{\alpha,t,x})\big)^2\,ds \geq R\Big\}\,.$$

We know from Theorem 4.6 that M must be a supermartingale. On the other hand the term given by the stochastic integral is a martingale (when cosidered stopped at τ_R). So $(M_{t\wedge\tau})_t$ can only be a supermartingale if for any stopping time $\hat\tau \leq T$ we have

$$\int_t^{\hat\tau\wedge\tau_R} f^{\alpha_s}(s,X_s) + (\partial_t v + L^{\alpha_s}v)(s,X_s)\,ds \leq 0\,.$$

Taking $R\to\infty$ we can replace $\hat\tau \wedge \tau_R$ by just $\hat\tau$. Since the starting point (t,x) and control α were arbitrary we get, using continuity of $\partial_t v + L^a v + f^a$ in (t,x) and a stopping time argument (see Exercise 4.9), that

$$(\partial_t v + L^a v + f^a)(t,x) \leq 0 \;\;\forall t,x,a\,.$$

Taking the supremum over $a \in A$ we get

$$\partial_t v(t,x) + \sup_{a\in A}[(L^a v + f^a)(t,x)] \leq 0 \;\;\forall t,x\,.$$

We now need to show that in fact the inequality cannot be strict. We proceed by setting up a contradiction. Assume that there is (t,x) such that

$$\partial_t v(t,x) + \sup_{a\in A}[(L^a v + f^a)(t,x)] < 0\,.$$

We will show that this contradicts the Bellman principle and hence we must have equality, thus completing the proof.

Now by continuity (recall that $v \in C^{1,2}([0,T)\times\mathbb{R}^d)$) we get that there must be $\varepsilon > 0$ and an associated $\delta > 0$ such that

$$\partial_t v + \sup_{a\in A}[(L^a v + f^a)] \leq -\varepsilon < 0 \;\text{ on } [t,t+\delta)\times B_\delta(x).$$

Let us fix $\alpha \in \mathcal{A}[t,T]$ and let $X_s := X_s^{t,x,\alpha}$. We define the stopping time

$$\tau = \tau^{t,x,\alpha} := \inf\{s > t : |X_s^{t,x,\alpha} - x| > \delta\} \wedge (t+\delta)\,.$$

Then

$$\begin{aligned}
&\int_t^\tau f^{\alpha_r}(r,X_r)\,dr + v(\tau,X_\tau)\\
&= v(t,x) + \int_t^\tau f^{\alpha_r}(r,X_r)\,dr + v(\tau,X_\tau) - v(t,x)\\
&= v(t,x) + \int_t^\tau \Big[\Big(\partial_t v + L^{\alpha_r}v + f^{\alpha_r}\Big)(r,X_r)\Big]\,dr + \int_t^\tau \Big[(\partial_x v)\,\sigma^{\alpha_r}\Big]\big(r,X_r\big)\,dW_r\\
&\leq v(t,x) - \varepsilon(\tau - t) + \int_t^\tau \Big[(\partial_x v)\,\sigma^{\alpha_r}\Big]\big(r,X_r\big)\,dW_r\,.
\end{aligned}$$

Now we take conditional expectation $\mathbb{E}^{t,x} := \mathbb{E}[\cdot|\mathcal{F}_t^X]$ on both sides of the last inequality, to get

$$\mathbb{E}^{t,x}\left[\int_t^\tau f^{\alpha_r}(r,X_r)\,dr + v(\tau,X_\tau)\right] \leq v(t,x) - \varepsilon\mathbb{E}^{t,x}\left[\tau^\alpha - t\right].$$

Now we first take infimum over all controls $\alpha \in \mathcal{A}[t,\tau]$ on the RHS and then supremum over all controls $\alpha \in \mathcal{A}[t,\tau]$ to get

$$\sup_{\alpha\in\mathcal{A}} \mathbb{E}^{t,x}\left[\int_t^\tau f^{\alpha_s}(s,X_s)ds + v(\tau,X_\tau)\right] \leq v(t,x) - \varepsilon \inf_\alpha \mathbb{E}^{t,x}\left[\tau^\alpha - t\right].$$

From Exercise 4.10 we have that $\inf_\alpha \mathbb{E}^{t,x}\left[\tau^\alpha - t\right] = m > 0$ and so

$$\sup_{\alpha\in\mathcal{A}} \mathbb{E}^{t,x}\left[\int_t^\tau f^{\alpha_s}(s,X_s)ds + v(\tau,X_\tau)\right] \leq v(t,x) - \varepsilon m < v(t,x).$$

But the Bellman principle states that:

$$v(t,x) = \sup_{\alpha\in\mathcal{A}} \mathbb{E}^{t,x}\left[\int_t^\tau f^{\alpha_s}(s,X_s)\,ds + v(\tau,X_\tau)\right].$$

Hence we've obtained a contradiction and completed the proof. $\square$

Exercise 4.8. Assume that X is a solution to $dX_s = b_s(X_s)\,dt + \sigma_s(X_s)\,dW_s$, $X_t = x$, $s \in [t,T]$ Assume that b and σ are bounded such that

$$\operatorname*{ess\,sup}_{\omega\in\Omega} \sup_{s\in[0,T]} \sup_{x\in\mathbb{R}^d} |b_s(\omega,x)| + |\sigma_s(\omega,x)| \leq K < \infty.$$

1. Show that for any $m \in \mathbb{N}$ there is $c > 0$ (depending on T, m and bound on b and σ) such that for all $x \in \mathbb{R}^d$ we have

$$\mathbb{E}|X_{s'}^{t,x} - X_s^{t,x}|^{2m} \leq c|s'-s|^m.$$

2. Use Kolmogorov's continuity, see Theorem A.25, to obtain Hölder continuity of sample paths of solutions.

Exercise 4.9. Let $X^{t,x}$ be the solution to the SDE $dX_s = b_s(X_s)\,dt + \sigma_s(X_s)\,dW_s$, $X_t = x$, $s \in [t,T]$ with b and σ like in Exercise 4.8.

Assume that $h : [0,T]\times\mathbb{R}^d \to \mathbb{R}$ is continuous. Assume that for any stopping time $\tau \geq t$

$$\int_t^\tau h(s,X_s^{t,x})\,ds \leq 0 \quad \forall (t,x) \in [0,T]\times\mathbb{R}^d.$$

Show that $h(t,x) \leq 0$ for all $(t,x) \in [0,T]\times\mathbb{R}^d$.

Exercise 4.10. Let $X^{t,x}$ be the solution to the controlled SDE $dX_s = b_s(X_s,\alpha_s)\,dt + \sigma_s(X_s,\alpha_s)\,dW_s$, $X_t = x$, $s \in [t,T]$ with α adapted to the filtration generated by W. Assume

$$\sup_{a\in A} \sup_{s\in[0,T]} \sup_{x\in\mathbb{R}^d} |b_s(x,a)| + |\sigma_s(x,a)| \leq K < \infty.$$

Let

$$\tau := \tau^{t,x,\alpha} := \inf\{s \geq t : |X_s^{t,x,\alpha} - x| \geq \delta\}.$$

Show that $\inf_\alpha \mathbb{E}[\tau^{t,x,\alpha} - t] > 0$.

Theorem 4.11 (HJB verification)**.** *If, on the other hand, some u in $C^{1,2}([0,T)\times\mathbb{R}^d)$ satisfies* (4.4) *and*

i) if we have that for all $(t,x) \in [0,T] \times \mathbb{R}^d$ there is some measurable function $a : [0,T] \times \mathbb{R}^d \to A$ such that

$$a(t,x) \in \arg\max_{a' \in A} \Big((L^{a'} u)(t,x) + f^{a'}(t,x) \Big) \,, \tag{4.5}$$

ii) and if

$$dX_s^* = b\big(s, X_s^*, a(s, X_s^*)\big)\, ds + \sigma\big(s, X_s^*, a(s, X_s^*)\big)\, dW_s, \quad X_t^* = x$$

admits a unique solution,

iii) and if the process

$$t' \mapsto \int_t^{t'} \partial_x u\big(s, X_s^*\big) \sigma\big(s, X_s^*, a(s, X_s^*)\big)\, dW_s \tag{4.6}$$

is a martingale in $t' \in [t,T]$,

then

$$\alpha_s^* := a\big(s, X_s^*\big) \qquad s \in [t,T]$$

is optimal for problem (P) and $v(t,x) = u(t,x)$.

Proof. Let $\alpha_s^* = a(s, X_s^*)$. Apply Itô's formula to u and X^* to see that

$$\begin{aligned}
&\int_t^T f_s^{\alpha_s^*}\big(X_s^*\big)\, ds + g(X_T^*) - u(t,x) = \int_t^T f_s^{\alpha_s^*}\big(X_s^*\big)\, ds + u(T, X_T^*) - u(t,x)\\
&= \int_t^T \Big[\partial_t u(s, X_s^*) + L^{\alpha_s^*}(s, X_s^*) u(s, X_s^*) + f_s^{\alpha_s^*}\big(X_s^*\big) \Big]\, ds\\
&\quad + \int_t^T \partial_x u\big(s, X_s^*\big) \sigma\big(s, X_s^*, a(s, X_s^*)\big)\, dW_s\\
&= \int_t^T \partial_x u\big(s, X_s^*\big) \sigma\big(s, X_s^*, a(s, X_s^*)\big)\, dW_s \,,
\end{aligned}$$

since for all (t,x) it holds that

$$\sup_{a \in A} [L^a(t,x) u(t,x) + f^a(t,x)] = L^{a(t,x)}(t,x) u(t,x) + f^{a(t,x)}(t,x) \,.$$

Hence, as the stochastic integral is a martingale by assumption,

$$\mathbb{E}\Big[\int_t^T f_s^{\alpha_s^*}\big(X_s^*\big)\, ds + g(X_T^*) - u(t,x) \Big] = 0 \,.$$

So

$$u(t,x) = \mathbb{E}\Big[\int_t^T f_s^{\alpha_s^*}\big(X_s^*\big)\, ds + g(X_T^*) \Big] \le \sup_{\alpha \in \mathcal{A}} \mathbb{E}\Big[\int_t^T f_s^{\alpha_s}\big(X^{t,x,\alpha}\big)\, ds + g(X^{t,x,\alpha}) \Big] = v(t,x) \,. \tag{4.7}$$

The same calculation with an arbitrary $\alpha \in \mathcal{A}$ and Itô formula applied to u and $X^{t,x,\alpha}$ leads to

$$\mathbb{E}\Big[\int_t^T f_s^{\alpha_s}\big(X_s^{t,x,\alpha}\big)\, ds + g(X_T^{t,x,\alpha}) - u(t,x)\Big] \leq 0\,.$$

Hence for any $\varepsilon > 0$ we have

$$v(t,x) \leq \varepsilon + \mathbb{E}\left[\int_t^T f_s^{\alpha_s^\varepsilon}\big(X_s^{t,x,\alpha^\varepsilon}\big)\, ds + g(X_T^{t,x,\alpha^\varepsilon})\right] \leq u(t,x)\,.$$

Hence $v(t,x) \leq u(t,x)$ and with (4.7) we can conclude that $v = u$.

Let

$$M_s := \int_t^s f_r^{\alpha_r^*}\big(X_r^*\big)\, dr + u\big(s, X_s^*\big)\,.$$

We would first like to see that this is a martingale. To that end, let us apply Itô's formula to v and X^* to see that

$$\begin{aligned} dM_s &= f_s^{\alpha_s^*}\big(X_s^*\big)\, ds + dv\big(s, X_s^*\big) \\ &= \Big[\partial_t v(s,X_s^*) + L^{\alpha_s^*}(s,X_s^*)v(s,X_s^*) + f_s^{\alpha_s^*}\big(X_s^*\big)\Big]\, ds + \partial_x v\big(s,X_s^*\big)\sigma\big(s,X_s^*,a(s,X_s^*)\, dW_s \\ &= \partial_x v\big(s,X_s^*\big)\sigma\big(s,X_s^*,a(s,X_s^*)\big)\, dW_s \end{aligned}$$

since $v = u$ satisfies (4.4). By assumption this stochastic integral is a martingale and hence M is also a martingale. By Theorem 4.6 α^* must be an optimal control process. $\square$

Theorem 4.11 is referred as the *verification theorem*. This is key for solving the control problem: if we know the value function v, then the dynamic optimization problem turns into a of static optimization problems at each point (t,x). Recall that (4.5) is calculated pointwise over (t,x).

4.3 Solving control problems using the HJB equation and verification theorem

Theorem 4.7 provides an approach to find optimal solutions:

1. Solve the HJB equation (4.4) (this is typically done by taking a lucky guess and in fact is rarely possible with pen and paper).
2. Find the optimal Markovian control rule $a(t,x)$ calculating (4.5). If you can, use calculus and anyway you probably had to do this in the step above anyway.
3. Check the optimally controlled SDE X^* has unique solution.
4. Verify the martingale condition.

This approach may end up with *failures*. Step one is to solve a fully non-linear second order PDE, that may not have a solution or may have many solutions.

In step two, given u that solves (4.4), the problem is a static optimization problem. This is generally much easier to solve.

If we can reach step three, then this step heavily depends on functions b and σ, for which we usually check case by case.

Example 4.12 (Merton problem with power utility and no consumption)**.** This is the classic finance application. The problem can be considered with multiple risky assets but we focus on the situation from Section 1.2.

Recall that we have risk-free asset B_t, risky asset S_t and that our portfolio has wealth given by

$$dX_s = X_s(\nu_s(\mu - r) + r)\, ds + X_s \nu_s \sigma\, dW_s\,,\ \ s \in [t,T]\,,\ \ X_t = x > 0\,.$$

Here ν_s is the control and it describes the fraction of our wealth invested in the risky asset. This can be negative (we short the stock) and it can be more than one (we borrow money from the bank and invest more than we have in the stock).

We take $g(x) := x^\gamma$ with $\gamma \in (0,1)$ a constant. Our aim is to maximize $J^\nu(t,x) := \mathbb{E}^{t,x}\left[g(X_T^\nu)\right]$. Thus our value function is

$$v(t,x) = \sup_{\nu \in \mathcal{U}} J^\nu(t,x) = \sup_{\nu \in \mathcal{U}} \mathbb{E}^{t,x}\left[g(X_T^\nu)\right]\,.$$

This should satisfy the HJB equation (Bellman PDE)

$$\begin{aligned} \partial_t v + \sup_u \left[\frac{1}{2}\sigma^2 u^2 x^2 \partial_{xx} v + x[(\mu - r)u + r]\partial_x v\right] &= 0 \ \text{ on } [0,T) \times (0,\infty) \\ v(T,x) = g(x) &= x^\gamma \ \ \forall x > 0\,. \end{aligned}$$

At this point our best chance is to guess what form the solution may have. We try $v(t,x) = \lambda(t) x^\gamma$ with $\lambda = \lambda(t) > 0$ differentiable and $\lambda(T) = 1$. This way at least the terminal condition holds. If this is indeed a solution then (using it in HJB) we have

$$\lambda'(t) + \sup_u \left[\frac{1}{2}\sigma^2 u^2 \gamma(\gamma - 1) + (\mu - r)\gamma u + r\gamma\right]\lambda(t) = 0 \ \ \forall t \in [0,T)\,,\ \ \lambda(T) = 1\,.$$

since $x^\gamma > 0$ for $x > 0$ and thus we were allowed to divide by this. Moreover we can calculate the supremum by observing that it is quadratic in u with negative leading term $(\gamma - 1)\gamma < 0$. Thus it is maximized when $u^* = \frac{\mu - r}{\sigma^2(1-\gamma)}$. The maximum itself is

$$\beta(t) := \frac{1}{2}\sigma^2 (u^*)^2 \gamma(\gamma - 1) + (\mu - r)\gamma u^* + r\gamma\,.$$

Thus

$$\lambda'(t) = -\beta(t)\lambda(t)\,,\ \ \lambda(T) = 1 \quad \Longrightarrow \lambda(t) = \exp\left(\int_t^T \beta(s)\, ds\right)\,.$$

Thus we think that the value function and the optimal control are

$$v(t,x) = \exp\left(\int_t^T \beta(s)\, ds\right) x^\gamma \ \text{ and } \ u^* = \frac{\mu - r}{\sigma^2(1-\gamma)}\,.$$

This now needs to be verified using Theorem 4.11. We've already carried out the maximization and found that $u^*(t,x) = \frac{\mu - r}{\sigma^2(1-\gamma)}$ is a constant and hence it's clearly measurable. Hence i) in Theorem 4.11 is satisfied.

Next, we note that the SDE for X^* always has a solution since plugging in the u^* we have an SDE with linear coefficients (thus Lipschitz) and Proposition 3.2 provides existence of unique solution. Hence ii) in Theorem 4.11 is satisfied.

Finally, we would like to check that $t \mapsto \int_0^t \partial_x v(s, X_s^*) X_s^* u^* \sigma \, dW s$ is a martingale. We note that $\partial_x v(s, X_s^*) = \gamma\lambda(s)(X_s^*)^{\gamma-1}$. To check that we have martingale it suffices to show that $\mathbb{E}\int_0^T |\gamma\lambda(s)(X_s^*)^{\gamma-1} X_s^* u^* \sigma|^2 \, dt < \infty$. Moreover

$$\mathbb{E}\int_0^T |\gamma\lambda(s)(X_s^*)^{\gamma-1} X_s^* u^* \sigma|^2 \, dt \le \gamma^2 \sup_{t\in[0,T]} \lambda(t)^2 (u^*)^2 \sigma^2 \mathbb{E}\int_0^T |(X_s^*)^\gamma|^2 \, dt \,.$$

We see that we have to focus on the moments of $(X^*)^{2\gamma}$. At this point we could use Proposition 3.5 but instead we proceed manually. From Itô's formula

$$\begin{aligned} d(X_s^*)^{2\gamma} &= 2\gamma(X_s^*)^{2\gamma-1} dX_s^* + \frac{1}{2} 2\gamma(2\gamma-1)(X_s^*)^{\gamma-2} d(X_s^*) d(X_s^*) \\ &= (X_s^*)^{2\gamma} \left[2\gamma[u^*(\mu - r) + r] \, ds + \gamma(\gamma-1) u^* \sigma \, dW_s\right] . \end{aligned}$$

We can solve this (like the SDE for geometric brownian motion) and see the solution is log-normal so it will have 1st and 2nd moment bounded uniformly in $t \in [0, T]$. Hence

$$\mathbb{E}\int_0^T |(X_s^*)^\gamma|^2 \, dt < \infty \,.$$

Hence we verified iii) of Theorem 4.11. Thus v is indeed the value function and u^* is indeed the optimal control.

Example 4.13 (1d Linear-quadratic control problem)**.** This example is a classic engineering application. Note that it can be considered in multiple spatial dimensions but here we focus on the one-dimensional case for simplicity. The multi-dimensional version is solved using HJB in Exercise 4.15 and also later using Pontryagin optimality principle, in Example 6.11.

We consider

$$dX_s = [H(s)X_s + M(s)\alpha_s] \, ds + \sigma(s) dW_s \,, s \in [t, T] \,, X_t = x \,.$$

Our aim is to maximize

$$J^\alpha(t, x) := \mathbb{E}^{t,x}\left[\int_t^T (C(s)X_s^2 + D(s)\alpha_s^2) \, ds + R X_T^2\right] ,$$

where $C = C(t) \le 0$, $R \le 0$ and $D = D(t)$, are given and deterministic and bounded in t s.t. for some $\delta > 0$ we have $D(t) + \delta < 0$ for all t. The interpretation is the following: since we are losing money at rate C proportionally to X^2, our aim is to make X^2 as small as possible as fast as we can. However controlling X costs us at a rate D proportionally to the strength of control we apply.

The value function is $v(t, x) := \sup_\alpha J^\alpha(t, x)$.

Let us write down the Bellman PDE (HJB equation) we would expect the value function to satisfy:

$$\begin{aligned} \partial_t v + \sup_a \left[\frac{1}{2}\sigma^2 \partial_x^2 v + [H\, x + M\, a]\partial_x v + C\, x^2 + D\, a^2\right] = 0 \text{ on } [0, T) \times \mathbb{R} \,, \\ v(T, x) = R x^2 \;\; \forall x \in \mathbb{R} \,. \end{aligned}$$

Since the terminal condition is $g(x) = Rx^2$ let us try $v(t, x) = S(t)x^2 + b(t)$ for some differentiable S and b. We re-write the HJB equation in terms of S and b: (omitting time dependence in H, M, σ, C and D), for $(t,x) \in [0,T) \times \mathbb{R}$,

$$S'(t)x^2 + b'(t) + \sigma^2 S(t) + 2H\, S(t)\, x^2 + C\, x^2 + \sup_a \left[2M\, a\, S(t)\, x + D\, a^2\right] = 0\,,$$
$$S(T) = R \text{ and } b(T) = 0\,.$$

For fixed t and x we can calculate $\sup_a[2M(t)aS(t)x + D(t)a^2]$ and hence write down the optimal control function $a^* = a^*(t,x)$. Indeed since $D < 0$ and since the expression is quadratic in a we know that the maximum is reached with $a^*(t,x) = -(D^{-1}M\,S)(t)x$.

We substitute a^* back in to obtain ODEs for $S = S(t)$ and $b = b(t)$ from the HJB equation. Hence

$$\left[S'(t) + 2H\, S(t) + C - D^{-1}M^2 S^2(t)\right] x^2 + b'(t) + \sigma^2 S(t) = 0\,,$$
$$S(T) = R \text{ and } b(T) = 0\,.$$

We collect terms in x^2 and terms independent of x and conclude that this can hold only if

$$S'(t) = D^{-1}M^2 S^2(t) - 2H\, S(t) - C\,, \;\; S(T) = R$$

and

$$b'(t) = -\sigma^2 S(t)\,, \;\; b(T) = 0\,.$$

The ODE for S is the *Riccati equation* which has unique solution for $S(T) = R$. We can obtain the expression for $b = b(t)$ by simply integrating:

$$b(T) - b(t) = -\int_t^T \sigma^2(r) S(r)\, dr\,.$$

Then

$$\alpha^*(t,x) = -(D^{-1}MS)(t)x \text{ and } v(t,x) = S(t)x^2 + \int_t^T \sigma^2(r)S(r)\, dr\,. \tag{4.8}$$

We'll do verification in the general multi-dimensional case in Example 4.15.

Exercise 4.14. This is a bit of linear algebra to make reading Example 4.15 easier.

1. Let $M \in \mathbb{R}^{d\times d}$ and $x \in \mathbb{R}^d$. Then $x^\top M\, x = x^\top M^\top x$.
2. Let $M \in \mathbb{R}^{d\times d}$. Show that $\partial_x[x^\top Mx] = M^\top x + Mx$ and if M is symmetric then $\partial_x[x^\top Mx] = 2Mx$.
3. Let $L, M \in \mathbb{R}^{d\times d}$ and $x \in \mathbb{R}^d$. Show that $x^\top L^\top M^\top x = x^\top M\, L\, x$.
4. Let $L \in \mathbb{R}^{d\times d}$, $M \in \mathbb{S}^d$ and show that $L^\top M + ML$ is symmetric.

Example 4.15 (Linear-quadratic control problem)**.** This is the multi-dimensional version and the only additional difficulty is from linear algebra. See also Example 6.11.

We consider an $\mathbb{R}^{d'}$-valued Wiener process W, $\mathbb{R}^m$-valued controls α_s, $H = H(t)$ taking values in $\mathbb{R}^{d\times d}$ and $M = M(t)$ taking values in $\mathbb{R}^{d\times m}$, $\sigma = \sigma(t)$ taking values in $\mathbb{R}^{d\times d'}$ with $\sup_{t\in[0,T]} |\sigma(t)| < \infty$, $x \in \mathbb{R}^d$ and a controlled SDE

$$dX_s = [H(s)X_s + M(s)\alpha_s]\, ds + \sigma(s) dW_s\,, s \in [t,T]\,, X_t = x\,.$$

Our aim is to *minimize*, for any $(t,x) \in [0,T] \times \mathbb{R}^d$ the objective

$$J^\alpha(t,x) := \mathbb{E}^{t,x}\left[\int_t^T \left[X_s^\top C(s) X_s + \alpha_s^\top D(s)\alpha_s\right) ds + X_T^\top R X_T\right],$$

over square integrable, adapted controls processes α, where $C = C(t)$, $D = D(t)$, R, are given, symmetric, measurable, deterministic and bounded in t. For any $x \in \mathbb{R}^d$ and $t \in [0,T]$ we have $x^\top C x \geq 0$, $x^\top R x \geq 0$. Moreover, there is $\delta > 0$ such that for all $a \in \mathbb{R}^m$ and $t \in [0,T]$ we have $a^\top D(t) a \geq \delta$.

The value function is $v(t,x) := \inf_\alpha J^\alpha(t,x)$ since we're minimizing.

Let us write down the Bellman PDE (HJB equation) we would expect the value function to satisfy:

$$\partial_t v + \inf_{a\in\mathbb{R}^m}\left[\frac{1}{2}\mathrm{tr}\left[\sigma\sigma^\top \partial_x^2 v\right] + [x^\top H^\top + a^\top M^\top]\partial_x v + x^\top C\, x + a^\top D\, a\right] = 0 \ \text{ on } [0,T)\times\mathbb{R}^d,$$
$$v(T,x) = x^\top R x \ \ \forall x \in \mathbb{R}^d.$$

Since the terminal condition is $g(x) = x^\top R x$ let us try $v(t,x) = x^\top S(t) x + b(t)$ for some $S \in C^1([0,T];\mathbb{S}^d_{\geq 0})$ and $b \in C^1([0,T];\mathbb{R})$.[8] We are asking for $S(t)$ to be symmetric and positive definite for all t which makes sense as that's the form of the terminal condition. Note that this means that $\partial_x v = 2S(t)x$. We re-write the HJB equation in terms of S and b: (omitting time dependence in S, b, H, M, σ, C and D), for $(t,x) \in [0,T) \times \mathbb{R}$,

$$x^\top S' x + b' + \mathrm{tr}[\sigma\sigma^\top S] + x^\top 2H^\top S\, x + x^\top C\, x + \inf_{a\in\mathbb{R}^m}\left[2a^\top M^\top S\, x + a^\top D\, a\right] = 0,$$
$$S(T) = R \ \text{ and } \ b(T) = 0.$$

For fixed t and x we can calculate $\inf_a [a^\top M^\top (S^\top + S)x + a^\top D a]$ using the first order condition

$$0 = 2M^\top S x + 2Da$$

and hence write down the optimal control function $a^* = a^*(t,x)$. Indeed, since D is invertible and since the expression is convex in a we know that the minimum is reached with $a^*(t,x) = -(D^{-1}M^\top S)(t)x$.

Note that $(a^*)^\top = -x^\top S M D^{-1}$ and so

$$2(a^*)^\top M^\top S\, x = -2x^\top S\, M D^{-1} M^\top S\, x$$

and

$$(a^*)^\top D\, a = (a^*)^\top = x^\top S\, M D^{-1} M^\top\, S\, x.$$

Thus

$$\inf_a [a^\top M^\top (S^\top + S)x + a^\top D a] = -x^\top S\, M D^{-1} M^\top S\, x.$$

Substituting back into the HJB and using $2H^\top S = H^\top(S^\top + S)$ and $x^\top H^\top S x = x^\top H^\top S^\top x = x^\top S\, H\, x$ we get

$$x^\top S' x + b' + \mathrm{tr}[\sigma\sigma^\top S] + x^\top (H^\top S + S\, H)\, x + x^\top C\, x - x^\top S\, M D^{-1} M^\top S\, x = 0.$$

[8] $\mathbb{S}^d_{\geq 0} := \{M \in \mathbb{R}^{d\times d} : \forall i,j = 1,\dots,d$ we have $M_{ij} = M_{ji}$ and $\forall x \in \mathbb{R}^d$ we have $x^\top M x \geq 0\}$.

Collecting terms we thus get ODEs for $S = S(t)$ and $b = b(t)$ from the HJB equation as follows

$$S'(t) = S(t)MD^{-1}M^\top S(t) - H^\top S(t) - S(t)\,H - C\,,\;\; S(T) = R$$

and

$$b'(t) = -\mathrm{tr}[\sigma\sigma^\top S(t)]\,,\;\; b(T) = 0\,.$$

The ODE for S is the *Riccati equation* which has unique symmetric positive definite solution for $S(T) = R$ under our assumptions.[9] We can obtain the expression for $b = b(t)$ by simply integrating:

$$b(T) - b(t) = -\int_t^T \mathrm{tr}[\sigma\sigma^\top(r)S(r)]\,dr\,.$$

Thus, in conclusion,

$$\alpha^*(t,x) = -(D^{-1}M^\top S)(t)x \;\text{ and }\; v(t,x) = x^\top S(t)x + \int_t^T \mathrm{tr}[\sigma\sigma^\top S](r)\,dr\,. \tag{4.9}$$

It is time to perform verification. We see that the control function is linear in x, continuous in t and hence jointly continuous and thus measurable. We will now check conditions of Theorem 4.11. The SDE with the optimal control is

$$dX_s^* = \rho(s)X_s^*\,ds + \sigma(s)dW_s\,,\;\; s\in[t,T]\,,\;\; X_t^* = x\,,$$

where $\rho := H - D^{-1}\,M^\top\,S\,M$. This is deterministic and bounded in time. The SDE thus satisfies the Lipschitz conditions and it has a unique strong solution for any t, x. Since $\partial_x v(r, X_r^*) = 2S(r)X_s^*$, since $\sup_{r\in[t,T]} S^2(r)$ is bounded (continuous function on a closed interval), $\sup_{r\in[t,T]} |\sigma(r)|^2 < \infty$ by assumption and since $\sup_{r\in[t,T]} \mathbb{E}[|X_r^*|^2] < \infty$ (moment estimate for SDEs with Lipschitz coefficients) we get

$$\mathbb{E}\int_t^T |S(r)|^2|X_s^*|^2|\sigma(r)|^2\,dr < \infty$$

and thus conclude that $s \mapsto \int_t^s (S(r)X_r^*)^\top\sigma(r)\,dW_r$ is a martingale. Thus Theorem 4.11 tells us that the value function and control given by (4.9) are indeed optimal.

4.4 Policy Improvement Algorithm

As in the controlled Markov chain case (see Section 2.3) one can solve the control problem using the policy improvement algorithm, stated below.

[9] We can see the symmetry of the solution by considering a Picard iteration; we know that at least locally the solution is a limit of Picard iteration procedure. We have $S(T) = R$ symmetric. Then if the previous iterate is symmetric then the ODE right-hand-side is also symmetric which means that the new value obtained by Picard iteration is also symmetric.

An important point is that $H^\top S + SH$ is symmetric when S is symmetric (but H doesn't need to be). If we didn't do the substitution $2x^\top H^\top Sx = x^\top(H^\top S + SH)x$ earlier we would have been left with $2H^\top S$ in the equation and $H^\top S$ is *only* symmetric if H is itself symmetric. In that case the solution of the quadratic equation wouldn't have been symmetric.

Algorithm 1 Policy improvement algorithm:

Initialisation: make a guess of the control $a^0 = a^0(t,x)$.
while difference between v^{n+1} and v^n is large **do**
Given a control $a^n = a^n(t,x)$ solve the *linear* PDE

$$\begin{aligned} \partial_t v^n + \frac{1}{2}\mathrm{tr}(\sigma\sigma^\top D_x^2 v^n) + b^{a^n} D_x v^n + f^{a^n} &= 0 \text{ on } [0,T) \times \mathbb{R}^d\,, \\ v^n(T,\cdot) &= g \quad \text{on } x \in \mathbb{R}^d\,. \end{aligned} \tag{4.10}$$

Update the control

$$a^{n+1}(t,x) \in \arg\max_{a\in A} \left[\left(b^a D_x v^n + f^a \right)(t,x) \right]\,. \tag{4.11}$$

end while
return v^n, a^{n+1}.

Of course to solve the linear PDE one would need to employ a numerical methods (e.g. finite differences). Convergence and rate of convergence of the policy improvement algorithm can be found e.g. in [11] and in references there.

4.5 Exercises

Exercise 4.16 (Optimal liquidation with no permanent market impact)**.** Solve the optimal liquidation problem of Section 1.3 in the case $\lambda = 0$ (i.e. there is no permanent price impact of our trading on the market price).

Exercise 4.17 (Unattainable optimizer)**.** Here is a simple example in which no optimal control exists, in a finite horizon setting, $T \in (0,\infty)$. Suppose that the state equation is

$$dX_s = \alpha_s\,ds + dW_s \quad s \in [t,T]\,, \quad X_t = x \in \mathbb{R}.$$

A control α is admissible ($\alpha \in \mathcal{A}$) if: α takes values in $\mathbb{R}$, is $(\mathcal{F}_t)_{t\in[0,T]}$-adapted, and $\mathbb{E}\int_0^T \alpha_s^2\,ds < \infty\,.$

Let $J(t,x,\alpha) := \mathbb{E}[|X_T^{t,x,\alpha}|^2]$. The value function is $v(t,x) := \inf_{\alpha\in\mathcal{A}} J(t,x,\alpha)$. Clearly $v(t,x) \geq 0$.

i) Show that for any $t \in [0,T]$, $x \in \mathbb{R}$, $\alpha \in \mathcal{A}$ we have $\mathbb{E}[|X_T^{t,x,\alpha}|^2] < \infty$.

ii) Show that if $\alpha_t := -cX_t$ for some constant $c \in (0,\infty)$ then $\alpha \in \mathcal{A}$ and

$$J^\alpha(t,x) = J^{cX}(t,x) = \frac{1}{2c} - \frac{1-2cx^2}{2c} e^{-2c(T-t)}.$$

Hint: with such an α, the process X is an Ornstein-Uhlenbeck process, see Exercise 1.16.

iii) Conclude that $v(t,x) = 0$ for all $t \in [0,T)$, $x \in \mathbb{R}$.

iv) Show that there is no $\alpha \in \mathcal{A}$ such that $J(t,x,\alpha) = 0$. *Hint:* Suppose that there is such a α and show that this leads to a contradiction.

v) The associated HJB equation is

$$\partial_t v + \inf_{a\in\mathbb{R}} \left\{ \frac{1}{2}\partial_{xx} v + a\partial_x v \right\} = 0, \quad \text{on } [0,T)\times\mathbb{R}.$$
$$v(T,x) = x^2\,.$$

Show that there is no value $\alpha\in\mathbb{R}$ for which the infimum is attained.

Conclusions from Exercise 4.17: The value function $v(t,x) = \inf_{\alpha\in\mathcal{A}} J(t,x,\alpha)$ satisfies $v(t,x) = 0$ for all $(t,x)\in[0,T]\times\mathbb{R}$ but there is no admissible control α which attains the v (i.e. there is no $\alpha^*\in\mathcal{A}$ such that $v(t,x) = J(t,x,\alpha^*)$).

The goal in this problem is to bring the state process as close as possible to zero at the terminal time T. However, as defined above, there is no cost of actually controlling the system. We can set α arbitrarily large without any negative consequences. From a modelling standpoint, there is often a trade-off between costs incurred in applying control and our overall objective. Compare this with Example 4.15.

Exercise 4.18 (Merton problem with exponential utility and no consumption)**.** We return to the portfolio optimization problem, see Section 1.2. Unlike in Example 4.12 we consider the utility function $g(x) := -e^{-\gamma x}$, $\gamma > 0$ a constant. We will also take $r = 0$ for simplicity and assume there is no consumption ($C = 0$). With X_t denoting the wealth at time time t we have the value function given by

$$v(t,x) = \sup_{\pi\in\mathcal{U}} \mathbb{E}\left[g\left(X_T^{\pi,t,x,} \right) \right]\,.$$

i) Write down the expression for the wealth process in terms of π, the amount of wealth invested in the risky asset and with $r = 0$, $C = 0$.

ii) Write down the HJB equation associated to the optimal control problem. Solve the HJB equation by inspecting the terminal condition and thus suggesting a possible form for the solution. Write down the optimal control explicitly.

iii) Use verification theorem to show that the solution and control obtained in previous step are indeed the value function and optimal control.

Exercise 4.19 ([17]*p252, Prob. 4.8)**.** Solve the problem

$$\max_{\nu} \mathbb{E}\big[-\int_0^T \nu^2(t)\frac{e^{-X(t)}}{2}\,dt + e^{X(T)} \big],$$

where ν takes values in $\mathbb{R}$, subject to $dX(t) = \nu(t)e^{-X(t)}\,dt + \sigma\,dW(t)$, $X(0) = x_0\in\mathbb{R}$, $\sigma\in(0,\infty)$, σ, x_0 are fixed numbers.

Hint: Try a solution of the HJB equation of the form $v(t,x) = \phi(t)e^x + \psi(t)$.

For more exercises, see [17, Exercise 4.13, 4.14, 4.15].

4.6 Solutions to Exercises

Solution (to Exercise 4.8)**.** First part: Clearly

$$|X_{s'}^{t,x} - X_s^{t,x}|^{2m} \le 2^{m+1}\left| \int_s^{s'} b(X_u)\,du \right|^{2m} + 2^{m+1}\left| \int_s^{s'} \sigma(X_u)\,dW_u \right|^{2m}.$$

From Hölder's inequality we have

$$\left|\int_s^{s'} b(X_u)\,du\right|^{2m} \le (s'-s)^m \left(\int_t^T |b(X_u)|^2\,du\right)^m \le (s'-s)^m \left(\int_t^T K^2\,du\right)^m \le T^m K^{2m}(s'-s)^m .$$

From the Burkholder–Davis–Gundy inequality, Hölder's inequality and then with the growth assumption and moment bound

$$\begin{aligned}\mathbb{E}\left[\left|\int_s^{s'}\sigma(X_u)\,dW_u\right|^{2m}\right] &\le \mathbb{E}\left[\sup_{s\le r\le s'}\left|\int_s^{r}\sigma(X_u)\,dW_u\right|^{2m}\right] \le c_m\mathbb{E}\left[\left(\int_s^{s'}|\sigma(X_u)|^2\,du\right)^m\right]\\
&\le c_m\mathbb{E}\left[\left(\int_s^{s'} 1^{\frac{m}{m-1}}\,du\right)^{m-1}\left(\int_s^{s'}|\sigma(X_u)|^{2m}\,du\right)\right]\\
&\le c_m(s'-s)^{m-1}\left(\int_s^{s'}K^{2m}\,du\right) \le c_m K^{2m}|s'-s|^m .\end{aligned}$$

Altogether

$$\mathbb{E}|X^{t,x}_{s'} - X^{t,x}_s|^{2m} \le K^{2m}(T^m + c_m)|s'-s|^m ,$$

where c_m is the constant from the Burkholder–Davis–Gundy inequality.

Second part: Using the meaning of α,β from Theorem A.25 we have $\alpha = m-1$ and $\beta = 2m$ so there is a version of X which satisfies that there is a r.v. $C(\omega) > 0$ such that for a.a. ω, any x and any $t \le s \le s' \le T$ that

$$|X^{t,x}_{s'}(\omega) - X^{t,x}_s(\omega)| \le C(\omega)|s'-s|^\delta$$

with $\delta \in (0, \frac{m-1}{2m})$. That is we get $\frac{1}{2} - \varepsilon$ Hölder continuity for any $\varepsilon > 0$.

Solution (to Exercise 4.9)**.** Assume, that there is $(t,x) \in [0,T]\times\mathbb{R}^d$ such that $h(t,x) > 0$. Then there is $\varepsilon > 0$ such that $h(t,x) > 2\varepsilon > 0$. By continuity of h there is $\delta > 0$ so that

$$|h(s,y) - h(t,x)| < \varepsilon \;\; \forall (s,y) \in [t,t+\delta)\times B_\delta(x) .$$

Let

$$\tau := \inf\{s \ge t : |X^{t,x}_s - x| \ge \delta\} \;\text{ and }\; \hat\tau := (t+\delta)\wedge\tau .$$

Then for all $s \in [t,\hat\tau)$ we have $h(s, X^{t,x}_s) > \varepsilon$. Hence

$$\int_t^{\hat\tau} h(s,X_s)\,ds > \int_t^{\hat\tau}\varepsilon\,ds = \varepsilon(\hat\tau - t) .$$

Note that from Exercise 4.8 we have that $\delta = |x - X_\tau| \le C|\tau - t|^\gamma$ and so $\left(\frac{\delta}{C}\right)^{1/\gamma} \le \tau - t$. Hence $\hat\tau - t \ge \delta \wedge \left(\frac{\delta}{C}\right)^{1/\gamma} \ge \left(\frac{\delta}{C}\right)^{1/\gamma}$ if we assume $\delta < 1$ and $C > 1$.

$$\int_t^\tau h(s, X^{t,x}_s)\,ds > \varepsilon(\hat\tau - t) > 0,$$

is a contradiction. This implies that $h(t,x) \le 0$.

Solution (to Exercise 4.10)**.** From Exercise 4.8 we have that $\delta = |x - X^{t,x,\alpha}_\tau| \le C(\tau^{t,x,\alpha} - t)^\gamma$. Here $C = C(\omega)$ but does not depend on α, only on the bound K. Hence

$$\inf_\alpha \mathbb{E}[\tau^{t,x,\alpha} - t] \ge \mathbb{E}\left[\left(\tfrac{\delta}{C}\right)^{1/\gamma}\right] > 0,$$

since an integral over a random variable which is strictly positive must be strictly positive.

Solution (to Exercise 4.14)**.**

1. $$x^\top M\,x = \sum_{i,j} x_i M_{ij} x_j = \sum_{i,j} x_i (M^\top)_{ji} x_j = \sum_{i,j} x_j (M^\top)_{ji} x_i = x^\top M^\top x .$$

2. Let $I \in R^{d\times d}$ be the identity matrix. Then

$$\begin{aligned}\partial_{x_i}[x^\top M x] &= \partial_{x_i}\Big[\sum_{j,k} x_j M_{jk} x_k\Big] = \sum_{j,k} x_j M_{jk} I_{ik} + \sum_{j,k} I_{ji} M_{jk} x_k = \sum_j x_j M_{ji} + \sum_k M_{ik} x_k\\
&= \sum_j (M^\top)_{ij} x_j + \sum_k M_{ik} x_k = (M^\top x + Mx)_i .\end{aligned}$$

   If $M = M^\top$ then we get $\partial_{x_i}[x^\top M x] = 2(Mx)_i$.

3. From part 1. we have

$$x^\top L^\top M^\top x = x^\top (L^\top M^\top)^\top x = x^\top M L\, x\,.$$

4. We have $M = M^\top$. Then

$$(L^\top M + ML)^\top = (L^\top M + ML)^\top = M^\top L + L^\top M = L^\top M + ML\,.$$

Solution (to Exercise 4.16)**.** Recall that we have: the inventory process:

$$dQ_u = -\alpha_u\, du \text{ with } Q_t = q > 0 \text{ initial inventory}$$

and the asset price, an $\mathbb{R}$-valued process:

$$dS_u = -\lambda\, \alpha_u\, du + \sigma\, dW_u\,, \;\; S_t = S\,.$$

Moreover the temporary price impact means that the execution price is

$$\hat{S}_t = S_t - \kappa\alpha_t\,.$$

Here $\lambda \geq 0, \sigma \geq 0$ and $\kappa \geq 0$ are constant. For given $T > 0$, $\theta > 0$ we wish to maximize (over adapted, square integrable trading strategies α), the expected amount gained in sales, whilst penalising the terminal inventory (with $\theta > 0$):

$$J(t,q,S,\alpha) := \mathbb{E}\left[\int_t^T \left[S_u^{t,S,\alpha}\, \alpha_u - \kappa\alpha_u^2\right] du + Q_T^{t,q,\alpha}\, S_T^{t,S,\alpha} - \theta\, |Q_T^{t,q,\alpha}|^2\right].$$

The goal is to find

$$V(t,q,S) := \sup_\alpha J(t,q,S,\alpha)\,.$$

From Theorem 4.7 we can write down the HJB equation for $V = V(t,S,q)$. To do that we have to figure out what the "drift" and "diffusion" coefficients are for this problem. We effectively have $x = (q,S)$, in the notation of Theorem 4.7 we have $d = 2$, $d' = 1$ and the drift is

$$(a,t,q,S) = (a,t,x) \mapsto b^a(t,x) = \begin{pmatrix} -a \\ -\lambda a \end{pmatrix}$$

while the diffusion matrix[10] is

$$(a,t,q,S) = (a,t,x) \mapsto \sigma^a(t,x) = \begin{pmatrix} 0 \\ \sigma \end{pmatrix}.$$

Note that

$$\sigma^a (\sigma^a)^\top = \begin{pmatrix} 0 & 0 \\ 0 & \sigma^2 \end{pmatrix}$$

and so

$$L^a V(t,q,S) = \tfrac{1}{2}\sigma^2 \partial_{SS} V(t,q,S) - a\partial_q V(t,q,S) - \lambda a \partial_S V(t,q,S)\,.$$

Finally we note that the "running gain" is

$$f^a(t,x) = f^a(t,q,S) = (S - \kappa a)a\,.$$

Hence, in our case when $\lambda = 0$, the HJB equation is

$$\partial V_t + \frac{1}{2}\sigma^2 \partial_{SS} V + \sup_{a\in A}\{(S-\kappa a)a - a\partial_q V\} = 0 \text{ on } [0,T)\times\mathbb{R}\times\mathbb{R}\,, \tag{4.12}$$

with the terminal condition

$$V(T,q,S) = qS - \theta q^2 \quad \forall (q,S) \in \mathbb{R}\times\mathbb{R}\,. \tag{4.13}$$

Next we note that

$$a \mapsto (S - \partial_q V)a - \kappa a^2 \text{ attains its maximum with } a^* = \frac{S - \partial_q V}{2\kappa}\,.$$

Hence the HJB equation (4.12) becomes

$$\partial V_t + \frac{1}{2}\sigma^2 \partial_{SS} V + \frac{1}{4\kappa}\left(S - \partial_q V\right)^2 = 0 \text{ on } [0,T)\times\mathbb{R}\times\mathbb{R}\,. \tag{4.14}$$

[10] One has to be careful to distinguish the constant σ and the diffusion coefficient which is a matrix-valued function σ^a

We now have to "guess" an ansatz for V and, observing the similarities here with the linear-quadratic case of Example 4.15, we try
$$V(t,q,S) = \beta(t)qS + \gamma(t)q^2 .$$
With $\beta(T) = 1$ and $\gamma(T) = -\theta$ we have the terminal condition (4.13) satisfied. To proceed we calculate the partial derivatives of V and substitute those into the HJB (4.14) to obtain

$$\beta'(t)qS + \gamma'(t)q^2 + \frac{1}{4\kappa}\big[S - \beta(t)S + 2\gamma(t)q\big]^2 = 0 \quad \forall (t,q,S) \in [0,T) \times \mathbb{R} \times \mathbb{R} . \tag{4.15}$$

This is equivalently

$$\begin{aligned}
&\beta'(t)qS + \gamma'(t)q^2 \\
&\qquad + \frac{1}{4\kappa}\big[S^2 - 2\beta(t)S^2 - 4\gamma(t)qS + \beta(t)^2S^2 + 4\beta(t)\gamma(t)qS + 4\gamma(t)^2q^2\big] \\
=&0 \quad \forall (t,q,S) \in [0,T) \times \mathbb{R} \times \mathbb{R} .
\end{aligned}$$

This has to hold for all S^2, q^2 and qS. Starting with S^2 terms we get that

$$1 - 2\beta(t) + \beta(t)^2 = 0 \quad \forall t \in [0,T)$$

which can only be true if $\beta(t) = 1$ (since $\beta(T)$ must be 1 and we need β differentiable). Considering now the qS term we have ($\beta'(t) = 0$ since we now have $\beta(t) = 1$):

$$-4\gamma(t) + 4\gamma(t) = 0 \quad \forall t \in [0,T)$$

which holds regardless of choice of γ. Finally we have the q^2 terms which lead to

$$\gamma'(t) + \frac{1}{\kappa}\gamma(t)^2 = 0 \quad \forall t \in [0,T) .$$

We recall the terminal condition $\gamma(T) = -\theta$ and solve this ODE[11] thus obtaining

$$\gamma(t) = -\left(\frac{1}{\theta} + \frac{1}{\kappa}(T-t)\right)^{-1} .$$

This fully determines the value function

$$V(t,q,S) = qS + \gamma(t)q^2$$

and the optimal control

$$a^*(t,q,S) = -\frac{1}{\kappa}\gamma(t)q .$$

We note that the optimal control is independent of S and in fact the entire control problem does not depend on the volatility parameter σ.

Solution (to Exercise 4.17)**.**

i) We use the fact that $\mathbb{E}\int_0^T \alpha_r^2\,dr < \infty$ for admissible control. We also use that $(a+b)^2 \le 2a^2 + 2b^2$. Then for, any $s \in [t,T]$,

$$\mathbb{E}[X_s^2] \le 4x^2 + 4\mathbb{E}\left(\int_t^s \alpha_r\,dr\right)^2 + 2\mathbb{E}(W_s - W_t)^2 .$$

With Hölder's inequality we get

$$\mathbb{E}[X_s^2] \le 4x^2 + 4(s-t)\mathbb{E}\int_t^s \alpha_r^2\,dr + 2(s-t) \le c_T\left(1 + x^2 + \mathbb{E}\int_0^T \alpha_r^2\,dr\right) < \infty . \tag{4.16}$$

ii) Substitute $\alpha_s = -cX_s$. The Ornstein-Uhlenbeck SDE, see Exercise 1.16, has solution

$$X_T = e^{-c(T-t)}x + \int_t^T e^{-c(T-r)}\,dW_r .$$

We square this, take expectation (noting that the integrand in the stochastic integral is deterministic and square integrable):

$$\mathbb{E}X_T^2 = e^{-2c(T-t)}x^2 + \mathbb{E}\left(\int_t^T e^{-c(T-r)}\,dW_r\right)^2 .$$

[11] You can for instance recall that if $f(t) = -\frac{1}{t}$ then $f'(t) = \frac{1}{t^2}$ and so $f'(t) = f(t)^2$. Manipulating expressions of this type can lead you to the correct solution.

With Itô's isometry we get

$$\mathbb{E}X_T^2 = e^{-2c(T-t)}x^2 + \int_t^T e^{-2c(T-r)}\,dr\,.$$

Now we just need to integrate to obtain $J^\alpha(t,x) = J^{cX}(t,x) = \mathbb{E}X_T^2$.

iii) We know that $v(t,x) \geq 0$ already. Moreover

$$v(t,x) = \inf_{\alpha\in\mathcal{U}} J^\alpha(t,x) \leq \lim_{c\nearrow\infty} J^{cX}(t,x) = \lim_{c\nearrow\infty}\left[\frac{1}{2c} - \frac{1-2cx^2}{2c}e^{-2c(T-t)}\right] = 0\,.$$

iv) Assume that an optimal $\alpha^* \in \mathcal{A}$ exists so that $\mathbb{E}[X_T^{\alpha^*,t,x}] = J^{\alpha^*}(t,x) = 0$ for any $t < T$ and any x. We will show this leads to contradiction.
First of all, we can calculate using Itô formula that

$$d|X_s^*|^2 = 2X_s^*\alpha_s^*\,ds + 2X_s^*\,dW_s + ds\,.$$

Hence

$$0 = \mathbb{E}[(X_T^*)^2] = x^2 + \mathbb{E}\int_t^T (2X_s^*\alpha_s^* + 1)\,ds + 2\mathbb{E}\int_t^T X_s^*\,dW_s\,.$$

But since α^* is admissible we have $\int_t^T \mathbb{E}(X_s^*)^2\,ds < \infty$ due to (4.16). This means that the stochastic integral is a martingale and hence its expectation is zero. We now use Fatou's lemma and take the limit as $t \nearrow T$. Then

$$-x^2 = 2\liminf_{t\nearrow T}\mathbb{E}\int_t^T (X_s^*\alpha_s^* + 1)\,ds \geq 2\mathbb{E}\left[\liminf_{t\nearrow T}\int_t^T (X_s^*\alpha_s^* + 1)\,ds\right] = 0\,.$$

So $-x^2 \geq 0$. This cannot hold for all $x \in \mathbb{R}$ and so we have contradiction.

v) If $\partial_x v(t,x) \neq 0$, then $a = \pm\infty$. If $\partial_x V(t,x) = 0$, then a is undefined. One way or another there is no real number attaining the infimum.

Solution (to Exercise 4.18)**.** The wealth process (with the control expressed as π, the amount of wealth invested in the risky asset and with $r = 0$, $C = 0$), is given by

$$dX_s = \pi_s\mu\,ds + \pi_s\sigma\,dW_s\,,\quad s \in [t,T]\,,\quad X_t = x > 0\,. \tag{4.17}$$

The associated HJB equation is

$$\partial_t v + \sup_{p\in\mathbb{R}}\left[\frac{1}{2}p^2\sigma^2\partial_{xx}v + p\,\mu\,\partial_x v\right] = 0 \text{ on } [0,T)\times\mathbb{R},$$
$$v(T,x) = g(x)\ \ \forall x \in \mathbb{R}\,.$$

We make a guess that $v(t,x) = \lambda(t)g(x) = -\lambda(t)e^{-\gamma x}$ for some differentiable function $\lambda = \lambda(t) \geq 0$. Since we can factor out the non-negative $\lambda(t)e^{-\gamma x}$ we get

$$-\lambda'(t)e^{-\gamma x} + \sup_{p\in\mathbb{R}}\left[-\frac{1}{2}p^2\sigma^2\gamma^2 + p\,\mu\,\gamma\right]\lambda(t)e^{-\gamma x} = 0 \text{ on } [0,T)\times\mathbb{R},\ \lambda(T) = 1.$$

We can divide by $e^{-\gamma x} \neq 0$ the HJB equation will hold provided that

$$-\lambda'(t) + \sup_{p\in\mathbb{R}}\left[-\frac{1}{2}p^2\sigma^2\gamma^2 + p\,\mu\,\gamma\right]\lambda(t) = 0 \text{ on } [0,T),\ \lambda(T) = 1.$$

The supremum is attained for $p^* = \frac{\mu}{\sigma^2\gamma}$ since the expression we are maximizing is quadratic in p with negative leading order term. Thus $\lambda'(t) = \beta\lambda(t)$ and $\lambda(T) = 1$ with

$$\beta := -\tfrac{1}{2}(p^*)^2\sigma^2\gamma^2 + p^*\,\mu\,\gamma = \tfrac{1}{2}\tfrac{\mu^2}{\sigma^2}\,.$$

We can solve the ODE for λ to obtain

$$\lambda(t) = e^{-\beta(T-t)}$$

and hence our candidate value function and control are

$$v(t,x) = -e^{-\frac{1}{2}\frac{\mu^2}{\sigma^2}(T-t)}e^{-\gamma x} \text{ and } p^* = \frac{\mu}{\sigma^2\gamma}\,.$$

We now need to use Theorem 4.11 to be able to confirm that these are indeed the value function and optimal control.

First of all the solution for optimal X^* always exists since we just need to integrate in the expression (4.17) taking $\pi_t := p^*$. We note that the resulting process is Gaussian.

Now $\partial_x v(s, X_s^*) = \lambda(t)\,\gamma e^{-\gamma X_s^*}$. We can now use what we know about moment generating functions of normal random variables to conclude that

$$\int_t^T \lambda(s)^2\, e^{-2\gamma X_s^*}\, ds < \infty.$$

The process

$$\bar{t} \mapsto \int_t^{\bar{t}} \lambda(s)\, e^{-\gamma X_s^*}\, dW_s$$

is thus a true martingale and the verification is complete.

Solution (to Exercise 4.19)**.**

$$\psi(t) = 0, \qquad \phi(t) = \frac{\sigma^2}{Ce^{\sigma^2 t/2} - 1}, \qquad C = (1+\sigma^2)e^{-\sigma^2 T/2}.$$

5 Applications in algorithmic trading and market making

In this section we extend results from Section 4 (Bellman principle aka DPP and the Hamilton–Jacobi–Bellman equation) to the simplest possible jump-diffusion setting. This will allow us to talk about controlled counting processes (integer valued) which is useful for modelling algorithmic trading and market making problems. Those who find this interesting can find lot more applications and details in [4].

5.1 Poisson process with controlled jump intensity

A stochastic process $(N_t)_{t\geq 0}$ is a homogenous Poisson process with intensity $\lambda > 0$ if the following conditions hold true.

1. $N_0 = 0$, $N_t \in \mathbb{Z}$ and for any $s \leq t$ we have $N_s \leq N_t$.
2. It has independent increments: for $t_1 \leq t_2 \leq t_3 \leq t_4$ the random variables $N_{t_4} - N_{t_3}$ is independent of $N_{t_2} - N_{t_1}$.
3. For any $t \geq 0$ and $\delta > 0$ we have

$$\mathbb{P}(N_{t+\delta} - N_t = n) = \frac{(\lambda\delta)^n}{n!}e^{-\lambda\delta}\,, \;\; n \in \{0\} \cup \mathbb{N}\,.$$

We see that in any infinitesimally small interval $[t, t+dt)$ the process either stays at a given value (level) or jumps up with a jump size 1 (so it stays integer-valued and increasing as per point 1). We can write $\mathbb{P}(N_{t+dt} = n+1 | N_t = n) = \lambda\, dt$. It can be shown that the process has a "cadlag"[12] modification, that is there is a modification which is continuous from the right (i.e. $\lim_{\delta \searrow 0} N_{t+\delta} = N_t$) with left-hand limits (i.e. $\lim_{\delta\searrow 0} N_{t-\delta}$ exists and we call the limit N_{t-}).

Condition 3. says that the increment $N_{t+\delta} - N_t$ is distributed as Poisson random variable with intensity $\lambda\delta$. This implies that $\mathbb{E}\big[N_{t+\delta} - N_t\big] = \lambda\delta$ which in turn, together with condition 2. implies that the process $(\hat{N}_t)_{t\geq 0}$ defined $\hat{N}_t := N_t - \lambda t$ is a martingale.

[12] It stands for continue à droite et limite gauche i.e. right continuous with left hand limits.

We also see that the process is generating a sequence of jump times $(\tau_i)_{i=1}^{\infty}$ and that $\tau_{i+1} - \tau_i \sim \mathrm{Exp}(\lambda)$.

Recall that for any process the infinitesimal generator is the operator which takes a bounded measurable function f and outputs

$$(Lf)(k) := \lim_{\delta \searrow 0} \frac{1}{\delta} \mathbb{E}\left[f(N_{t+\delta}) - f(N_t) | N_t = k\right] .$$

For the homogenous Poisson process with intensity $\lambda > 0$ this can be calculated as follows. First note that

$$\begin{aligned}
\mathbb{E}\left[f(N_{t+\delta}) - f(N_t) | N_t = k\right] &= \sum_{n=0}^{\infty} \frac{(\lambda\delta)^n}{n!} e^{-\lambda\delta} \big(f(k+n) - f(k)\big) \\
&= e^{-\lambda\delta} \lambda\delta \big(f(k+1) - f(k)\big) + e^{-\lambda\delta} \sum_{n=2}^{\infty} \frac{(\lambda\delta)^n}{n!} \big(f(k+n) - f(k)\big) .
\end{aligned}$$

We trivially see that

$$\frac{1}{\delta} e^{-\lambda\delta} \lambda\delta \big(f(k+1) - f(k)\big) \to \lambda\big(f(k+1) - f(k)\big) \text{ as } \delta \searrow 0 .$$

What remains is to consider

$$\begin{aligned}
I_\delta &:= \left| e^{-\lambda\delta} \sum_{n=2}^{\infty} \frac{(\lambda\delta)^n}{n!} \big(f(k+n) - f(k)\big) \right| \leq 2\|f\|_\infty e^{-\lambda\delta} \left(\sum_{n=0}^{\infty} \frac{(\lambda\delta)^n}{n!} - 1 - \lambda\delta \right) \\
&= 2\|f\|_\infty e^{-\lambda\delta} (e^{\lambda\delta} - 1 - \lambda\delta) ,
\end{aligned}$$

where in the last equality we recongized the Taylor series expansion of the exponential function. Using Taylor's theorem again we see that

$$2\|f\|_\infty e^{-\lambda\delta}(e^{\lambda\delta} - 1 - \lambda\delta) = 2\|f\|_\infty e^{-\lambda\delta} \int_0^{\lambda\delta} e^y (\lambda\delta - y)\, dy \leq 2\|f\|_\infty e^{-\lambda\delta} \int_0^{\lambda\delta} e^{\lambda\delta} (\lambda\delta - y)\, dy = \lambda^2 \|f\|_\infty \delta^2 .$$

Hence $\frac{1}{\delta} I_\delta \to 0$ as $\delta \searrow 0$ and hence

$$(Lf)(k) = \lambda\big(f(k+1) - f(k)\big)$$

is the infinitesimal generator for a homogenous Poisson process with intensity $\lambda > 0$.

Let's now have a look at what an Itô formula might look like for this pure jump process for a function $f = f(t, k)$ which is bounded, measurable and continuously differentiable in t. First let us note that $f(t, N_{t+\delta}) - f(t, N_t)$ converges to either 0 if $\Delta N_t := N_t - N_{t-} = 0$ or to $f(t, N_{t-} + 1) - f(N_{t-})$ if $\Delta N_t = 1$. Next note that from Taylor's theorem applied to the time component we get

$$\begin{aligned}
f(t+\delta, N_{t+\delta}) - f(t, N_{t-}) &= f(t+\delta, N_{t+\delta}) - f(t, N_{t+\delta}) + f(t, N_{t+\delta}) - f(t, N_{t-}) \\
&= f(t, N_{t+\delta}) - f(t, N_{t-}) + (\partial_t f)(t, N_{t+\delta})\delta + \tfrac{1}{2}(\partial_t^2 f)(t, N_{t+\delta})\delta^2 + \cdots
\end{aligned}$$

Given $T > 0$ and $M \in \mathbb{N}$ we take $\delta = T/N$ and get

$$\begin{aligned}
f(T, N_T) - f(0, N_0) &= \sum_{n=1}^{M} f(n\delta, N_{n\delta}) - f((n-1)\delta, N_{(n-1)\delta-}) \\
&= \sum_{n=1}^{M} f((n-1)\delta, N_{n\delta}) - f((n-1)\delta, N_{(n-1)\delta-}) + \sum_{n=1}^{M} (\partial_t f)((n-1)\delta, N_{n\delta})\delta \\
&\quad + \delta \sum_{n=1}^{M} \tfrac{1}{2}(\partial_t^2 f)((n-1)\delta, N_{n\delta})\delta + \cdots .
\end{aligned}$$

Taking the limit as $M \to \infty$ we get

$$f(T, N_T) - f(0, N_0) = \sum_{\Delta N_t > 0,\ 0 \le t \le T} \big(f(t, N_t) - f(t, N_{t-})\big) + \int_0^T (\partial_t f)(t, N_t)\, dt\,.$$

We will write

$$\sum_{\Delta N_t > 0,\ 0 \le t \le T} \big(f(t, N_t) - f(t, N_{t-})\big) = \int_0^T \big(f(t, N_{t-} + 1) - f(t, N_{t-})\big)\, dN_t\,.$$

At first glance the above doesn't look like the Itô formula we're used to. We expect to see

$$f(T, X_T) - f(0, X_0) = \int_0^T (\partial_t + \text{“infinitesimal generator of } X\text{”}) f(t, X_t)\, dt + \text{“local martingale term”}\,.$$

But in fact we can write the above in this way by recalling that $d\hat{N}_t = dN_t - \lambda\, dt$:

$$f(T, N_T) - f(0, N_0) = \int_0^T \big(f(t, N_{t-} + 1) - f(t, N_{t-})\big)\, dN_t + \int_0^T (\partial_t f)(t, N_t)\, dt$$
$$= \int_0^T \big(f(t, N_{t-} + 1) - f(t, N_{t-})\big)\, d\hat{N}_t + \int_0^T \big(f(t, N_{t-} + 1) - f(t, N_{t-})\big)\lambda\, dt + \int_0^T (\partial_t f)(t, N_t)\, dt$$

and so

$$f(T, N_T) - f(0, N_0) = \int_0^T (\partial_t f + Lf)(t, N_t)\, dt + \int_0^T \big(f(t, N_{t-} + 1) - f(t, N_{t-})\big)\, d\hat{N}_t\,.$$

Let us move away from $\lambda > 0$ constant and allow the intensity to be itself a stochastic process $(\lambda_t)_{t \ge 0}$ adapted to some filtration $(\mathcal{F}_t)_{t \ge 0}$. A stochastic process $(N_t)_{t \ge 0}$ is called doubly stochastic Poisson process (or a Cox process) with stochastic intensity $(\lambda_t)_{t \ge 0}$ if the following conditions hold true.

1. $N_0 = 0$, $N_t \in \mathbb{Z}$ and for any $s \le t$ we have $N_s \le N_t$.
2. It has independent increments: for $t_1 \le t_2$ the random variable $N_{t_2} - N_{t_1}$ is independent of $\mathcal{F}_{t_1 -}$.
3. For any $t \ge 0$ and $\delta > 0$ we have

   $$\mathbb{P}(N_{t+\delta} - N_t = n) = \frac{\Lambda(t, t+\delta)^n}{n!} e^{-\Lambda(t, t+\delta)}\,,\ \ n \in \{0\} \cup \mathbb{N}\,,$$

   where

   $$\Lambda(t, t+\delta) := \int_t^{t+\delta} \lambda_s\, ds$$

We can write the last condition as $\mathbb{P}(N_{t_- + dt} - N_{t_-} > 0) = \lambda_{t_-}\, dt$. Again we define $\hat{N}_t := N_t - \int_0^t \lambda_s\, ds$ and note that $(\hat{N}_t)_{t \ge 0}$ is a martingale.

We can carry out similar analysis to what we've done in the case of constant intensity. In particular it would be possible to prove the following Itô formula:

Lemma 5.1. *Let $N = (N_t)_{t\geq 0}$ be a doubly stochastic Poisson process with intensity $(\lambda_t)_{t\geq 0}$. Let $f = f(t,k)$ be a measurable function which is continuously differentiable in t. Then*

$$f(T,N_T) - f(0,N_0) = \int_0^T (\partial_t f + L_t f)(t,N_t)\,dt + \int_0^T \big(f(t,N_{t-}+1) - f(t,N_{t-})\big)\,d\hat{N}_t\,,$$

where $(L_t f)(t,k) = \lambda_t\big(f(t,k+1) - f(t,k)\big)$.

Note that while $(N_t)_{t\geq 0}$ has finite variation it has quadratic variation equal to itself so $dN_t\,dN_t = dN_t$ and $dt\,dN_t = dN_t\,dt = 0$. Moreover if W is a Wiener process independent of $(N_t)_{t\geq 0}$ then $dW_t\,dN_t = dN_t\,dW_t = 0$.

Theorem 5.2. *Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued process given by*

$$dX_t = \alpha_t\,dt + \beta_t\,dW_t + \gamma_t dN_t\,,$$

where $(\alpha_t)_{t\geq 0}$ is $\mathbb{R}^d$-valued, $(\beta_t)_{t\geq 0}$ is $\mathbb{R}^{d\times d'}$-valued, $(W_t)_{t\geq 0}$ is an $\mathbb{R}^{d'}$-valued Wiener process, $(\gamma_t)_{t\geq 0}$ is $\mathbb{R}^{d\times p}$-valued and $(N_t)_{t\geq 0}$ is $\mathbb{R}^p$-valued Poisson process with intensities $(\lambda_t)_{t\geq 0}$ that are $\mathbb{R}^p$-valued. Let $dX^c_t = \alpha_t\,dt + \beta_t\,dW_t$. Then for $f \in C^{1,2}([0,\infty)\times\mathbb{R}^d)$ we have

$$\begin{aligned} df(t,X_t) = \partial_t f(t,X_t)\,dt + \sum_{i=1}^d \partial_{x^i} f(t,X_t) d(X^c)^i_t + \frac{1}{2}\sum_{i,j=1}^d \partial_{x^i x^j} f(t,X_t)\,d(X^c)^i_t\,d(X^c)^j_t \\ + \sum_{j=1}^p \Big(f(t,X_{t-} + \gamma^{\cdot j}_t) - f(t,X_{t-})\Big)\,dN^j_t\,. \end{aligned}$$

Here $\gamma^{\cdot j}_t$ denotes the j-th column vector (in $\mathbb{R}^d$) of the matrix γ_t.

Example 5.3. Let us consider

$$dX_t = X_t\Big(\mu dt + \sigma dW_t + \kappa(dN^u_t - dN^d_t)\Big)\,, X_0 = x > 0 \in \mathbb{R}$$

μ, σ, $\kappa \in (0,1)$ and $\lambda > 0$ real constants with N^u and N^d two independent Poisson processes with intensity λ. Let $N_t = N^u_t - N^d_t$. Then by the Itô formula (Theorem 5.2) we have (ignoring for now that $x \mapsto \ln(x)$ isn't $C^2(\mathbb{R})$)

$$\begin{aligned} d\ln(X_t) &= \frac{1}{X_t}X_t(\mu\,dt + \sigma\,dW_t) - \frac{1}{2}\frac{1}{X_t^2}\sigma^2 X_t^2\,dW_t\,dW_t + (\ln(X_t) - \ln(X_{t-})\,dN_t \\ &= (\mu - \tfrac{1}{2}\sigma^2)\,dt + \sigma\,dW_t \\ &\qquad + \Big(\ln(X_t + \kappa X_t) - \ln(X_t)\Big)\,dN^u_t + \Big(\ln(X_t - \kappa X_t) - \ln(X_t)\Big)\,dN^d_t \\ &= (\mu - \tfrac{1}{2}\sigma^2)\,dt + \sigma\,dW_t + \ln\left(\tfrac{X_t + \kappa X_t}{X_t}\right)dN^u_t + \ln\left(\tfrac{X_t - \kappa X_t}{X_t}\right)dN^d_t \\ &= (\mu - \tfrac{1}{2}\sigma^2)\,dt + \sigma\,dW_t + \ln(1+\kappa)\,dN^u_t + \ln(1-\kappa)\,dN^d_t\,. \end{aligned}$$

Hence

$$\ln X_t - \ln(X_0) = (\mu - \tfrac{1}{2}\sigma^2)t + \sigma W_t + \ln(1+\kappa)N^u_t + \ln(1-\kappa)N^d_t$$

and so

$$X_t = x\exp\Big((\mu - \tfrac{1}{2}\sigma^2)t + \sigma W_t + \overline{\kappa}N^u_t - \underline{\kappa}N^d_t\Big)\,,$$

where $\overline{\kappa} := \ln(1+\kappa) > 0$ and $\underline{\kappa} := -\ln(1-\kappa) > 0$.

In general we will work with intensities of the form $\lambda_t = \lambda(t, N_{t-}, \alpha_{t-})$ where $\lambda = \lambda(t,n,a)$ is a measurable function and $\alpha = (\alpha_t)_{t\geq 0}$ is some adapted stochastic process which will play the role of control.

5.2 Controlled diffusions with jumps

We will now take the doubly stochastic Poisson process introduced in Section 5.1 and use it drive a controlled diffusion with jumps, building on Section 3.2.

We assume there is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\mathbb{F} := (\mathcal{F}_t)_{t\in[0,T]}$ for some fixed $T > 0$.

Assume that A is the space of actions (complete separable metric space or just a subset of $\mathbb{R}^m$). Admissible controls will be processes α prog. meas. w.r.t. $\mathbb{F}$ with some integrability properties (e.g. m'-moments, this will depend on applications). We will denote the set of admissible controls by $\mathcal{A}$.

Take N^α to be a d''-dimensional doubly stochastic Poisson process controlled by α with intensity $\lambda_t := \lambda(t, N_{t-}, \alpha_{t-})$ for some $\mathbb{R}^{d''}$-valued measurable function $\lambda = \lambda(t, n, a)$.

Let W be a d'-dimensional Wiener process with increments independent of $\mathbb{F}$ (so that it's a martingale w.r.t. this filtration).

The controlled SDE for $\mathbb{R}^d$-valued X with jumps we wish to consider is

$$\begin{aligned} dX_s^{t,x,\alpha} &= b(s, X_s^{t,x,\alpha}, \alpha_s)\, ds + \sigma(s, X_s^{t,x,\alpha}, \alpha_s)\, dW_s \\ &\quad + \gamma(s, X_s^{t,x,\alpha}, \alpha_s)\, dN_s^\alpha\,, \quad s \in [t,T]\,, \quad X_t^{t,x,\alpha} = x\,. \end{aligned} \tag{5.1}$$

So $b : [0,T]\times\mathbb{R}^d\times A \to \mathbb{R}^d$, $\sigma : [0,T]\times\mathbb{R}^d\times A \to \mathbb{R}^{d\times d'}$ and $\gamma : [0,T]\times\mathbb{R}^d\times A \to \mathbb{R}^{d\times d''}$. We shall not go into details on existence and uniqueness of solutions for these equations like in Section 3.2 as we wish to see the applications.

Assume that we are given reward functions $b : [0,T]\times\mathbb{R}^d\times A \to \mathbb{R}$ and $g : \mathbb{R}^d \to \mathbb{R}$. Our objective is to maximize

$$J(t,x,\alpha) = \mathbb{E}^{t,x}\left[\int_t^T f(s, X_s^\alpha, \alpha_s)\, ds + g(X_T^\alpha)\right]\,, \tag{5.2}$$

over $\alpha \in \mathcal{A}$, subject to X^α solving (5.1). We will also define the value function for this control problem as

$$v(t,x) := \sup_{\alpha\in\mathcal{A}} J(t,x,\alpha)\,. \tag{5.3}$$

5.3 Bellman principle and Bellman PDE for diffusions with jumps

As we have seen in Section 4 a very useful tools for solving the control problems and for devising algorithms are the Bellman principle / Dynamic Programming Principle (DPP) and the Bellman or Hamilton–Jacobi–Bellman PDE. We will now state, without proofs, what they are for the case of controlled jump diffusions.

Theorem 5.4 (Bellman principle for controlled jump diffusions)**.** *Let appropriate assumptions on b, σ, γ, f and g hold. Let $(t,x) \in [0,T]\times\mathbb{R}^d$. Let τ be a stopping time such that $t \le \tau \le T$. Then*

$$v(t,x) = \sup_{\alpha\in\mathcal{A}} \mathbb{E}^{t,x}\left[\int_t^\tau f(s, X_s^\alpha, \alpha_s)\, ds + v(\tau, X_\tau^\alpha)\right]\,.$$

Given the preceding discussion we can see that the infinitesimal generator is

$$\begin{aligned} L_t^a v(t,x) &= b(t,x,a,)\partial_x v(t,x) + \tfrac{1}{2}\mathrm{tr}\big[(\sigma\sigma^\top)(t,x,a)\partial_{xx} v(t,x)\big] \\ &\quad + \sum_{j=1}^{d''} \lambda_j(t,x,a)\Big(v\big(t, x + \gamma_{\cdot j}(t,x,a)\big) - v(t,x)\Big) \end{aligned}$$

for $v \in C^{1,2}((0,T)\times\mathbb{R}^d)$. We now have what we need to state the Bellman PDE we expect v given by (5.3) to satisfy.

Theorem 5.5 (Hamilton–Jacobi–Bellman PDE for controlled jump diffusions)**.** *Assume that the Bellman principle holds and b, σ and γ are sufficiently regular. If v given by* (5.3) *is in $C^{1,2}((0,T)\times\mathbb{R}^d)$ then*

$$\begin{aligned} \partial_t v + \sup_{a\in A}\Big(L^a v + f^a\Big) = 0 \quad \text{on } [0,T)\times\mathbb{R}^d \\ v(T,x) = g(x) \quad \forall x\in\mathbb{R}^d\,. \end{aligned} \tag{5.4}$$

If we solve (5.4) then we should again be applying a verification theorem similar to Theorem 4.11 to ensure that we really obtained the value function (5.3) for our control problem. Since we haven't given conditions under which (5.1) has unique solution we won't be able to use such a theorem - hence we omit it. But bear in mind that without the verification we don't really know what we found by solving (5.4).

Exercise 5.6 (Merton's problem with jumps and logarithmic utility)**.** Let us consider a risk-free asset $dB_u = rB_u\,dt$ for $u\in[t,T]$ and $B_t = b$ and a risky asset of the form

$$dS_u = S_u\Big(\mu du + \sigma dW_u + \kappa(dN^u_u - dN^d_u)\Big)\,,\;\; u\in[t,T]\,,\;\; S_t = S > 0\,.$$

Here W is a Wiener process and N^u and N^d are Poisson processes with intensity $\lambda \geq 0$; all three are independent.

Let $X_T^{t,x,\nu}$ be the wealth of an investor who started with $x>0$ at $t\in[0,T]$ following strategy ν which specifies the fraction of total wealth invested in the risky asset. Our aim in this exercise is to find

$$v(t,x) := \sup_\nu \mathbb{E}_{t,x,\nu}\left[\ln(X_T)\right]\,.$$

i) Assume an investor invests ξ "units" (can be fraction / negative) into the risky asset and the remaining wealth in the risk-free asset. Show that their wealth evolves as

$$dX_u = X_u\Big(\nu_u(\mu-r)+r\Big)\,du + X_u\nu_u\Big(\sigma\,dW_u + \kappa(dN^u_u - dN^d_u)\Big)\,,\;\; u\in[t,T]\,,\;\; X_t = x\,,$$

where $\nu=\nu_u$ denotes the fraction of wealth invested in the risky asset.

ii) Use Itô formula to show that with $Y_t := \ln X_t$ we have

$$dY_u = \Big(\nu_u(\mu-r)+r-\tfrac12\sigma^2\nu_u^2\Big)\,dt + \nu_u\sigma\,dW_u + \ln(1+\kappa\nu_{u-})dN^u_u + \ln(1-\kappa\nu_{u-})dN^d_u\,.$$

iii) Thus we may equivalently search for

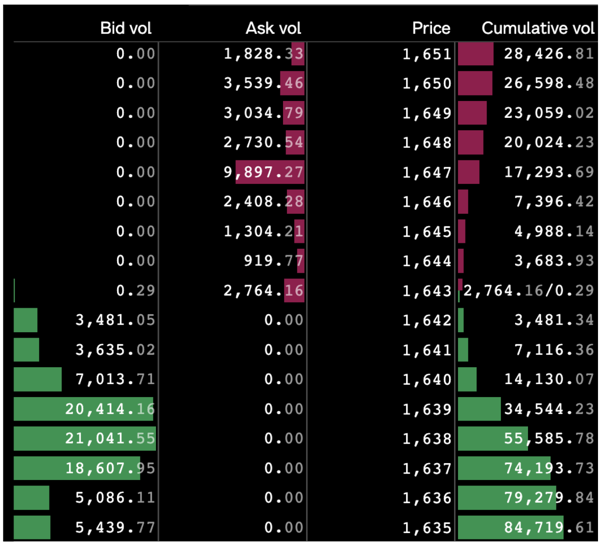
$$w(t,y) := \sup_\nu \mathbb{E}_{t,y,\nu}\left[Y_T\right]$$

with $y=\ln x$ so that $v(t,x) = w(t,\ln x)$. Write down the Bellman equation for this problem, in particular write what the infinitesimal generator for the process Y is.

iv) Let us assume $w(t,y) = \psi(t)y + \gamma(t)$ with $\psi,\gamma\in C^1([0,T])$ and $\psi>0$. Use this ansatz to solve the problem up to the static maximization

$$\hat{A}_{\mu,r,\sigma,\kappa,\lambda} := \max_a\Big[a(\mu-r) + r - \tfrac12\sigma^2a^2 + \lambda\ln(1-a^2\kappa^2)\Big]\,.$$

Since this doesn't, in general, have a neat solution write down v in terms of $\hat{A}$, T, t, and x.



Bid vol	Ask vol	Price	Cumulative vol
0.00	1,828.33	1,651	28,426.81
0.00	3,539.46	1,650	26,598.48
0.00	3,034.79	1,649	23,059.02
0.00	2,730.54	1,648	20,024.23
0.00	9,897.27	1,647	17,293.69
0.00	2,408.28	1,646	7,396.42
0.00	1,304.21	1,645	4,988.14
0.00	919.77	1,644	3,683.93
0.29	2,764.16	1,643	2,764.16/0.29
3,481.05	0.00	1,642	3,481.34
3,635.02	0.00	1,641	7,116.36
7,013.71	0.00	1,640	14,130.07
20,414.16	0.00	1,639	34,544.23
21,041.55	0.00	1,638	55,585.78
18,607.95	0.00	1,637	74,193.73
5,086.11	0.00	1,636	79,279.84
5,439.77	0.00	1,635	84,719.61

Figure 5.1: Typical limit order book (LOB). Left half: shows the cumulative ask volume (red) and the bid volume (green) at various price levels (y-axis) on either side of the mid price of about 1643. Right half: shows volumes at various bid (green) and ask (red) levels with mid-price about 1 643.

5.4 Optimal execution with limit orders

In this section we will consider the problem of optimal order execution in a limit order book (LOB) market placing limit orders only. This is similar to Section 1.3 where we considered optimal execution (liquidation) using market orders only.

A *market* order specifies whether it's to buy or to sell and then desired volume. It then executes against existing limit order volume on the order book at the *volume weighted average price (VWAP)* implied by the state of the book. E.g. in Figure 5.1 a market buy order for volume of 5000 will achieve VWAP of

$$(2764.16 \times 1643 + 919.77 \times 1644 + 1304.21 \times 1645 + 11.86 \times 1646)/5000 \approx 1643.71\,.$$

A *limit* order specifies whether it's a buy or sell order, price and volume. Once submitted it will be placed on the order book at the price level specified. An exception is if it's a buy order and there is already sell volume at that price or if it's a sell order and there is already buy volume at that level. In that case the portion of the volume that would cross trades and what remains (if any) is placed on the book.

Let us now consider the simplest possible model for inventory liquidation. We have

i) An agent wishing to sell $\mathcal{N}$ units of a certain asset until time $T > 0$.

ii) The market mid price is driven by $dS_r = \sigma dW_r$ with W a Wiener process, $\sigma > 0$ constant.

iii) The agent places sell LO for volume Δ at distance $(\delta_t)_{t\in[0,T]}$ taking values in $\mathbb{R}^+$ from the mid price (we ignore that LOBs have ticks). So the price at which the LO sits is $S_t + \delta_t$. We assume that the agent continuously amends the order so it stays at that level. This is the control. We must have that $\mathcal{N}$ is divisible by Δ.

iv) The number of market buy orders that arrive in the market is given by Poisson process $(M_t)_{t\in[0,T]}$ with rate $\lambda > 0$.

v) Not all market buy orders hit the volume of our agent. In fact their volume only trades with probability $e^{-\kappa\delta}$, with $\kappa > 0$ a constant. We call this the fill probability. If our agent's volume trades then they immediately replace it distance $\delta > 0$ from the mid.

vi) We count the market buy orders which hit our agent's volume as $(N_t^\delta)_{t\in[0,T]}$. This is a doubly stochastic Poisson process with the controlled intensity $\lambda e^{-\kappa\delta_t}$.

As a consequence the agent's inventory, which is $(Q_t)_{t\in[0,T]}$ with $Q_0 = \mathcal{N}$ evolves as

$$dQ_t^\delta = -\Delta dN_t^\delta \,.$$

In fact it is $Q_t^\delta = \mathcal{N} - \Delta N_t^\delta$. The agent's cash balance changes as

$$dX_t = \Delta(S_t + \delta)\, dN_t^\delta \,.$$

You may be asking: where did the order book enter the picture? The answer is in the fill probability $e^{-\kappa\delta}$. If you look at Figure 5.1 it is clear that the further the sell order is from mid price (on the red side of the book) the lower the probability that market orders will "eat up" the volume in front of it and that it will trade. There is more discussion in [4, Chapter 8.1] justifying this modelling choice.

Assume agent wishes to maximize their cash while disposing of as much inventory as possible. Let $\tau_{t,q,\delta} := T \wedge \min(s \geq t : Q_s^{t,q,\delta} = 0)$. Then their objective can be expressed as

$$J(t,S,x,q,\delta) := \mathbb{E}_{t,S,x,q,\delta}\left[X_\tau^\delta + S_\tau Q_\tau^\delta - \alpha(Q_\tau^\delta)^2\right],$$

where $\alpha > 0$ captures the unsold inventory penalty. The agent's tradeoff is thus clear: choose high δ and sell at (possibly much) better price than midprice but with high δ run the risk of not selling all by time T. The value function is

$$v(t,S,x,q) = \sup_{\delta\in\mathcal{A}} J(t,S,x,q,\delta)\,.$$

The admissible strategies are progressively measurable, square integrable and non-negative. Using Theorem 5.4 we can write the corresponding Bellman PDE on the domain in space $\mathcal{D} = \mathbb{R}\times\mathbb{R}^+\times\{\Delta, 2\Delta,\ldots,\mathcal{N}\}$ as

$$\begin{aligned}
\partial_t v + \tfrac{1}{2}\sigma^2\partial_{SS}v + \sup_{\delta\geq 0}\left[\lambda e^{-\kappa\delta}\big(v(t,S,x+\Delta(S+\delta),q-\Delta) - v\big)\right] = 0 \ \text{ on } \ [0,T)\times\mathcal{D}\,,\\
v(t,x,S,0) = x \ \ \forall t\in[0,T)\,, x\in\mathbb{R}^+\,, S\in\mathbb{R}\,,\\
v(T,x,S,q) = x + qS - \alpha q^2 \ \ \forall (x,S,q)\in\mathcal{D}\,.
\end{aligned}$$

The boundary at $t = T$ corresponds to the situation when the agent ran out of time, they have cash x, accounting value of qS for the unsold inventory q and penalty αq^2 for the unsold inventory. The boundary at $q = 0$ corresponds to the situation when the agent successfully sold everything and they hold cash x.

An ansatz which allows us to match the boundary conditions is

$$v(t,S,x,q) = x + qS + \theta(t,q)$$

as long as $\theta(T,q) = -\alpha q^2$. Let us now solve the Bellman PDE with this ansatz. We note that

$$x + \Delta(S+\delta) + S(q-\Delta) + \theta(t,q-\Delta) - x - qS - \theta = \Delta\delta + \theta(t,q-\Delta) - \theta$$

and hence

$$\begin{aligned}\partial_t\theta + \sup_{\delta\geq 0}\Big[\lambda e^{-\kappa\delta}\big(\Delta\delta + \theta(t,q-\Delta) - \theta\big)\Big] = 0 \;\text{ on }\; [0,T)\times\mathcal{D}\,,\\ \theta(t,0) = 0 \;\; \forall t\in[0,T),\\ \theta(T,q) = -\alpha q^2 \;\; \forall q = \{0,\Delta,2\Delta,\ldots,\mathcal{N}\}\,.\end{aligned}$$

Letting $g(t,q) = \frac{\theta(t,q-\Delta)-\theta}{\Delta}$ we observe that we need to maximize the function

$$\delta \mapsto e^{-\kappa\delta}(\delta + g)\,.$$

Calculating derivatives we see this is concave on $[0,\infty)$ and from the first order condition we see that

$$e^{-\kappa\delta} - (\delta+g)\kappa e^{-\kappa\delta} = 0$$

i.e.

$$\delta^*(t,q) = \frac{1}{\kappa} - g(t,q)\,.$$

To substitute this back into the equation for θ we see that

$$\sup_{\delta\geq 0}\Big[e^{-\kappa\delta}\big(\delta + g(t,q)\big)\Big] = e^{-1}e^{\kappa g(t,q)}\frac{1}{\kappa}$$

and so

$$\begin{aligned}\partial_t\theta + \frac{1}{\kappa}\lambda\Delta e^{-1}\exp\Big(\kappa\frac{\theta(t,q-\Delta)-\theta}{\Delta}\Big) = 0 \;\text{ on }\; [0,T)\times\{\Delta,2\Delta,\ldots,\mathcal{N}\}\,,\\ \theta(t,0) = 0 \;\; \forall t\in[0,T),\\ \theta(T,q) = -\alpha q^2 \;\; \forall q = \{0,\Delta,2\Delta,\ldots,\mathcal{N}\}\,.\end{aligned}$$

There is one more trick needed and we are done: let $w(t,q) := e^{\kappa\Delta^{-1}\theta(t,q)}$ so that $\frac{\kappa}{\Delta}\theta(t,q) = \ln w(t,q)$. Then $\partial_t\theta = \frac{1}{\kappa}\frac{1}{w}\partial_t w$ and the equation for w is

$$\frac{\Delta}{\kappa}\frac{1}{w(t,q)}\partial_t w(t,q) + \frac{1}{\kappa}\lambda\Delta e^{-1}\frac{w(t,q-\Delta)}{w(t,q)} = 0 \;\text{ on }\; [0,T)\times\{\Delta,2\Delta,\ldots,\mathcal{N}\}$$

which is

$$\partial_t w(t,q) + \lambda e^{-1}w(t,q-\Delta) = 0 \;\text{ on }\; [0,T)\times\{\Delta,2\Delta,\ldots,\mathcal{N}\}\,.$$

The terminal condition is $w(T,q) = e^{-\kappa\Delta^{-1}\alpha q^2}$ and $w(t,0) = 1$.

5.5 Optimal limit order spread in market making

We will now consider a problem which arises in market making. A market maker is an agent who's objective is to stay market neutral (so ideally 0 position) while profiting from the bid-ask spread while providing liquidity (volume that can be traded) at all (or most) times. A possible model is the following.

i) We fix a finite time horizon $T > 0$.

ii) The market mid price is driven exogenously by $dS_r = \sigma dW_r$ with W a Wiener process, $\sigma > 0$ constant.

iii) The agent places a buy LO for volume Δ at distance $(\delta^b_t)_{t\in[0,T]}$ and sell LO for volume Δ at distance $(\delta^a_t)_{t\in[0,T]}$ from the mid price (we ignore that LOBs have ticks). Both δ^b_t and δ^t take values in $\mathbb{R}^+$. So the price at which the buy LO sits is $S_t - \delta^b_t$ and the price at which the sell LO sits is $S_t + \delta^a_t$ (buy low, sell high). We assume that the agent continuously amends the order so it stays at that level as the mid price moves. In this case $\delta_t = (\delta^b_t, \delta^a_t)$ is the control.

iv) The number of market buy / sell orders arriving in the market is given by two independent Poisson process $(M^b_t)_{t\in[0,T]}$ and $(M^a_t)_{t\in[0,T]}$ with rates $\lambda^b > 0$, $\lambda^a > 0$

v) Not all market buy orders hit the volume of our agent. In fact their volume only trades with a buy LO with probability $e^{-\kappa^b\delta}$, with $\kappa^b > 0$ a constant and with a sell LO with probability $e^{-\kappa^a\delta}$, with $\kappa^a > 0$ a constant. We call this the fill probabilities. If our agent's volume trades then they immediately replace it distance $\delta^b > 0$ or $\delta^a > 0$ from the mid.

vi) The agent is allowed to only carry finite inventory, so $Q_t \in [\underline{k}\Delta, \bar{k}\Delta]$ for some $\underline{k}, \bar{k} \in \mathbb{Z}$. In case $Q_t = \underline{k}\Delta$ we set $\delta^a_t = \infty$ and if $Q_t = \bar{k}\Delta$ we set $\delta^b_t = \infty$. With the fill probabilities above this means that we place no orders when we've hit the boundary and hence our inventory cannot decrease or increase respectively.

vii) We count the market buy orders which hit our agent's buy volume as $(N^{b,\delta}_t)_{t\in[0,T]}$ and those that hit our agent's sell volume as $(N^{a,\delta}_t)_{t\in[0,T]}$. These are doubly stochastic Poisson processes with the controlled intensities $\lambda e^{-\kappa^b\delta^b_t}$ and $\lambda e^{-\kappa^a\delta^a_t}$ respectively.

As a consequence the agent's inventory, which is $(Q_t)_{t\in[0,T]}$ with $Q_0 = \mathcal{N}$ evolves as

$$dQ^\delta_u = \Delta dN^{b,\delta}_u - \Delta dN^{a,\delta}_u\,,\ \ u \in [t,T]\,,\ \ Q_t = q\,.$$

The agent's cash balance changes as

$$dX^\delta_u = -\Delta(S_u - \delta^b_u)\,dN^{b,\delta}_u + \Delta(S_u + \delta^a_u)\,dN^{a\delta}_u\,,\ \ u \in [t,T]\,,\ X_t = x\,.$$

The objective of the agent should express that they want to maximize cash, they don't want to hold inventory, they account for the mark-to-market value of their inventory at time T and that they penalise any terminal time inventory holdings. Thus their aim is to maximize:

$$J(t,S,x,q,\delta) := \mathbb{E}_{t,S,x,q,\delta}\left[X_T + S_T\,Q_T - \alpha|Q_T|^2 - \phi\int_t^T |Q_u|^2\,du\right].$$

The value function of the control problem is

$$v(t,S,x,q) := \sup_\delta J(t,S,x,q,\delta)\,.$$

Recall that our state process is

$$\begin{aligned}
&dS_u = \sigma dW_u\,,\ \ u \in [t,T]\,,\ \ S_t = S\,,\\
&dX^\delta_u = -\Delta(S_u - \delta^b_u)\,dN^{b,\delta}_u + \Delta(S_u + \delta^a_u)\,dN^{a,\delta}_u\,,\ \ u \in [t,T]\,,\ X_t = x\,,\\
&dQ^\delta_u = \Delta dN^{b,\delta}_u - \Delta dN^{a,\delta}_u\,,\ \ u \in [t,T]\,,\ \ Q_t = q\,.
\end{aligned}$$

The infinitesimal generator is

$$
\begin{aligned}
L^{\delta} f(S,x,q) =& \frac{1}{2}\partial_{SS} f(S,x,q) \\
&+ \lambda^b e^{-\kappa^b \delta^b} \Big(f(S, x-\Delta(S-\delta^b), q+\Delta) - f(S,x,q) \Big) \\
&+ \lambda^a e^{-\kappa^a \delta^a} \left(f(S, x+\Delta(S+\delta^a), q-\Delta) - f(S,x,q) \right) .
\end{aligned}
$$

Hence the Bellman equation for v is

$$
\begin{aligned}
0 =& \partial_t v + \frac{1}{2}\partial_{SS} v - \phi |q|^2 \\
&+ \lambda^b \sup_{\delta^b} e^{-\kappa^b \delta^b} \Big(v(t,S,x-\Delta(S-\delta^b), q+\Delta) - v \Big) \mathbf{1}_{\{q<\bar{k}\Delta\}} \\
&+ \lambda^a \sup_{\delta^a} e^{-\kappa^a \delta^a} \left(v(t,S,x+\Delta(S+\delta^a), q-\Delta) - v \right) \mathbf{1}_{\{q>\underline{\mathrm{k}}\Delta\}} ,
\end{aligned}
\tag{5.5}
$$

in $[0,T)\times\mathbb{R}\times\mathbb{R}\times\{k\Delta : k \in [\underline{\mathrm{k}},\bar{k}]\cap\mathbb{Z}\}$ with the terminal time condition $v(T,S,x,q) = x + Sq - \alpha|q|^2$.

Solving the Bellman equation for the market making problem

We will assume that $\lambda^a = \lambda^b > 0$ and $\kappa^a = \kappa^b = \kappa$. We will use the ansatz $v(t,S,x,q) = x + Sq + \theta(t,q)$ with $\theta \in C^1([0,T])$ and $\theta(T,q) = -\alpha|q|^2$. Let us first work out the difference terms in (5.5). The first one is:

$$
\begin{aligned}
& v(t,S,x-\Delta(S-\delta^b),q+\Delta) - v = x - \Delta(S-\delta^b) + S(q+\Delta) + \theta(t,q+\Delta) - x - Sq - \theta(t,q) \\
&= -\Delta S + \Delta\delta^b + Sq + S\Delta + \theta(t,q+\Delta) - Sq - \theta(t,q) = \Delta\delta^b + \theta(t,q+\Delta) - \theta(t,q) .
\end{aligned}
$$

The second one is

$$
v(t,S,x+\Delta(S+\delta^a),q-\Delta) - v = \Delta\delta^a + \theta(t,q-\Delta) - \theta(t,q) .
$$

Hence the Bellman equation (5.5) transforms to

$$
\begin{aligned}
0 = \partial_t \theta - \phi|q|^2 &+ \Delta\lambda \sup_{\delta^b} e^{-\kappa\delta^b} \Big(\delta^b + g^b(t,q) \Big) \mathbf{1}_{\{q<\bar{k}\Delta\}} \\
&+ \Delta\lambda \sup_{\delta^a} e^{-\kappa\delta^a} \left(\delta^a + g^a(t,q) \right) \mathbf{1}_{\{q>\underline{\mathrm{k}}\Delta\}} ,
\end{aligned}
\tag{5.6}
$$

where

$$
g^b(t,q) = \tfrac{\theta(t,q+\Delta)-\theta(t,q)}{\Delta} , \quad g^a(t,q) = \tfrac{\theta(t,q-\Delta)-\theta(t,q)}{\Delta} .
$$

From the first order condition we see

$$
0 = \frac{d}{d\delta^b} e^{-\kappa\delta} \Big(\delta^b + g^b(t,q) \Big) = e^{-\kappa\delta} - \kappa e^{-\kappa\delta} \Big(\delta^b + g^b(t,q) \Big)
$$

and hence

$$
\delta_t^{b,*} = \frac{1}{\kappa} - g^b(t,q) .
$$

Similarly

$$
\delta_t^{a,*} = \frac{1}{\kappa} - g^a(t,q) .
$$

Substituting these back into (5.6) we have

$$0 = \partial_t\theta - \phi|q|^2 + \Delta\lambda e^{-1} e^{\kappa g^b(t,q)} \frac{1}{\kappa} \mathbf{1}_{\{q<\bar{k}\Delta\}} + \Delta\lambda e^{-1} e^{\kappa g^a(t,q)} \frac{1}{\kappa} \mathbf{1}_{\{q>\underline{\mathrm{k}}\Delta\}}$$

with $t \in [0,T)$, $q \in [\underline{\mathrm{k}}\Delta, \bar{k}\Delta] \cap \{k\Delta : k \in \mathbb{Z}\}$ and with the terminal condition $\theta(T,q) = -\alpha|q|^2$. While this is still a non-linear equation we can transform it into a linear one. Let $\theta(t,q) = \frac{\Delta}{\kappa} \log w(t,q)$ or equivalently $w(t,q) = e^{\kappa\Delta^{-1}\theta(t,q)}$. Notice that $\partial_t\theta(t,q) = \frac{\Delta}{\kappa} \frac{1}{w(t,q)} \partial_t w(t,q)$ and the equation for w is

$$0 = \frac{\Delta}{\kappa} \frac{1}{w} \partial_t w - \phi|q|^2 + \frac{\Delta}{\kappa} \lambda e^{-1} \frac{w(t,q+\Delta)}{w} \mathbf{1}_{\{q<\bar{k}\Delta\}} + \frac{\Delta}{\kappa} \lambda e^{-1} \frac{w(t,q-\Delta)}{w} \mathbf{1}_{\{q>\underline{\mathrm{k}}\Delta\}} ,$$

with $t \in [0,T)$, $q \in [\underline{\mathrm{k}}\Delta, \bar{k}\Delta] \cap \{k\Delta : k \in \mathbb{Z}\}$ and with the terminal condition $w(T,q) = e^{-\kappa\Delta^{-1}\alpha|q|^2}$. This simplifies to

$$0 = \partial_t w - \frac{\kappa}{\Delta} \phi|q|^2 w + \lambda e^{-1} w(t,q+\Delta) + \lambda e^{-1} w(t,q-\Delta) , \quad t \in [0,T) , \quad q \in (\underline{\mathrm{k}}\Delta, \bar{k}\Delta) \cap \{k\Delta : k \in \mathbb{Z}\} .$$

We can write this an ODE system. Let $N = \bar{k} - \underline{\mathrm{k}} + 1$. Let

$$\mathbf{A}_{mn} = \begin{cases} \lambda e^{-1} & \text{if } m = n-1 \text{ and } n \le N \\ -\frac{\kappa}{\Delta}\, \phi \,|m\,\Delta|^2 & \text{if } m = n \\ \lambda e^{-1} & \text{if } m = n+1 \text{ and } n \ge 1 . \end{cases} \tag{5.7}$$

Then $\mathbf{W}(t) = (w(t,k\Delta))_{k\in[\underline{\mathrm{k}},\bar{k}]\cap\mathbb{Z}}$ satisfies the linear ODE

$$\mathbf{W}'(t) + \mathbf{A}\mathbf{W}(t) = 0 , \quad t \in [0,T] ,$$

with $\mathbf{W}(T)_k = e^{-\kappa\Delta^{-1}\alpha|k\Delta|^2}$. Note that the solution is, in terms of the matrix exponential, $\mathbf{W}(t) = \exp((T-t)\mathbf{A})\mathbf{W}(T)$.

5.6 Solution to Exercises

Solution (to Exercise 5.6). i) An agent running a self-financing strategy in those two assets has portfolio evolution given by

$$\begin{aligned} dX_u &= \xi_u\, dS_u + \tfrac{X_u - \xi_u S_u}{B_u}\, dB_u \\ &= \tfrac{X_u \nu_u}{S_u}\, dS_u + \tfrac{X_u - X_u\nu_u}{B_u} dB_u \\ &= X_u\Big(\nu_u\big(\mu\, du + \sigma\, dW_u + \kappa(dN_u^u - dN_u^d)\big) + (1-\nu_u) r\, dt\Big) \\ &= X_u\Big(\nu_u\big(\mu - r\big) + r\Big)\, du + X_u\nu_u\Big(\sigma\, dW_u + \kappa(dN_u^u - dN_u^d)\Big) . \end{aligned}$$

ii) For $Y_t := \ln X_t$ we get

$$d\ln X_t = \frac{1}{X_t} dX_t^c - \frac{1}{2}\frac{1}{X_t^2} dX_t^c dX_t^c + (\ln X_t - \ln X_{t-}) dJ_t ,$$

where $dJ_t = X_t\nu_t\kappa(dN_t^u - dN_t^d)$. Hence

$$\begin{aligned} &(\ln X_t - \ln X_{t-}) dJ_t = (\ln X_t - \ln X_{t-}) X_t\nu_t\kappa(dN_t^u - dN_t^d) \\ &= X_t\nu_t\kappa(\ln X_t - \ln X_{t-}) dN_t^u - X_t\nu_t\kappa(\ln X_t - \ln X_{t-}) dN_t^d \\ &= (\ln(X_{t-} + X_{t-}\nu_{t-}\kappa) - \ln X_{t-}) dN_t^u + (\ln(X_{t-} - X_{t-}\nu_{t-}\kappa) - \ln X_{t-})) dN_t^d \\ &= \ln\Big(\tfrac{X_{t-} + X_{t-}\nu_{t-}\kappa - X_{t-}}{X_{t-}}\Big) dN_t^u + \ln\Big(\tfrac{X_{t-} - X_{t-}\nu_{t-}\kappa - X_{t-}}{X_{t-}}\Big) dN_t^d \\ &= \ln(1 + \kappa\nu_{t-}) dN_t^u + \ln(1 - \kappa\nu_{t-}) dN_t^d . \end{aligned}$$

Hence

$$dY_u = \Big(\nu_u(\mu - r) + r - \tfrac{1}{2}\sigma^2\nu_u^2\Big)\, dt + \nu_u\sigma\, dW_u + \ln(1 + \kappa\nu_{u-}) dN_u^u + \ln(1 - \kappa\nu_{u-}) dN_u^d .$$

iii) The diffusion generator for Y is

$$L^a f(t,y) = \big(a(\mu-r)+r-\tfrac{1}{2}\sigma^2a^2\big)\partial_y f(t,y)+\tfrac{1}{2}\sigma^2a^2\partial_{yy}f(t,y) \\ +\lambda\Big(f(t,y+\ln(1+a\kappa))-f(t,y)\Big)+\lambda\Big(f(t,y+\ln(1-a\kappa))-f(t,y)\Big)\,.$$

The Bellman equation is given by Theorem 5.5 and since there is no running gain it reads

$$\partial_t w + \sup_a L^a w = 0 \;\text{ in }\; [0,T)\times(0,\infty)\,,$$

with the terminal condition $w(T,y)=y$.

iv) We need $\psi(T)=1$, $\gamma(T)=0$ and the Bellaman equation becomes (pulling $\psi>0$ out of the supremum over a):

$$\psi'(t)y+\gamma'(t)+\psi(t)\sup_a\Big[a(\mu-r)+r-\tfrac{1}{2}\sigma^2a^2+\lambda\Big((y+\ln(1+a\kappa))-2y+(y+\ln(1-a\kappa))\Big)\Big]=0$$

for $(t,y)\in[0,T)\times\mathbb{R}^+$. Note that

$$\lambda\Big((y+\ln(1+a\kappa))-2y+(y+\ln(1-a\kappa))\Big)=\lambda\ln(1-a^2\kappa^2)\,.$$

We thus need to maximize

$$a\mapsto a(\mu-r)+r-\tfrac{1}{2}\sigma^2a^2+\lambda\ln(1-a^2\kappa^2)\,.$$

Two observations: one this is independent of (t,y) and two it's concave (we can see that with a bit of work) so the maximum is achieved. The expression isn't tidy in the general case so let's just write

$$\hat{A}_{\mu,r,\sigma,\kappa,\lambda} := \max_a\Big[a(\mu-r)+r-\tfrac{1}{2}\sigma^2a^2+\lambda\ln(1-a^2\kappa^2)\Big]\,.$$

We can collect terms with and without y in the Bellman equation to see that

$$\psi'(t)=0\,,\;\; t\in[0,T]\,,\;\psi(T)=1\,,$$
$$\gamma'(t)+\psi(t)\hat{A}=0\,,\;\; t\in[0,T]\,,\;\gamma(T)=0\,.$$

So $\psi(t)=1$ for all $t\in[0,T]$ and $\gamma(t)=\hat{A}(T-t)$. Then $w(t,y)=y+\hat{A}(T-t)$ and so $v(t,x)=\ln x+\hat{A}(T-t)$.

6 Pontryagin maximum principle and backward stochastic differential equations

In the previous part, we developed the dynamic programming theory for the stochastic control problem with Markovian system.

We introduce another approach called Pontryagin optimality principle, originally due to Pontryagin in the deterministic case. We will also study this approach to study the control problem (P).

6.1 Non-rigorous Derivation of Pontryagin's Maximum Principle

Consider the control problem

$$
(P)\quad \begin{cases} \text{Maximize, over } \alpha \in \mathcal{A} \text{ the functional} \\ J(\alpha) := \mathbb{E}\left[\int_0^T f\big(s, X_s^{\alpha,0,x}, \alpha_s\big)\, ds + g\big(X_T^{\alpha,0,x}\big)\right], \\ \text{where } X_t^{\alpha,0,x} \text{ uniquely solves, for } t \in [0,T], \text{ the controlled SDE} \\ X_t = x + \int_0^t b(s, X_s, \alpha_s)\, ds + \int_0^t \sigma(s, X_s, \alpha_s)\, dW_s\,. \end{cases}
$$

Going back to what we know about calculus, if a maximum of a certain functional exists then it satisfies a "first order condition" which postulates that the derivative is zero "in every direction". Let us try to derive such first order condition.

The simplest interesting case: $f = 0$, $\sigma = 0$, $d = 1$ and $A \subseteq R$. So we just have a deterministic problem with 1-dimensional state and 1-dimensional control. Let $x \in \mathbb{R}$ and $\alpha \in \mathcal{A}$ be fixed. Let $\beta \in \mathcal{A}$. Then

$$
\frac{d}{d\varepsilon} J(\alpha + \varepsilon(\beta - \alpha))\Big|_{\varepsilon=0} = (\partial_x g)(X_T^\alpha) \frac{d}{d\varepsilon} X_T^{\alpha+\varepsilon(\beta-\alpha)}\Big|_{\varepsilon=0}.
$$

We see that to proceed we need to calculate how the controlled process changes with a change of control. Let us write $V_t := \frac{d}{d\varepsilon} X_t^{\alpha+\varepsilon(\beta-\alpha)}\big|_{\varepsilon=0}$. From the equation for the controlled process we see that

$$
V_t = \int_0^t (\partial_x b)(s, X_s^\alpha, \alpha_s) V_s\, ds + \int_0^t (\partial_a b)(s, X_s^\alpha, \alpha_s)(\beta_s - \alpha_s)\, ds\,.
$$

Note that the equation for V is affine and can be solved using an integrating factor[13] so that

$$
V_t = \int_0^t \exp\left(\int_s^t (\partial_x b)(r, X_r^\alpha, \alpha_r)\, dr\right) (\partial_a b)(s, X_s^\alpha, \alpha_s)(\beta_s - \alpha_s)\, ds\,.
$$

Hence

$$
\frac{d}{d\varepsilon} J(\alpha+\varepsilon(\beta-\alpha))\Big|_{\varepsilon=0} = (\partial_x g)(X_T^\alpha) \int_0^T \exp\left(\int_s^T (\partial_x b)(r, X_r^\alpha, \alpha_r)\, dr\right) (\partial_a b)(s, X_s^\alpha, \alpha_s)(\beta_s - \alpha_s)\, ds\,.
$$

[13] This works in the 1-dimensional case here. In higher dimension there still is an integrating factor but it no longer has the explicit form written above.

To simplify notation let us introduce the "backward process"

$$Y_t^\alpha := (\partial_x g)(X_T^\alpha) \exp\left(\int_t^T (\partial_x b)(r, X_r^\alpha, \alpha_r)\, dr\right)$$

and note that that the backward process satisfies an equation which we will refer to as the "backward equation":

$$dY_t = -(\partial_x b)(t, X_t^\alpha, \alpha_t) Y_t^\alpha\, dt\,, \;\; t \in [0,T]\,, \;\; Y_T^\alpha = (\partial_x g)(X_T^\alpha).$$

Furthermore let $H(t,x,y,a) := b(t,x,a)y$ (we will refer to H as the "Hamiltonian") so that $(\partial_a b)(t, X_t^\alpha, \alpha_t)Y_t^\alpha = (\partial_a H)(t, X_t^\alpha, Y_t^\alpha, \alpha_t)$. Then

$$\frac{d}{d\varepsilon} J(\alpha + \varepsilon(\beta - \alpha))\bigg|_{\varepsilon=0} = \int_0^T (\partial_a H)(s, X_s^\alpha, Y_s^\alpha, \alpha_s)(\beta_s - \alpha_s)\, ds\,. \tag{6.1}$$

If α is a (locally) optimal control then for any other control β we have, for any $\varepsilon > 0$, that

$$J(\alpha + \varepsilon(\beta - \alpha)) \leq J(\alpha)$$

and so

$$\begin{aligned}
0 &\geq \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\Big(J(\alpha + \varepsilon(\beta - \alpha)) - J(\alpha)\Big) \\
&= \frac{d}{d\varepsilon} J(\alpha + \varepsilon(\beta - \alpha))\bigg|_{\varepsilon=0} = \int_0^T (\partial_a H)(t, X_t^\alpha, Y_t^\alpha, \alpha_t)(\beta_t - \alpha_t)\, dt\,.
\end{aligned}$$

So

$$0 \geq \int_0^T (\partial_a H)(t, X_t^\alpha, Y_t^\alpha, \alpha_t)(\beta_t - \alpha_t)\, dt\,.$$

Finally from the definition of the derivative we get that for almost all $t \in [0,T]$ we have

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\Big(H(t, X_t^\alpha, Y_t^\alpha, \alpha_t + \varepsilon(\beta_t - \alpha_t)) - H(t, X_t^\alpha, Y_t^\alpha, \alpha_t)\Big) = (\partial_a H)(t, X_t^\alpha, Y_t^\alpha, \alpha_t)(\beta_t - \alpha_t)\,.$$

Hence there are $\varepsilon > 0$ (small) such that

$$H(t, X_t^\alpha, Y_t^\alpha, \alpha_t + \varepsilon(\beta_t - \alpha_t)) \leq H(t, X_t^\alpha, Y_t^\alpha, \alpha_t)\,.$$

From this we can conclude that any optimal control "locally maximizes the Hamiltonian" and any optimal control, together with the forward and backward processes must satisfy

$$\begin{cases}
\quad \alpha_t \in \arg\max_{a \in A} H(t, X_t^\alpha, Y_t^\alpha, a)\,, \\
dX_t^\alpha = b(t, X_t^\alpha, \alpha_t)\, dt\,, \;\; t \in [0,T]\,, \;\; X_0^\alpha = x\,, \\
dY_t^\alpha = -(\partial_x H)(t, X_t^\alpha, Y_t^\alpha, \alpha_t)\, dt\,, \;\; t \in [0,T]\,, \;\; Y_T^\alpha = (\partial_x g)(X_T^\alpha)\,.
\end{cases} \tag{6.2}$$

Stochastic case but with $f = 0$ for simplicity. Now $\sigma \neq 0$ and we have processes in higher dimensions. Then

$$\frac{d}{d\varepsilon} J(\alpha + \varepsilon(\beta - \alpha))\bigg|_{\varepsilon=0} = \mathbb{E}\left[(\partial_x g)(X_T^\alpha) \frac{d}{d\varepsilon} X_T^{\alpha + \varepsilon(\beta - \alpha)}\bigg|_{\varepsilon=0}\right]\,.$$

As before, to proceed, we need to calculate how the SDE changes with a change of control. Let us write $V_t := \frac{d}{d\varepsilon} X_s^{\alpha+\varepsilon(\beta-\alpha)}\big|_{\varepsilon=0}$. From the SDE we see that

$$
\begin{aligned}
V_t = &\int_0^t (\partial_x b)(s, X_s^\alpha, \alpha_s) V_s \, ds + \int_0^t (\partial_a b)(s, X_s^\alpha, \alpha_s)(\beta_s - \alpha_s) \, ds \\
&+ \int_0^t (\partial_x \sigma)(s, X_s^\alpha, \alpha_s) V_s \, dW_s + \int_0^t (\partial_x \sigma)(s, X_s^\alpha, \alpha_s)(\beta_s - \alpha_s) \, dW_s \, .
\end{aligned}
$$

Now we will work with a backward equation directly, setting $Y_T^\alpha = (\partial_x g)(X_T^\alpha)$ as before. It's not so clear what its dynamics should be so let it just be a general Itô process

$$
dY_t = U_t \, dt + Z_t \, dW_t \, , \;\; t \in [0,T] \, , \;\; Y_T^\alpha = (\partial_x g)(X_T^\alpha) \, .
$$

Recall that we are interested in $\mathbb{E}\big[Y_T V_T\big]$. We see that

$$
\begin{aligned}
\mathbb{E}\big[Y_T V_T\big] =& \mathbb{E}\left[\int_0^T Y_t \, dV_t + \int_0^T V_t \, dY_t + \int_0^T dY_t \, dV_t\right] \\
=& \mathbb{E} \int_0^T Y_t (\partial_x b)(t, X_t^\alpha, \alpha_t) V_t \, dt + \mathbb{E} \int_0^T Y_t (\partial_a b)(t, X_t^\alpha, \alpha_t)(\beta_t - \alpha_t) \, dt \\
&+ \mathbb{E} \int_0^T V_t U_t \, dt + \mathbb{E} \int_0^T Z_t (\partial_x \sigma)(t, X_t^\alpha, \alpha_t) V_t \, dt + \mathbb{E} \int_0^T Z_t (\partial_a \sigma)(t, X_t^\alpha, \alpha_t)(\beta_t - \alpha_t) \, dt \, .
\end{aligned}
\tag{6.3}
$$

We would like to derive something that looks like (6.1) i.e. keep $\beta - \alpha$ terms but avoid V in the dynamics. To that end we choose

$$
U_t = -(\partial_x b)(t, X_t^\alpha, \alpha_t) Y_t - (\partial_x \sigma)(t, X_t^\alpha, \alpha_t) Z_t \, .
$$

Then

$$
\begin{aligned}
&\frac{d}{d\varepsilon} J(\alpha + \varepsilon(\beta - \alpha))\bigg|_{\varepsilon=0} = \mathbb{E}\left[Y_T V_T\right] \\
&= \mathbb{E} \int_0^T (\partial_a b)(t, X_t^\alpha, \alpha_t) Y_t (\beta_t - \alpha_t) \, ds + \mathbb{E} \int_0^T (\partial_a \sigma)(t, X_t^\alpha, \alpha_t) Z_t (\beta_t - \alpha_t) \, dt \, .
\end{aligned}
$$

So if we let the Hamiltonian to be

$$
H(t, x, y, z, a) := b(t, x, a) y + \sigma(t, x, a) z
$$

then

$$
\frac{d}{d\varepsilon} J(\alpha + \varepsilon(\beta - \alpha))\bigg|_{\varepsilon=0} = \mathbb{E} \int_0^T (\partial_a H)(t, X_t^\alpha, Y_t^\alpha, Z_t^\alpha, \alpha_t)(\beta_t - \alpha_t) \, dt \, . \tag{6.4}
$$

Arguing as before, see the argument going from (6.1) and (6.2), we arrive at

$$
\begin{cases}
\quad \alpha_t \in \arg\max_{a \in A} H(t, X_t^\alpha, Y_t^\alpha, Z_t^\alpha, a) \, , \\
dX_t^\alpha = b(t, X_t^\alpha, \alpha_t) \, dt + \sigma(t, X_t^\alpha, \alpha_t) \, dW_t \, , \;\; t \in [0,T] \, , \;\; X_0^\alpha = x \, , \\
dY_t^\alpha = -(\partial_x H)(t, X_t^\alpha, Y_t^\alpha, \alpha_t) \, dt + Z_t^\alpha \, dW_t \, , \;\; t \in [0,T] \, , \;\; Y_T^\alpha = (\partial_x g)(X_T^\alpha) \, .
\end{cases}
\tag{6.5}
$$

You may ask why not take $Z = 0$, since then (6.3) would be simplified. We would have been able to only keep $\beta - \alpha$ terms as we wished. But without Z we would have no reason to hope that $(Y_t)_{t \in [0,T]}$ is $\mathcal{F}_t$-adapted since we are specifying a terminal condition. This in turn would mean that α is not $\mathcal{F}_t$-adapted which would render the optimality criteria useless.

6.2 Deriving a Numerical Method from Pontryiagin's maximum principle

Imagine we have chosen a control $\alpha \in \mathcal{A}$ and solved the corresponding forward and backward equations so that we have X^α and (Y^α, Z^α). How would we go about finding a control that's "better"?

Looking at (6.4) we would say that a better control would make the derivative positive (so that we change the control in a direction of the maximum): this means we will choose β such that

$$\beta_t - \alpha_t = \gamma(\partial_a H)(t, X_t^\alpha, Y_t^\alpha, Z^t\alpha, \alpha_t)$$

for some $\gamma > 0$ because then

$$\frac{d}{d\varepsilon} J(\alpha + \varepsilon(\beta - \alpha))\bigg|_{\varepsilon=0} = \gamma \mathbb{E} \int_0^T |(\partial_a H)(t, X_t^\alpha, Y_t^\alpha, Z_t^\alpha, \alpha_t)|^2 \, dt \geq 0\,.$$

Of course a full algorithm would need to solve the forward SDE (easy) and the backward SDE (not so easy) and carry out the update step repeatedly. In the deterministic case solving the forward and backward ODEs is easy and the method is known as "method of successive approximations" (MSA). It can be shown that a modification of this algorithm converges under appropriate conditions, see [12].

6.3 Backward Stochastic Differential Equations (BSDEs)

For a deterministic differential equation

$$\frac{dx(t)}{dt} = b(t, x(t)) \;\; t \in [0,T]\,, \;\; x(T) = a$$

we can reverse the time by changing variables. Let $\tau := T - t$ and $y(\tau) = x(t)$. Then we have

$$\frac{dy(\tau)}{d\tau} = -b(T-\tau, y(\tau)) \;\; \tau \in [0,T]\,, \;\; y(0) = a\,.$$

So the *backward* ODE is equivalent to a *forward* ODE.

The same argument would fail for SDEs since the time-reversed SDE would not be adapted to the appropriate filtration and the stochastic integrals will not be well defined.

Recall the martingale representation theorem (see Theorem A.24), which says any $\xi \in L^2_{\mathcal{F}_T}$ can be uniquely represented by

$$\xi = \mathbb{E}[\xi] + \int_0^T \phi_t \, dW_t\,.$$

If we define $M_t = \mathbb{E}[\xi] + \int_0^t \phi_s dW_s$, then M_t satisfies

$$dM_t = \phi_t \, dW_t\,, \;\; M_T = \xi\,.$$

This leads to the idea that a solution to a *backward* SDE must consist of two processes (in the case above M and ϕ).

Consider the backward SDE (BSDE)

$$dY_t = g_t\big(Y_t, Z_t\big)\, dt + Z_t \, dW_t, \qquad Y_T = \xi\,.$$

We shall give a few examples when this has explicit solution.

Example 6.1. Assume that $g = 0$. In this case, $Y_t = \mathbb{E}[\xi|\mathcal{F}_t]$ and Z is the process given by the martingale representation theorem.

Example 6.2. Assume that $g_t(y,z) = \gamma_t$. In this case, take $\hat{\xi} := \xi - \int_0^T \gamma_t\,dt$. We get the solution $(\hat{Y},\hat{Z})$ to

$$d\hat{Y}_t = \hat{Z}_t\,dW_t\,,\ \ \hat{Y}_T = \hat{\xi}$$

as

$$\hat{Y}_t = \mathbb{E}\left[\hat{\xi}|\mathcal{F}_t\right] = \mathbb{E}\left[\xi - \int_0^T \gamma_t\,dt\middle|\mathcal{F}_t\right]$$

and we get Z from the martingale representation theorem. Then with $Y_t := \hat{Y}_t + \int_0^t \gamma_s\,ds$, $Z_t := \hat{Z}_t$ we have a solution (Y,Z) so in particular

$$Y_t = \mathbb{E}\left[\xi - \int_0^T \gamma_t\,dt\middle|\mathcal{F}_t\right] + \int_0^t \gamma_s\,ds = \mathbb{E}\left[\xi - \int_t^T \gamma_s\,ds\middle|\mathcal{F}_t\right]\,.$$

Example 6.3. Assume that $g_t(y,z) = \alpha_t y + \beta_t z + \gamma_t$ and $\alpha = \alpha_t$, $\beta = \beta_t$, $\gamma = \gamma_t$ are real-valued adapted processes that satisfy certain integrability conditions (those will become clear). Assume that W is only real valued.[14] We will construct a solution using an exponential transform and a change of measure.

Consider a new measure $\mathbb{Q}$ given by the Radon–Nikodym derivative

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\frac{1}{2}\int_0^T \beta_s^2\,ds - \int_0^T \beta_s\,dW_s\right)$$

and assume that $\mathbb{E}\big[\frac{d\mathbb{Q}}{d\mathbb{P}}\big] = 1$. Then, due to Girsanov's Theorem A.23, the process given by $W_t^{\mathbb{Q}} = W_t + \int_0^t \beta_s\,ds$ is a $\mathbb{Q}$-Wiener process. Consider the BSDE

$$d\bar{Y}_t = \bar{\gamma}_t\,dt + \bar{Z}_t\,dW_t^{\mathbb{Q}}\,,\ \ \bar{Y}_T = \bar{\xi}\,, \tag{6.6}$$

where $\bar{\gamma}_t := \gamma_t \exp\left(-\int_0^t \alpha_s\,ds\right)$ and $\bar{\xi} := \xi\exp\left(-\int_0^T \alpha_s\,ds\right)$. We know from Example 6.2 that this BSDE has a solution $(\bar{Y},\bar{Z})$ and in fact we know that

$$\bar{Y}_t = \mathbb{E}^{\mathbb{Q}}\left[\xi e^{-\int_0^T \alpha_s\,ds} - \int_t^T \gamma_s e^{-\int_0^s \alpha_r\,dr}\,ds\middle|\mathcal{F}_t\right]\,.$$

We let $Y_t := \bar{Y}_t\exp\left(\int_0^t \alpha_s\,ds\right)$ and $Z_t := \bar{Z}_t\exp\left(\int_0^t \alpha_s\,ds\right)$. Now using the Itô product rule with (6.6) and the equation for $W^{\mathbb{Q}}$ we can check that

$$\begin{aligned} dY_t &= d\left(\bar{Y}_t e^{\int_0^t \alpha_s\,ds}\right) = \alpha_t Y_t\,dt + e^{\int_0^t \alpha_s\,ds}\,d\bar{Y}_t = \alpha_t Y_t\,dt + \gamma_t\,dt + Z_t\,dW_t^{\mathbb{Q}} \\ &= (\alpha_t Y_t + \beta_t Z_t + \gamma_t)\,dt + Z_t\,dW_t \end{aligned}$$

and moreover $Y_T = \xi$. In particular we get

$$Y_t = \mathbb{E}^{\mathbb{Q}}\left[\xi e^{-\int_t^T \alpha_s\,ds} - \int_t^T \gamma_s e^{-\int_t^s \alpha_r\,dr}\,ds\middle|\mathcal{F}_t\right]\,. \tag{6.7}$$

To get the solution as an expression in the original measure we need to use the Bayes formula for conditional expectation, see Proposition A.49. We obtain

$$Y_t = \frac{\mathbb{E}\left[\left(\xi e^{-\int_t^T \alpha_s\,ds} - \int_t^T \gamma_s e^{-\int_t^s \alpha_r\,dr}\,ds\right)e^{-\frac{1}{2}\int_0^T \beta_s^2\,ds - \int_0^T \beta_s\,dW_s}\middle|\mathcal{F}_t\right]}{\mathbb{E}\left[e^{-\frac{1}{2}\int_0^T \beta_s^2\,ds - \int_0^T \beta_s\,dW_s}\middle|\mathcal{F}_t\right]}\,.$$

[14] It is possible to find a solution in higher dimensions but the integrating factor doesn't have this nice explicit form and we cannot apply Girsanov's theorem in the same way. See [9, Section 3].

Proposition 6.4 (Boundedness of solutions to linear BSDEs)**.** *Consider the linear backward SDE with $g_t(y,z) = \alpha_t y + \beta_t z + \gamma_t$. If α, β, γ and ξ are all bounded then the process Y in the solution pair (Y,Z) is bounded.*

Proof. This proof is left as exercise. □

Example 6.5 (BSDE and replication in the Black-Scholes market)**.** In a standard Black-Scholes market model we have a risk-free asset $dB_t = rB_t\,dt$ and risky assets

$$dS_t = \mathrm{diag}(\mu)S_t\,dt + \sigma S_t\,dW_t\,.$$

Here μ is the drift (column) vector of the risky asset rate, σ is the volatility matrix.

Let π denote the cash amounts invested in the risky assets and Y the replicating portfolio value (so $Y - \sum_{i=1}^m \pi^{(i)}$ is invested in the risk-free asset). Then the self-financing property says that (interpreting $1/S$ to be $\mathrm{diag}(1/S_1,\ldots,1/S_m)$)

$$dY_t = \pi_t^\top \mathrm{diag}\left(\frac{1}{S_t}\right)dS_t + \frac{Y_t - \sum_{i=1}^m \pi_t^{(i)}}{B_t}\,dB_t$$

i.e.

$$dY_t = \left[rY_t + \pi_t^\top(\mu - r)\right]dt + \pi_t^\top \sigma\,dW_t\,.$$

We can define $Z_t = \pi_t^\top\sigma$ and if σ^{-1} exists then $\pi_t^\top = Z_t\sigma^{-1}$ and so

$$dY_t = \left[rY_t + Z_t\sigma^{-1}(\mu^\top - r)\right]dt + Z_t\,dW_t.$$

For any payoff ξ at time T, the replication problem is to solve the BSDE given by this differential coupled with $Y_T = \xi$. If $\xi \in L^2_{\mathcal{F}_T}$ the equation admits a unique square-integrable solution (Y,Z). Hence the cash amount invested in the risky assets, required in the replicating portfolio is $\pi_t^\top = Z_t\sigma^{-1}$, and the replication cost (contingent claim price) at time t is Y_t.

We see that this is a BSDE with linear driver and so from Example 6.3 we have, due to (6.7), that

$$Y_t = \mathbb{E}^{\mathbb{Q}}\left[\xi e^{-r(T-t)}\middle|\mathcal{F}_t\right]\,,$$

where

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{-\frac{1}{2}|\sigma^{-1}(\mu-r)|^2T - (\mu^\top - r)(\sigma^{-1})^\top W_T}\,.$$

In other words we see that $\mathbb{Q}$ is the usual risk-neutral measure we get in Black–Scholes pricing.

A standard backward SDE (BSDE) is formulated as

$$dY_t = g_t\big(Y_t, Z_t\big)\,dt + Z_t\,dW_t, \qquad Y(T) = \xi, \tag{6.8}$$

where $g = g_t(\omega, y, z)$ must be such that $g_t(y,z)$ is at least $\mathcal{F}_t$-measurable for any fixed t, y, z. We will refer to g is called as the *generator* or *driver* of the Backward SDE.

Definition 6.6. Given $\xi \in L^2(\mathcal{F}_T)$ and a generator g, a pair of $(\mathcal{F}_t)_{t\in[0,T]}$-adapted processes (Y,Z) is called as a solution for (6.8) if

$$Y_t = \xi - \int_t^T g_s\big(Y_s, Z_s\big)\,ds - \int_t^T Z_s\,dW_s, \quad \forall t \in [0,T].$$

Theorem 6.7 (Existence and uniqueness for BSDEs)**.** *Suppose $g = g_t(y,z)$ satisfies*

(i) We have $g(0,0) \in \mathcal{H}$.

(ii) There exists a constant $L > 0$ such that

$$|g_t(y,z) - g_t(\overline{y},\overline{z})| \le L(|y-\overline{y}| + |z-\overline{z}|)\ ,\ a.s.\ \forall t \in [0,T]\,, \forall y,z,\overline{y},\overline{z}\,.$$

Then for any $\xi \in L^2_{\mathcal{F}_T}$, there exists a unique $(Y,Z) \in \mathcal{H}\times\mathcal{H}$ solving the BSDE (6.8).

Recall that $\mathcal{H}$ is the space introduced in Definition A.18.

Proof. We consider the map $\Phi = \Phi(U,V)$ for (U,V) in $\mathcal{H}\times\mathcal{H}$. Given (U,V) we define $(Y,Z) = \Phi(U,V)$ as follows. Let $\hat{\xi} := \xi - \int_0^T g_s(U_s,V_s)\,ds$. Then

$$\begin{aligned}\mathbb{E}\int_0^T |g_s(U_s,V_s)|^2\,ds &\le \mathbb{E}\int_0^T [2|g_s(U_s,V_s) - g_s(0,0)|^2 + 2|g_s(0,0)|^2]\,ds \\ &\le \mathbb{E}\int_0^T [2L^2(|U_s|^2 + |V_s|^2) + 2|g_s(0,0)|^2]\,ds < \infty,\end{aligned} \tag{6.9}$$

since U and V and $g(0,0)$ are in $\mathcal{H}$. So $\hat{\xi} \in L^2(\mathcal{F}_T)$ and we know that for $\hat{Y}_t := \mathbb{E}[\hat{\xi}|\mathcal{F}_t]$ there is Z such that

$$d\hat{Y}_t = Z_t\,dW_t\,,\ \ \hat{Y}_T = \hat{\xi}\,.$$

Take $Y_t := \hat{Y}_t + \int_0^t g_s(U_s,V_s)\,ds$. Then

$$Y_t = \xi - \int_t^T g_s(U_s,V_s)\,ds - \int_t^T Z_s\,dW_s\,. \tag{6.10}$$

The next step is to show that $(U,V) \mapsto \Phi(U,V) = (Y,Z)$ described above is a contraction on an appropriate Banach space.

We will assume, for now, that $|\xi| \le N$ and that $|g| \le N$. We consider (U,V) and (U',V'). From these we obtain $(Y,Z) = \Phi(U,V)$ and $(Y',Z') = \Phi(U',V')$. We will write

$$(\bar{U},\bar{V}) := (U-U',V-V')\,,\ \ (\bar{Y},\bar{Z}) := (Y-Y',Z-Z')\,,\ \ \bar{g} := g(U,V) - g(U',V')\,.$$

Then

$$d\bar{Y}_s = \bar{g}_s\,ds + \bar{Z}_s dW_s$$

and with Itô formula we see that

$$d\bar{Y}_s^2 = 2\bar{Y}_s\bar{g}_s\,ds + 2\bar{Y}_s\bar{Z}_s\,dW_s + \bar{Z}_s^2 ds\,.$$

Hence, for some $\beta > 0$,

$$d(e^{\beta s}\bar{Y}_s^2) = e^{\beta s}\left[2\bar{Y}_s\bar{g}_s\,ds + 2\bar{Y}_s\bar{Z}_s\,dW_s + \bar{Z}_s^2\,ds + \beta\bar{Y}_s^2\,ds\right]\,.$$

Noting that, due to (6.10), we have $\bar{Y}_T = Y_T - Y'_T = 0$, we get

$$0 = \bar{Y}_0^2 + \int_0^T e^{\beta s}\left[2\bar{Y}_s\bar{g}_s + \bar{Z}_s^2 + \beta\bar{Y}_s^2\right]\,ds + \int_0^T 2e^{\beta s}\bar{Y}_s\bar{Z}_s\,dW_s\,.$$

Since $Z \in \mathcal{H}$ we have

$$\mathbb{E}\int_0^T 4e^{2\beta s}|\bar{Y}_s|^2|\bar{Z}_s|^2\,ds \le e^{2\beta T}4N^2(1+T)^2\mathbb{E}\int_0^T |\bar{Z}_s|^2\,ds < \infty$$

and so, the stochastic integral being a martingale, we get

$$\mathbb{E}\int_0^T e^{\beta s}\left[\bar{Z}_s^2 + \beta\bar{Y}_s^2\right]\,ds = -\mathbb{E}\bar{Y}_0^2 - \mathbb{E}\int_0^T e^{\beta s}2\bar{Y}_s\bar{g}_s\,ds \le 2\mathbb{E}\int_0^T e^{\beta s}|\bar{Y}_s||\bar{g}_s|\,ds\,.$$

Using the Lipschitz continuity of g and Young's inequality (with $\varepsilon = 1/4$) we have

$$\begin{aligned} e^{\beta s}|\bar{Y}_s||\bar{g}_s| &\le e^{\beta s}|\bar{Y}_s|L(|\bar{U}_s| + |\bar{V}_s|) \le 2L^2e^{\beta s}|\bar{Y}_s|^2 + \frac{1}{8}e^{\beta s}(|\bar{U}_s| + |\bar{V}_s|)^2 \\ &\le 2L^2e^{\beta s}|\bar{Y}_s|^2 + \frac{1}{4}e^{\beta s}(|\bar{U}_s|^2 + |\bar{V}_s|^2)\,. \end{aligned}$$

We can now take $\beta = 1 + 4L^2$ and we obtain

$$\mathbb{E}\int_0^T e^{\beta s}\left[\bar{Z}_s^2 + \bar{Y}_s^2\right]\,ds \le \frac{1}{2}\mathbb{E}\int_0^T e^{\beta s}(|\bar{U}_s|^2 + |\bar{V}_s|^2)\,ds\,. \tag{6.11}$$

We now need to remove the assumption that $|\xi| \le N$ and $|g| \le N$. To that end consider $\xi^N := -N \wedge \xi \vee N$ and $g^N := -N \wedge g \vee N$ (so $|\xi^N| \le N$ and $|g^N| \le N$). We obtain Y^N, Z^N as before. Note that

$$Y_t = \mathbb{E}[\xi|\mathcal{F}_t] = \mathbb{E}\left[\lim_{N\to\infty}\hat{\xi}^N|\mathcal{F}_t\right] = \lim_{N\to\infty}Y_t^N$$

due to Lebesgue's dominated convergence for conditional expectations. Indeed, we have $|\hat{\xi}^N| \le |\xi| + \int_0^T |g_s(U_s,V_s)|\,ds$ and this is in L^2 due to (6.9). Moreover

$$\begin{aligned} \mathbb{E}\int_0^T |Z_t^N - Z_t|^2\,dt &= \mathbb{E}\left(\int_0^T (Z_t^N - Z_t)\,dW_t\right)^2 = \mathbb{E}\left(Y_T^N - Y_T + Y_0 - Y_0^N\right)^2 \\ &\le 2\mathbb{E}|Y_T^N - Y_T|^2 + 2\mathbb{E}|Y_0 - Y_0^N|^2 \to 0 \text{ as } N\to\infty \end{aligned}$$

due to Lebesgue's dominated convergence theorem. Then from (6.11) be have, for each N,

$$\mathbb{E}\int_0^T e^{\beta s}\left[\bar{|}Z_s^N|^2 + \bar{|}Y_s^N|^2\right]\,ds \le \frac{1}{2}\mathbb{E}\int_0^T e^{\beta s}(|\bar{U}_s|^2 + |\bar{V}_s|^2)\,ds\,.$$

But since the RHS is independent of N, we obtain (6.11) but now without the assumption that $|\xi| \le N$ and $|g| \le N$. Consider now the Banach space $(\mathcal{H}\times\mathcal{H}, \|\cdot\|)$, with

$$\|(Y,Z)\| := \mathbb{E}\int_0^T e^{\beta s}\left[Z_s^2 + Y_s^2\right]\,ds\,.$$

From (6.11) we have

$$\|\Phi(U,V) - \Phi(U',V')\| \le \frac{1}{2}\|(U,V) - (U',V')\|\,.$$

So the map $\Phi : \mathcal{H}\times\mathcal{H} \to \mathcal{H}\times\mathcal{H}$ is a contraction and due to Banach's Fixed Point Theorem there is a unique (Y^*,Z^*) which solves the equation $\Phi(Y^*,Z^*) = (Y^*,Z^*)$. Hence

$$Y_t^* = \xi - \int_t^T g_s(Y_s^*,Z_s^*)\,ds - \int_t^T Z_s^*\,dW_s$$

due to (6.10). $\square$

Theorem 6.8. *Let (Y^1, Z^1) and (Y^2, Z^2) be solutions to BSDEs with generators and terminal conditions g^1, ξ^1 and g^2, ξ^2 respectively. Assume that $\xi^1 \leq \xi^2$ a.s. and that $g_t^2(Y_t^2, Z_t^2) \leq g_t^1(Y_t^1, Z_t^1)$ a.e. on $\Omega \times (0,T)$. Assume finally that the generators satisfy the assumption of Theorem 6.7 and $\xi^1, \xi^2 \in L^2(\mathcal{F}_T)$. Then $Y^1 \leq Y^2$.*

Proof. We note that the BSDE satisfied by $\bar{Y} := Y^2 - Y^1$, $\bar{Z} := Z^2 - Z^1$ is

$$d\bar{Y}_t = [g_t^2(Y_t^2, Z_t^2) - g_t^1(Y_t^1, Z_t^1)]\,dt + \bar{Z}_t\,dW_t\,,\ \ \bar{Y}_T = \bar{\xi} := \xi^2 - \xi^1\,.$$

This is

$$\begin{aligned} d\bar{Y}_t =& [g_t^2(Y_t^2, Z_t^2) - g_t^2(Y_t^1, Z_t^2) + g_t^2(Y_t^1, Z_t^2) - g_t^2(Y_t^1, Z_t^1) + g_t^2(Y_t^1, Z_t^1) - g_t^1(Y_t^1, Z_t^1)]\,dt \\ &+ \bar{Z}_t\,dW_t\,,\ \ \bar{Y}_T = \bar{\xi} \end{aligned}$$

which we can re-write as

$$d\bar{Y}_t = [\alpha_t \bar{Y}_t + \beta_t \bar{Z}_t + \gamma_t]\,dt + \bar{Z}_t\,dW_t\,,\ \ \bar{Y}_T = \bar{\xi}\,,$$

where

$$\alpha_t := \frac{g_t^2(Y_t^2, Z_t^2) - g_t^2(Y_t^1, Z_t^2)}{Y_t^2 - Y_t^1}\mathbb{1}_{Y_t^1 \neq Y_t^2}\,,\ \ \beta_t := \frac{g_t^2(Y_t^1, Z_t^2) - g_t^2(Y_t^1, Z_t^1)}{Z_t^2 - Z_t^1}\mathbb{1}_{Z_t^1 \neq Z_t^2}$$

and where

$$\gamma_t := g_t^2(Y_t^1, Z_t^1) - g_t^1(Y_t^1, Z_t^1)\,.$$

Due to the Lipschitz assumption on g^2 we get that α and β are bounded and since Y^i, Z^i are in $\mathcal{H}$ we get that $\gamma \in \mathcal{H}$. Thus we have an affine BSDE for $(\bar{Y}, \bar{Z})$ and the conclusion follows from (6.7) since we get

$$\bar{Y}_t = \mathbb{E}^{\mathbb{Q}}\left[\underbrace{\bar{\xi} e^{-\int_t^T \alpha_s\,ds}}_{\geq 0} - \underbrace{\int_t^T \gamma_s e^{-\int_t^s \alpha_r\,dr}\,ds}_{\leq 0}\middle|\mathcal{F}_t\right] \geq 0$$

from the assumptions that $\xi^1 \leq \xi^2$ a.s. and that $g_t^2(Y_t^2, Z_t^2) \leq g_t^1(Y_t^1, Z_t^1)$ a.e. $\square$

6.4 Pontryagin's Maximum Principle as Sufficient Condition

We now return to the optimal control problem (P). Recall that given running gain f and terminal gain g our aim is to optimally control

$$dX_t^\alpha = b_t(X_t, \alpha_t)\,dt + \sigma_t(X_t, \alpha_t)\,dW_t,\ \ t \in [0,T]\,,\ \ X_0^\alpha = x,$$

where $\alpha \in \mathcal{A}$ and we assume that Assumption 3.9 holds. Recall that by optimally controlling the process we mean a control which will maximize

$$J(\alpha) := \mathbb{E}\left[\int_0^T f(t, X_t^\alpha, \alpha_t)\,dt + g(X_T^\alpha)\right]$$

over $\alpha \in \mathcal{A}$. Unlike in Chapter 4 we can consider the process starting from time 0 (because we won't be exploiting the Markov property of the SDE) and unlike in Chapter 4 we will assume that A is a subset of $\mathbb{R}^m$.

We define the *Hamiltonian* $H : [0,T] \times \mathbb{R}^d \times A \times \mathbb{R}^d \times \mathbb{R}^{d \times d'} \to \mathbb{R}$ of the system as

$$H_t(x, a, y, z) := b_t(x, a)\,y + \mathrm{tr}[\sigma_t^\top(x, a)\,z] + f_t(x, a)\,.$$

Assumption 6.9. Assume that $x \mapsto H_t(x,a,y,z)$ is differentiable for all a,t,y,z with derivative bounded uniformly in a,t,y,z. Assume that g is differentiable in x with the derivative having at most linear growth (in x).

Consider the *adjoint BSDEs* (one for each $\alpha \in \mathcal{A}$)

$$dY_t^\alpha = -\partial_x H_t(t, X_t, \alpha_t, Y_t^\alpha, Z_t^\alpha)\, dt + Z_t\, dW_t\,, \;\; Y_T^\alpha = \partial_x g(X_T^\alpha)\,.$$

Note that under Assumption 6.9 and 3.9

$$\mathbb{E}[|\partial_x g(X_T^\alpha)|^2] \le \mathbb{E}[(K(1+|X_T^\alpha|)^2] < \infty,$$

Hence, due to Theorem 6.7, the adjoint BSDEs have unique solutions (Y^α, Z^α).

We will now see that it is possible to formulate a sufficient optimality criteria based on the properties of the Hamiltonian and based on the adjoint BSDEs. This is what is known as the *Pontryagin's Maximum Principle.* Consider two control processes, $\alpha, \beta \in \mathcal{A}$ and the two associated controlled diffusions, both starting from the same initial value, labelled X^α, X^β. Then

$$J(\beta) - J(\alpha) = \mathbb{E}\left[\int_0^T \left[f(t, X_t^\beta, \beta_t) - f(t, X_t^\alpha, \alpha_t)\right]\, dt + g(X_T^\beta) - g(X_T^\alpha)\right]\,.$$

We will need to assume that g is concave (equivalently assume $-g$ is convex). Then $g(x) - g(y) \ge \partial_x g(x)(x-y)$ and so (recalling what the terminal condition in our adjoint equation is)

$$\mathbb{E}\left[g(X_T^\beta) - g(X_T^\alpha)\right] \ge \mathbb{E}\left[(X_T^\beta - X_T^\alpha)\partial_x g(X_T^\beta)\right] = \mathbb{E}\left[(X_T^\beta - X_T^\alpha) Y_T^\beta\right]\,.$$

We use Itô's product rule and the fact that $X_0^\alpha = X_0^\beta$. Let us write $\Delta b_t := b_t(X_t^\beta, \beta_t) - b_t(X_t^\alpha, \alpha_t)$ and $\Delta\sigma_t := \sigma_t(X_t^\beta, \beta_t) - \sigma_t(X_t^\alpha, \alpha_t)$. Then we see that

$$\begin{aligned}\mathbb{E}\left[(X_T^\beta - X_T^\alpha) Y_T^\beta\right] &\ge \mathbb{E}\Bigg[\int_0^T -(X_t^\beta - X_t^\alpha)\partial_x H_t(X_t^\beta, \beta_t, Y_t^\beta, Z_t^\beta)\, dt \\ &+ \int_0^T \Delta b_t\, Y_t^\beta\, dt + \int_0^T \operatorname{tr}\left[\Delta\sigma_t^\top Z_t^\beta\right]\, dt\Bigg]\,.\end{aligned}$$

Note that we are missing some details here, because the second stochastic integral term that we dropped isn't necessarily a martingale. However with a stopping time argument and Fatou's Lemma the details can be filled in (and this is why we have an inequality). We also have that for all y, z,

$$\begin{aligned} f(t, X_t^\beta, \beta_t) &= H_t(X_t^\beta, \beta_t, y, z) - b_t(X_t^\beta, \beta_t) y - \operatorname{tr}[\sigma_t^\top(X_t^\beta, \beta_t) z]\,, \\ f(t, X_t^\alpha, \alpha_t) &= H_t(X_t^\alpha, \alpha_t, y, z) - b_t(X_t^\alpha, \alpha_t) y - \operatorname{tr}[\sigma_t^\top(X_t^\alpha, \alpha_t) z] \end{aligned}$$

and so

$$f(t, X_t^\beta, \beta_t) - f(t, X_t^\alpha, \alpha_t) = \Delta H_t - \Delta b_t Y_t^\beta - \operatorname{tr}(\Delta\sigma_t^\top Z_t^\beta)$$

where

$$\Delta H_t := H_t(X_t^\beta, \beta_t, Y_t^\beta, Z_t^\beta) - H_t(X_t^\alpha, \alpha_t, Y_t^\beta, Z_t^\beta)\,.$$

Thus

$$\mathbb{E}\left[\int_0^T \left[f(t, X_t^\beta, \beta_t) - f(t, X_t^\alpha, \alpha_t)\right]\, dt\right] = \mathbb{E}\left[\int_0^T \left[\Delta H_t - \Delta b_t Y_t^\beta - \operatorname{tr}(\Delta\sigma_t^\top Z_t^\beta)\right]\, dt\right]\,.$$

Altogether

$$J(\beta) - J(\alpha) \geq \mathbb{E}\left[\int_0^T \left[\Delta H_t - (X_t^\beta - X_t^\alpha)\partial_x H_t(X_t^\beta, \beta_t, Y_t^\beta, Z_t^\beta)\right]\, dt\right]$$

If we now assume that $(x,a) \mapsto H_t(x, a, Y_t^\beta, Z_t^\beta)$ is differentiable and concave for any t, y, z then

$$\Delta H_t \geq (X_t^\beta - X_t^\alpha)\partial_x H_t(X_t^\beta, \beta_t, Y_t^\beta, Z_t^\beta) + (\beta_t - \alpha_t)\partial_a H_t(X_t^\beta, \beta_t, Y_t^\beta, Z_t^\beta)$$

and so

$$J(\beta) - J(\alpha) \geq \mathbb{E}\left[\int_0^T (\beta_t - \alpha_t)\partial_a H_t(X_t^\beta, \beta_t, Y_t^\beta, Z_t^\beta)\, dt\right] .$$

Finally we assume that β_t is a control process which satisfies

$$H_t(X_t^\beta, \beta_t, Y_t^\beta, Z_t^\beta) = \max_{a \in A} H_t(X_t^\beta, a, Y_t^\beta, Z_t^\beta) < \infty \text{ a.s. for almost all } t.$$

Then $J(\beta) \geq J(\alpha)$ for arbitrary α. In other words, such control β is optimal. Hence we have proved the following theorem.

Theorem 6.10 (Pontryagin's Maximum Principle). *Let Assumptions 3.9 and 6.9 holds, let $\subset \mathbb{R}^m$. Let g be concave. Let $\beta \in \mathcal{A}$ and let X^β be the associated controlled diffusion and (Y^β, Z^β) the solution of the adjoint BDSE. If $\beta \in \mathcal{A}$ is such that*

$$H_t(X_t^\beta, \beta_t, Y_t^\beta, Z_t^\beta) = \max_{a \in A} H_t(X_t^\beta, a, Y_t^\beta, Z_t^\beta) < \infty \;\; \textit{a.s. for almost all } t. \tag{6.12}$$

holds and if

$$(x,a) \mapsto H_t(x, a, Y_t^\beta, Z_t^\beta)$$

is differentiable and concave then $J(\beta) = \sup_\alpha J(\alpha)$ i.e. β is an optimal control.

We can see that the Pontryagin maximum principle gives us a sufficient condition for optimality.

Example 6.11 (Linear–quadratic control revisited). Take W to be $\mathbb{R}^{d'}$-valued Wiener process and let the space where controls take values to be $A = \mathbb{R}^m$. Consider $X_t = X_t^{\alpha,x}$ taking values in $\mathbb{R}^d$ given by

$$dX_t = [L(t)X_t + M(t)\alpha_t]\, dt + \sigma(t)\, dW_t \text{ for } t \in [0,T] ,\;\; X_0 = x ,$$

where $L = L(t) \in \mathbb{R}^{d\times d}, M = M(t) \in \mathbb{R}^{d\times m}$ and $\sigma = \sigma(t) \in \mathbb{R}^{d\times d'}$ are bounded, measurable, deterministic functions of t.

Further[15] let $C = C(t) \in \mathbb{R}^{d\times d}$, $D = D(t) \in \mathbb{R}^{m\times m}$, $F = F(t) \in \mathbb{R}^{d\times m}$ be deterministic, integrable functions of t and $R \in \mathbb{R}^{d\times d}$ be such that C, D and R are symmetric, $C = C(t) \leq 0, R \leq 0$ and $D = D(t) \leq -\delta < 0$ with some constant $\delta > 0$. The aim will be to maximize

$$J^\alpha(x) := \mathbb{E}^{x,\alpha}\left[\int_0^T \left[X_t^\top C(t)X_t + \alpha_t^\top D(t)\alpha_t + 2X_t^\top F(t)\alpha_t\right]\, dt + X_T^\top R X_T\right]$$

[15] For any matrix $\mathcal{M}$ we will write $\mathcal{M} \leq 0$ to denote negative definite matrices i.e. matrices such that for any $\xi \in \mathbb{R}^n$ we have $\xi^\top \mathcal{M}\xi \leq 0$. Similarly $\mathcal{M} < 0$ denotes strictly negative definite matrices i.e. those such that for any $\xi \in \mathbb{R}^n$ we have $\xi^\top \mathcal{M}\xi < 0$.

over all adapted processes α such that $\mathbb{E}\int_0^T \alpha_t^2\,dt < \infty$ (we will call these admissible). The Hamiltonian is

$$H_t(x,a,y,z) = x^\top L(t)\,y + y^\top M(t)\,a + \operatorname{tr}\left[\sigma(t)^\top z\right] + x^\top C(t)\,x + a^\top D(t)\,a + 2x^\top F(t)\,a\,.$$

We see that as function of (a,x) it is a sum of linear and quadratic functions and hence differentiable. Moreover since $C \le 0$ and $D < 0$ we see that it is concave.

We see that

$$\partial_x H_t(x,a,y,z) = L(t)y + 2C(t)x + 2F(t)a$$

and so the adjoint BSDE $(\hat{Y},\hat{Z})$ for the optimal control $\hat{\alpha}$ is

$$d\hat{Y}_t = -\left[L(t)\hat{Y}_t + 2C(t)\hat{X}_t + 2F(t)\hat{\alpha}\right]\,dt + \hat{Z}_t\,dW_t \ \text{ for } \ t\in[0,T]\,,\ \ \hat{Y}_T = 2R\hat{X}_T\,.$$

Note that $x \mapsto x^\top R x$ is concave (since $R \le 0$) and so the Pontryagin's maximum principle applies. If $\hat{\alpha}$ is the optimal control, $\hat{X}$ is the associated diffusion and $(\hat{Y},\hat{Z})$ is the solution to the adjoint BSDE for $\hat{\alpha}$ then the maximum principle says that

$$H_t(\hat{X}_t,\hat{\alpha}_t,\hat{Y}_t,\hat{Z}_t) = \max_{a\in\mathbb{R}} H_t(\hat{X}_t,a,\hat{Y}_t,\hat{Z}_t)\,.$$

In this case the maximum is achieved when (Hamiltonian is quadratic in a with negative leading coefficient so we just differentiate w.r.t. a and see for which value this is 0):

$$0 = M(t)^\top \hat{Y}_t + 2D(t)\,a + 2F(t)^\top \hat{X}_t$$

i.e.

$$\hat{\alpha}_t = -\frac{1}{2}D(t)^{-1}\left(M(t)^\top \hat{Y}_t + 2F(t)^\top \hat{X}_t\right)\,.$$

Inspecting the terminal condition for the adjoint BSDE leads us to "guess" that we should have $\hat{Y}_t = 2S(t)\hat{X}_t$ for some $S \in C^1([0,T];\mathbb{R}^{d\times d})$, S symmetric and $S \le 0$ with $S(T) = R$. We rewrite the optimal control with our guess for $\hat{Y}$:

$$\hat{\alpha}_t = -D(t)^{-1}\left(M(t)^\top S(t) + F(t)^\top\right)\hat{X}_t$$

and we can also write the optimally controlled SDE:

$$d\hat{X}_t = \left\{L(t) + M(t)\left[-D(t)^{-1}\left(M(t)^\top S(t) + F(t)^\top\right)\right]\right\}\hat{X}_t\,dt + \sigma(t)\,dW_t\,. \tag{6.13}$$

Since our guess is that $\hat{Y}_t = 2S(t)\hat{X}_t$ we have, due to Itô's formula

$$\begin{aligned} d\hat{Y}_t &= 2S'(t)\hat{X}_t\,dt + 2S(t)\,d\hat{X}_t \\ &= 2S'(t)\hat{X}_t\,dt + 2S(t)\left\{L(t) + M(t)\left[-D(t)^{-1}\left(M(t)^\top S(t) + F(t)^\top\right)\right]\right\}\hat{X}_t\,dt \\ &\quad + 2S(t)\sigma(t)\,dW_t\,. \end{aligned}$$

On the other hand the adjoint equation for $\hat{Y}$ gives

$$d\hat{Y}_t = -2\left[L(t)S(t) + C(t) - F(t)D(t)^{-1}\left(M(t)^\top S(t) + F(t)^\top\right)\right]\hat{X}_t\,dt + \hat{Z}_t\,dW_t\,.$$

Since both must hold we get that $\hat{Z}_t = 2S(t)\sigma(t)$ and that

$$S'(t) + S(t)L(t) + S(t)M(t)\left[-D(t)^{-1}\left(M(t)^\top S(t) + F(t)^\top\right)\right]$$
$$= -L(t)S(t) - C(t) + F(t)D(t)^{-1}\left(M(t)^\top S(t) + F(t)^\top\right)$$

so that for $t \in [0, T]$

$$S'(t) = [S(t)M(t) + F(t)]D(t)^{-1}\left(M(t)^\top S(t) + F(t)^\top\right) - L(t)S(t) - S(t)L(t) - C(t)$$

with $S(T) = R$. Under our assumptions the Ricatti equation has a unique solution such that S is symmetric and $S \le 0$.

The equation (6.13) for $\hat{X}$ is linear and clearly has unique solution and all the moments are bounded.

We observe (recalling $\hat{Y}_t = 2S(t)\hat{X}_t$) that

$$2\hat{X}_T^\top R\hat{X}_T = 2\hat{X}_T^\top S(T)\hat{X}_T = \hat{X}_T^\top \hat{Y}_T = \hat{X}_T^\top \hat{Y}_T - \hat{X}_0^\top \hat{Y}_0 + \hat{X}_0^\top \hat{Y}_0 = \int_0^T d(\hat{X}_t^\top \hat{Y}_t) + 2x^\top S(0)x. \tag{6.14}$$

Let us write $\psi(t) := -D(t)^{-1}\left(M(t)^\top S(t) + F(t)^\top\right)$, so that

$$d\hat{X}_t = [L(t) + M(t)\psi(t)]\,\hat{X}_t\,dt + \sigma(t)\,dW_t\,,$$
$$d\hat{Y}_t = -2\,[L(t)S(t) + C(t) + F(t)\psi(t)]\,\hat{X}_t\,dt + 2S(t)\sigma\,dW_t\,.$$

Moreover[16]

$$\begin{aligned}\frac{1}{2}d(\hat{X}_t^\top \hat{Y}_t) =& \frac{1}{2}\left(\hat{X}_t^\top S(t)d\hat{X}_t + \hat{X}_t^\top d\hat{Y}_t + d(\hat{X}_t^\top)d\hat{Y}_t\right)\\ =& \hat{X}_t^\top S(t)L(t)\hat{X}_t\,dt + \hat{X}_t^\top S(t)M(t)\psi(t)\hat{X}_t\,dt + \hat{X}_t^\top S(t)\sigma(t)\,dW_t\\ &- \hat{X}_t^\top L(t)S(t)\hat{X}_t\,dt - \hat{X}_t^\top C(t)\hat{X}_t - \hat{X}_t^\top F(t)\psi(t)\hat{X}_t\\ &+ \hat{X}_t^\top S(t)\sigma(t)\,dW_t + \mathrm{tr}[\sigma(t)(S(t)\sigma(t))^\top]\,dt\,.\end{aligned}$$

Hence,

$$\begin{aligned}\frac{1}{2}d(\hat{X}_t^\top \hat{Y}_t) =& + \hat{X}_t^\top S(t)M(t)\psi(t)\hat{X}_t\,dt\\ &- \hat{X}_t^\top C(t)\hat{X}_t - \hat{X}_t^\top F(t)\psi(t)\hat{X}_t\\ &+ 2\hat{X}_t^\top S(t)\sigma(t)\,dW_t + \mathrm{tr}[\sigma(t)(S(t)\sigma(t))^\top]\,dt\,.\end{aligned} \tag{6.15}$$

We also have

$$J^{\hat{\alpha}}(x) = \mathbb{E}\left[\int_0^T \left(\hat{X}_t^\top C(t)\hat{X}_t + \hat{\alpha}_t^\top D(t)\hat{\alpha}_t + 2X_t^\top F(t)\hat{\alpha}_t\right)\,dt + R\hat{X}_T^2\right]. \tag{6.16}$$

Noting that $\alpha_t^\top D(t)\hat{\alpha}_t = -\hat{X}_t^\top (M(t)^\top S(t) + F(t)^\top)^\top \psi(t)\hat{X}_t$ and substituting (6.15) into (6.14) and using this in (6.16) we see that most terms cancel and hence

$$J^{\hat{\alpha}}(x) = \mathbb{E}\left[\int_0^T \mathrm{tr}\left[\sigma(t)(S(t)\sigma(t))^\top\right]\,dt + \int_0^T 2\hat{X}_t^\top S(t)\sigma(t)\,dW_t + x^\top S(0)\,x\right].$$

[16] We have S symmetric. Then

$$x^\top SLx = (S^\top x)^\top Lx = (Sx)^\top (Lx) = (Lx)^\top Sx = x^\top LSx$$

where the first and last equalities are $AB = (B^\top A^\top)^\top$, the second equality is symmetry of S and the third equality is properties of dot products.

Since the solution of the SDE for $\hat{X}$ has all moments bounded we have

$$\mathbb{E}\int_0^T 4|S(t)|^2|\sigma(t)|^4|\hat{X}_t|^2\,dt \leq N\int_0^T \mathbb{E}|\hat{X}_t|^2\,dt \leq N_T < \infty\,.$$

The stochastic integral is thus a martingale and so

$$v(x) = J^{\hat{\alpha}}(x) = x^\top S(0)x + \int_0^T \operatorname{tr}\left[\sigma(t)(S(t)\sigma(t))^\top\right]\,dt\,.$$

Example 6.12 (Minimum variance for given expected return)**.** We consider the simplest possible model for optimal investment: we have a risk-free asset B with evolution given by $dB_t = rB_t\,dt$ and $B_0 = 1$ and a risky asset S with evolution given by $dS_t = \mu S_t\,dt + \sigma S_t\,dW_t$ with S_0 given. For simplicity we assume that σ, μ, r are given constants, $\sigma \neq 0$ and $\mu > r$. The value of a portfolio with no asset injections / consumption is given by $X_0 = x$ and

$$dX_t^\alpha = \frac{\alpha_t}{S_t}\,dS_t + \frac{X_t - \alpha_t}{B_t}\,dB_t\,,$$

where α_t represents the amount invested in the risky asset. Then

$$dX_t^\alpha = (rX_t + \alpha_t(\mu - r))\,dt + \sigma\alpha_t\,dW_t\,. \tag{6.17}$$

Given a desired return $m > 0$ we aim to find a trading strategy which would minimize the variance of the return (in other words a strategy that gets as close to the desired return as possible). We restrict ourselves to α such that $\mathbb{E}\int_0^T \alpha_t^2\,dt < \infty$. Thus we seek

$$V(m) := \inf_\alpha \left\{\operatorname{Var}(X_T^\alpha) : \mathbb{E}X_T^\alpha = m\right\}\,. \tag{6.18}$$

See Exercise 6.13 to convince yourself that the set over which we wish to take infimum is non-empty. Conveniently, if, for $\lambda \in \mathbb{R}$, we can calculate

$$v(\lambda) := \inf_\alpha \mathbb{E}\left[|X_T^\alpha - \lambda|^2\right]$$

then [15, Proposition 6.6.5] tells us that

$$V(m) = \sup_{\lambda\in\mathbb{R}}\left[v(\lambda) - (m-\lambda)^2\right]\,.$$

Furthermore

$$v(\lambda) = -\sup_\alpha \mathbb{E}\left[-|X_T^\alpha - \lambda|^2\right]\,.$$

Thus our aim is to maximize

$$J_\lambda(\alpha) := \mathbb{E}\left[g(X_T^\alpha)\right] \;\text{ with }\; g(x) = -(x-\lambda)^2\,.$$

Since g is concave and differentiable we will try to apply Pontryagin's maximum principle. As there is no running gain (i.e. $f = 0$) and since X^α is given by (6.17) we have the Hamiltonian

$$H_t(x,a,y,z) = [rx + a(\mu - r)]y + \sigma a\,z\,.$$

This, being affine in (a,x), is certainly differentiable and concave. Moreover, if there is an optimal control β and if the solution of the adjoint BSDE is denoted (Y^β, Z^β) then

$$\max_a H_t(X_t^\beta, a, Y_t^\beta, Z_t^\beta) = \max_a \left[rX_t^\beta Y_t^\beta + a(\mu - r)Y_t^\beta + \sigma a Z_t^\beta\right]\,.$$

The quantity being maximized is linear in a and thus it will be finite if and only if the solution to the adjoint equation satisfies

$$(\mu - r)Y_t^\beta + \sigma Z_t^\beta = 0 \quad \text{a.s. for a.a } t. \tag{6.19}$$

From now on we omit the superscript β everywhere. Recalling the adjoint equation:

$$dY_t = -rY_t\,dt + Z_t\,dW_t \ \text{ and } Y_T = \partial_x g(X_T) = -2(X_T - \lambda). \tag{6.20}$$

To proceed we will need to make a guess at what the solution to the adjoint BSDE will look like. Since the terminal condition is linear in X_T we will try the ansatz $Y_t = \varphi(t)X_t + \psi(t)$ for some C^1 functions φ and ψ. Notice that this is rather different to the situation in Example 6.3, since there we obtain a solution but only in terms of an unknown process arising from the martingale representation theorem. With this ansatz we have, substituting the expression for Y on the r.h.s. of (6.20), that

$$dY_t = -r\varphi(t)X_t\,dt - r\psi(t)\,dt + Z_t\,dW_t \tag{6.21}$$

and on the other hand we can use the ansatz for Y and product rule on the l.h.s. of (6.20) to see

$$\begin{aligned} dY_t &= \varphi(t)\,dX_t + X_t\varphi'(t)\,dt + \psi'(t)\,dt \\ &= \varphi(t)\,[rX_t + \beta_t(\mu - r)]\,dt + \varphi(t)\sigma\beta_t\,dW_t + X_t\varphi'(t)\,dt + \psi'(t)\,dt\,. \end{aligned} \tag{6.22}$$

The second equality above came from (6.17) with β as the control. Then (6.21) and (6.22) can simultaneously hold only if $Z_t = \varphi(t)\sigma\beta_t$ and if

$$\varphi(t)\,[rX_t + \beta_t(\mu - r)] + X_t\varphi'(t) + \psi'(t) = -r\varphi(t)X_t - r\psi(t)\,.$$

This in turn will hold as long as

$$\beta_t = \frac{2r\varphi(t)X_t + r\psi(t) + \varphi'(t)X_t + \psi'(t)}{\varphi(t)(r - \mu)}\,. \tag{6.23}$$

On the other hand from the Pontryagin maximum principle we conculded (6.19) which, with $Y_t = \varphi(t)X_t + \psi(t)$ and $Z_t = \varphi(t)\sigma\beta_t$ says

$$(\mu - r)[\varphi(t)X_t + \psi(t)] + \sigma^2\varphi(t)\beta_t = 0,$$

i.e.

$$\beta_t = \frac{(r - \mu)[\varphi(t)X_t + \psi(t)]}{\sigma^2\varphi(t)}\,. \tag{6.24}$$

But (6.23) and (6.24) can both hold only if (collecting terms with X_t and without)

$$\begin{aligned} \varphi'(t) &= \left(\tfrac{(r-\mu)^2}{\sigma^2} - 2r\right)\varphi(t)\,, \quad \varphi(T) = -2 \\ \psi'(t) &= \left(\tfrac{(r-\mu)^2}{\sigma^2} - r\right)\psi(t)\,, \quad \psi(T) = 2\lambda\,. \end{aligned} \tag{6.25}$$

Note that the terminal conditions arose from Y_T (rather than from the equations for β). Also note that ψ clearly depends on λ but for now we omit this in our notation. Clearly

$$\varphi(t) = -2e^{-\left(\frac{(r-\mu)^2}{\sigma^2} - 2r\right)(T-t)} \quad\text{and}\quad \psi(t) = 2\lambda e^{-\left(\frac{(r-\mu)^2}{\sigma^2} - r\right)(T-t)}\,. \tag{6.26}$$

We note that from (6.24) we can write the control as Markov control

$$\beta(t,x) = -\frac{(\mu - r)[\varphi(t)x + \psi(t)]}{\sigma^2\varphi(t)}\,.$$

Thus X driven by this control is square integrable. Indeed β is a linear function in x and together with (6.17) and Proposition 3.6 we can conclude the square integrability. Thus we also have $\mathbb{E}\int_0^T \beta_t^2\,dt < \infty$ and so the control is admissable.

We still need to know

$$v(\lambda) = -J(\beta) = \mathbb{E}\left[|X_T - \lambda|^2\right]\,.$$

We cannot calculate this by solving for X as in Exercise 6.13 (try it). Instead we note that

$$\mathbb{E}|X_T - \lambda|^2 = \mathbb{E}\left[-\frac{1}{2}\varphi(T)X_T^2 - \psi(T)X_T + \lambda^2\right].$$

From Itô's formula for $\xi_t := -\frac{1}{2}\varphi(t)X_t^2 - \psi(t)X_t$ we get that

$$-d\xi_t = \left(\tfrac{1}{2}\varphi'(t)X_t^2 + \psi'(t)X_t\right)\,dt + [\varphi(t)X_t + \psi(t)]\,dX_t + \tfrac{1}{2}\varphi(t)\,dX(t)dX(t)\,.$$

And we have that

$$dX_t = (rX_t + \beta_t(\mu - r))\,dt + \sigma\beta_t\,dW_t\,.$$

Hence

$$\begin{aligned}-\mathbb{E}\xi_T = -\xi_0 + \mathbb{E}\int_0^T \Big(&\tfrac{1}{2}\varphi'(t)X_t^2 + \psi'(t)X_t\\ &+ r\varphi(t)X_t^2 + r\psi(t)X_t\\ &+ \beta_t(\mu - r)[\varphi(t)X_t + \psi(t)]\\ &+ \tfrac{1}{2}\varphi(t)\sigma^2\beta_t^2\Big)\,dt\,.\end{aligned}$$

From the optimality condition $(\mu - r)\beta_t[\varphi(t)X_t + \psi_t] + \sigma^2\varphi(t)\beta_t^2 = 0$ we get

$$\tfrac{1}{2}\sigma^2\varphi(t)\beta_t^2 = -\tfrac{1}{2}(\mu - r)\beta_t[\varphi(t)X_t + \psi_t]$$

and so

$$\begin{aligned}-\mathbb{E}\xi_T = -\xi_0 + \mathbb{E}\int_0^T \Big(&\tfrac{1}{2}\varphi'(t)X_t^2 + \psi'(t)X_t\\ &+ r\varphi(t)X_t^2 + r\psi(t)X_t\\ &+ \tfrac{1}{2}\beta_t(\mu - r)[\varphi(t)X_t + \psi(t)]\Big)\,dt\,.\end{aligned}$$

This is

$$-\mathbb{E}\xi_T = -\xi_0 + \tfrac{(r-\mu)^2}{\sigma^2}\mathbb{E}\int_0^T \left(\tfrac{1}{2}\varphi(t)X_t^2 + \psi(t)X_t - \tfrac{1}{2}\tfrac{\varphi(t)^2X_t^2 + 2\varphi(t)\psi(t)X_t + \psi(t)^2}{\varphi(t)}\right)\,dt\,.$$

So

$$\mathbb{E}\xi_T = \xi_0 + \tfrac{1}{2}\tfrac{(r-\mu)^2}{\sigma^2}\int_0^T \tfrac{\psi(t)^2}{\varphi(t)}\,dt\,.$$

Due to (6.26) we have

$$\mathbb{E}\xi_T = \xi_0 - \lambda^2 \frac{(r-\mu)^2}{\sigma^2} \int_0^T e^{-\frac{(r-\mu)^2}{\sigma^2}(T-t)}\, dt\,.$$

Hence

$$\mathbb{E}\xi_T = \xi_0 - \lambda^2 \left[1 - e^{-\frac{(r-\mu)^2}{\sigma^2} T} \right].$$

But $\mathbb{E}|X_T - \lambda|^2 = \mathbb{E}\xi_T + \lambda^2$ and so

$$\mathbb{E}|X_T - \lambda|^2 = \xi_0 + \lambda^2 e^{-\frac{(r-\mu)^2}{\sigma^2} T}.$$

Moreover $\xi_0 = -\frac{1}{2}\varphi(0)x^2 - \psi(x)x$ and so

$$\xi_0 = x^2 e^{-\left(\frac{(r-\mu)^2}{\sigma^2} - 2r\right)T} - 2x\lambda e^{-\left(\frac{(r-\mu)^2}{\sigma^2} - r\right)T}.$$

Finally

$$\mathbb{E}|X_T - \lambda|^2 = e^{-\frac{(r-\mu)^2}{\sigma^2} T} \left[x^2 e^{2rT} - 2x\lambda e^{rT} + \lambda^2 \right] = e^{-\frac{(r-\mu)^2}{\sigma^2} T} \left(\lambda - x e^{rT} \right)^2.$$

which means that

$$v(\lambda) = -\kappa \left(\lambda - x e^{rT} \right)^2,$$

where $\kappa := e^{-\frac{(r-\mu)^2}{\sigma^2} T} > 0$. We thus get

$$V(m) = \sup_{\lambda \in \mathbb{R}} \left[-\kappa \left(\lambda^2 - 2\lambda x e^{rT} + x^2 e^{2rT} \right) - \lambda^2 + 2\lambda m - m^2 \right].$$

This is achieved when

$$0 = -\kappa\lambda + \kappa x e^{rT} - \lambda + m$$

i.e. when $\lambda = \frac{\kappa x e^{rT} + m}{\kappa + 1}$.

6.5 Variational connection to HJB equation

In this section we want to formally derive the HJB equation as a necessary condition for optimality using a variational argument. We consider the control problem

$$(P) \quad \begin{cases} \text{Maximize, over } a : [0,T] \times \mathbb{R}^d \to \mathbb{R}^m \text{ the functional} \\ J(a) := \mathbb{E}\left[\int_0^T f\big(s, X_s^{a,0,\xi}, a_s(X_s^{a,0,\xi})\big)\, ds + g\big(X_T^{a,0,\xi}\big) \right], \\ \text{where } X_t^{a,0,\xi},\ t \in [0,T], \text{ is the weakly unique solution of} \\ X_t = \xi + \int_0^t b(s, X_s, a_s(X_s))\, ds + \int_0^t \sigma(s, X_s, a_s(X_s))\, dW_s\,. \end{cases}$$

Let us assume that b and σ are smooth and uniformly bounded and moreover $\sigma\sigma^\top \geq \lambda I$ for some $\lambda > 0$. Assume $\xi \sim p_0$. The controlled Fokker–Planck–Kolmogorov equation (FPKE), see Section A.2.2, for the density of $X_t^{a,0,\xi}$, $t \in [0,T]$ is

$$\partial_t p - \tfrac{1}{2}\nabla^2 \Big((\sigma\sigma^\top)(\cdot, a) p \Big) + \nabla \cdot (b(\cdot, a)p) = 0 \text{ in } (0,T) \times \mathbb{R}^d, \tag{6.27}$$

together with $p(0,\cdot) = p_0$ on $\mathbb{R}^d$. Note that $p = p^a$ in that it depends on the Markov control chosen.

We will use the convention that $\nabla = \nabla_x$ and integrals over domain of integration left unspecified are over $\mathbb{R}^d$.

We will now mimic the argument from Section 6.1 but instead of perturbing controlled processes we'll perturb the Markov controls and instead of working with the SDE we will work with the controlled FPKE. Let us fix Markov controls $(t,x) \mapsto a(t,x)$ and $(t,x) \mapsto a'(t,x)$. Then

$$J(a) = \int_0^T \int f_t(\cdot, a_t) p_t^a \, dx \, dt + \int g \, p_T^a \, dx$$

and hence (using product rule)

$$\begin{aligned} &\frac{d}{d\varepsilon} J(a + \varepsilon(a' - a))\Big|_{\varepsilon=0} \\ &= \int_0^T \int \left[f_s(\cdot, a_s) v_s + p_s^a (a'_s - a_s) \partial_a f_s(\cdot, a_s) \right] dx \, dt + \int g \, v_T \, dx \, , \end{aligned}$$

where $v_t := \frac{d}{d\varepsilon} p_t^{a+\varepsilon(a'-a)}$ and where a, a' are two fixed Markov control functions. To proceed we'd like an equation for v and to that end we need to take derivative in (6.27). Note that

$$\frac{d}{d\varepsilon} \left[b(\cdot, a_t + \varepsilon(a'_t - a_t)) p_t^{a_t + \varepsilon(a'_t - a_t)} \right] \Big|_{\varepsilon=0} = b(\cdot, a_t) v_t + p_t^a \nabla_a b(\cdot, a_t)(a'_t - a_t)$$

and

$$\frac{d}{d\varepsilon} \left[(\sigma\sigma^\top)(\cdot, a_t + \varepsilon(a'_t - a_t)) p_t^{a_t + \varepsilon(a'_t - a_t)} \right] \Big|_{\varepsilon=0} = (\sigma\sigma^\top)(\cdot, a_t) v_t + p_t^a \left(\nabla_a (\sigma\sigma^\top) \right) (\cdot, a_t)(a'_t - a_t) \, .$$

Hence from these and (6.27) we get

$$\begin{aligned} 0 = \partial_t v_t - \tfrac{1}{2} \nabla^2 &\left[(\sigma\sigma^\top)(\cdot, a_t) v_t + p_t^a \left(\nabla_a (\sigma\sigma^\top) \right) (\cdot, a_t)(a'_t - a_t) \right] \\ &+ \nabla \cdot \left[b(\cdot, a_t) v_t + p_t^a \nabla_a b(\cdot, a_t)(a'_t - a_t) \right] \text{ in } (0,T) \times \mathbb{R}^d \, . \end{aligned}$$

Next, we would like to find a "backward" equation for an unknown function, say $(t,x) \mapsto u_t(x)$, such that $u_T = g$ and in terms of the backward equation we have

$$\frac{d}{d\varepsilon} J(a + \varepsilon(a' - a))\Big|_{\varepsilon=0} = \int_0^T \int \ldots \text{Hamiltonian involving } u \ldots (a' - a) \, dx \, dt \, .$$

At this point we don't know exactly how the Hamiltonian should look and we don't know what the backward equation should be exactly. Nevertheless, since $v_0 = 0$, we get from the product rule that

$$\begin{aligned} &\frac{d}{d\varepsilon} J(a + \varepsilon(a' - a))\Big|_{\varepsilon=0} \\ &= \int_0^T \int \left[f_s(\cdot, a_s) v_s + p_s^a (a'_s - a_s) \partial_a f_s(\cdot, a_s) \right] dx \, ds + \int u_T v_T \, dx \\ &= \int_0^T \int \left[f_t(\cdot, a_t) v_t + p_t^a (a'_t - a_t) \partial_a f_t(\cdot, a_t) \right] dx \, dt + \int_0^T \int u_t \partial_t v_t \, dx \, dt + \int_0^T \int v_t \partial_t u_t \, dx \, dt \, . \end{aligned}$$

We now want to choose the evolution of u i.e. choose $\partial_t u$ such that all the terms *not involving* $(a'_t - a_t)$ cancel out. To that end we note that

$$
\begin{aligned}
\int u_t \partial_t v_t \, dx \, dt = & \int u_t \tfrac{1}{2}\nabla^2 \left[(\sigma\sigma^\top)(\cdot, a_t) v_t + p_t^a \left(\nabla_a (\sigma\sigma^\top) \right)(\cdot, a_t)(a'_t - a_t) \right] \, dx \\
& - \int u_t \nabla \cdot \left[b(\cdot, a_t) v_t + p_t^a \nabla_a b(\cdot, a_t)(a'_t - a_t) \right] \, dx \, .
\end{aligned}
$$

Since we are in the business of deciding what u should satisfy let us demand that $u_t, \nabla u_t \to 0$ as $x \to \infty$ "sufficiently fast" (we could quantify this if we quantify how p_t^a behaves at infinity). Then, using integration by parts

$$
\begin{aligned}
\int u_t \partial_t v_t \, dx \, dt = & \int \tfrac{1}{2} \left[\mathrm{tr}(\sigma\sigma^\top \nabla^2 u_t)(\cdot, a_t) v_t + p_t^a \left(\nabla_a \mathrm{tr}((\sigma\sigma^\top)(\cdot, a_t) \nabla^2 u_t) \right)(a'_t - a_t) \right] \, dx \\
& + \int \nabla u_t \cdot \left[b(\cdot, a_t) v_t + p_t^a \nabla_a b(\cdot, a_t)(a'_t - a_t) \right] \, dx \, .
\end{aligned}
$$

Hence if we take u such that

$$
\partial_t u = -\tfrac{1}{2}\mathrm{tr}(\sigma\sigma^\top \nabla^2 u_t)(\cdot, a_t) - b(\cdot, a_t) \cdot \nabla u_t - f_t(\cdot, a_t)
$$

then

$$
\begin{aligned}
& \frac{d}{d\varepsilon} J(a + \varepsilon(a' - a)) \Big|_{\varepsilon=0} \\
& = \int_0^T \int \left[\nabla_a f_t(\cdot, a_t) + \nabla_a b(\cdot, a_t) \nabla u_t + \tfrac{1}{2} \left(\nabla_a \mathrm{tr}((\sigma\sigma^\top)(\cdot, a_t) \nabla^2 u_t) \right)(a'_t - a_t) p_t^a \right] dx \, ds \\
& = \int_0^T \int \nabla_a H(\cdot, \nabla u_t, \nabla^2 u_t, a_t)(a'_t - a_t) p_t^a \, dx \, dt \, ,
\end{aligned}
$$

where we defined, for $x \in \mathbb{R}^d$, $y \in \mathbb{R}^d$, $z \in \mathbb{R}^{d \times d}$ and $a \in \mathbb{R}^m$ the Hamiltonian as

$$
H(x, y, z, a) := \tfrac{1}{2}\mathrm{tr}((\sigma\sigma^\top)(x, a) z) + b(x, a) \cdot y + f(x, a) \, . \tag{6.28}
$$

Now we observe that if a is a (locally) optimal control then for any small $\varepsilon > 0$ we have

$$
\begin{aligned}
0 \geq \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(J(a + \varepsilon(a' - a)) - J(a) \right) & = \frac{d}{d\varepsilon} J(a + \varepsilon(a' - a)) \Big|_{\varepsilon=0} \\
= \int_0^T \int \nabla_a H(\cdot, \nabla u_t, \nabla^2 u_t, a_t)(a'_t - a_t) p_t^a \, dx \, dt \, .
\end{aligned}
$$

Since we are assuming $\sigma\sigma^\top \geq \lambda I$ for $\lambda > 0$ we get that p_t^a has full support and so for Lebesgue-almost-every t, x we have $\varepsilon > 0$ such that

$$
0 \geq H(x, \nabla u_t(x), \nabla^2 u_t(x), a_t(x) + \varepsilon(a' - a_t(x))) - H(x, \nabla u_t(x), \nabla^2 u_t(x), a_t(x)) \, .
$$

Hence the optimal control and forward and backward equations satisfy

$$
\begin{aligned}
a_t(x) \in \operatorname*{arg\,max}_{a' \in \mathbb{R}^m} & H(x, \nabla u_t(x), \nabla^2 u_t(x), a') \;\; (t, x) \in (0, T) \times \mathbb{R}^d, \\
0 &= \partial_t p^a - \tfrac{1}{2}\nabla^2 \left((\sigma\sigma^\top)(\cdot, a) p^a \right) + \nabla \cdot (b(\cdot, a) p^a) \;\text{ in }\; (0, T) \times \mathbb{R}^d \, , \\
0 &= \partial_t u + \tfrac{1}{2}\mathrm{tr}(\sigma\sigma^\top \nabla^2 u_t)(\cdot, a_t) + b(\cdot, a_t) \cdot \nabla u_t - f_t(\cdot, a_t) \;\text{ in }\; (0, T) \times \mathbb{R}^d \\
u_T &= g \, .
\end{aligned}
$$

We note that this implies that the backward equation is

$$\partial_t u + \max_{a' \in \mathbb{R}^m} \Big[\tfrac{1}{2}\mathrm{tr}(\sigma\sigma^\top \nabla^2 u)(\cdot, a') + b(\cdot, a') \cdot \nabla u - f(\cdot, a') \Big] = 0 \;\text{ in }\; (0,T) \times \mathbb{R}^d$$

with the terminal condition $u_T = g$. This is exactly the HJB equation we've derived from the Bellman principle in (4.4).

What if we wanted to see if the HJB will arise from concavity-type consideration? Let us consider two Markov controls a' and a. Then

$$J(a') - J(a) = \int_0^T \int f_t(\cdot, a'_t) p'_t \, dx \, dt + \int g \, p'_T \, dx - \int_0^T \int f_t(\cdot, a_t) p_t \, dx \, dt - \int g \, p_T \, dx \,, \tag{6.29}$$

where for brevity $p^{a'} = p'$ and $p^a = p$ are the densities of the controlled SDEs controlled with Markov controls a' and a' respectively.

Let u be the solution to the linear equation

$$\partial_t u = -\tfrac{1}{2}\mathrm{tr}(\sigma\sigma^\top(\cdot, a)\nabla^2 u_t) - b(\cdot, a) \cdot \nabla u - f(\cdot, a) \;\text{ in }\; (0,T) \times \mathbb{R}^d \,,$$

with $u_T = g$. This is the equation for the value function arising from using the Markov control a. If we want to, we can write an equation for value function arising from a' and label it u'. This is the "backward" equation. The forward equations are given by (6.27) and we know that $p'_0 = p_0$ and so we can write

$$\langle g, p'_T - p_T \rangle = \langle u_T, p'_T - p_T \rangle - \langle u_0, p'_0 - p_0 \rangle = \int_0^T d\langle u_t, p'_t - p_t \rangle \,,$$

where we adopted the notation $\langle u, v \rangle = \int uv \, dx$ whenever the integral can be well defined. Then, using some kind of product rule and plugging in the "forward" and "backward" equations we should have, that

$$\begin{aligned}
&\int_0^T d\langle u_t, p'_t - p_t \rangle = \int_0^T \langle u_t, d(p'_t - p_t) \rangle + \int_0^T \langle p'_t - p_t, du_t \rangle \\
&= \int_0^T \Big\langle u_t, \Big[\tfrac{1}{2}\nabla^2 \Big((\sigma\sigma^\top)(\cdot, a'_t) p'_t \Big) - \nabla \cdot (b(\cdot, a'_t) p'_t) - \tfrac{1}{2}\nabla^2 \Big((\sigma\sigma^\top)(\cdot, a_t) p_t \Big) + \nabla \cdot (b(\cdot, a_t) p_t) \Big] \Big\rangle \, dt \\
&\quad + \int_0^T \Big\langle p'_t - p_t, \Big[-\tfrac{1}{2}\mathrm{tr}(\sigma\sigma^\top(\cdot, a)\nabla^2 u_t) - b(\cdot, a) \cdot \nabla u - f(\cdot, a) \Big] \Big\rangle \, dt \,.
\end{aligned}$$

If we can justify that integration-by-parts leads to no boundary terms then

$$\Big\langle u_t, \tfrac{1}{2}\nabla^2 \Big((\sigma\sigma^\top)(\cdot, a'_t) p'_t \Big) \Big\rangle = \Big\langle \tfrac{1}{2}\mathrm{tr}((\sigma\sigma^\top)(\cdot, a'_t)\nabla^2 u_t), p'_t \Big\rangle$$

and

$$-\Big\langle u_t, \nabla \cdot (b(\cdot, a'_t) p'_t) \Big\rangle = \Big\langle \nabla u_t \cdot b(\cdot, a'_t), p'_t \Big\rangle$$

and similarly for the terms without the prime and so

$$\begin{aligned}
\langle g, p'_T - p_T \rangle = &\int_0^T \Big\langle \tfrac{1}{2}\mathrm{tr}(\sigma\sigma^\top(\cdot, a'_t)\nabla^2 u_t), p'_t \Big\rangle + \Big\langle \nabla u_t \cdot b(\cdot, a'_t), p'_t \Big\rangle \, dt \\
&- \int_0^T \Big\langle p'_t, \Big[\tfrac{1}{2}\mathrm{tr}(\sigma\sigma^\top(\cdot, a_t)\nabla^2 u_t) + b(\cdot, a_t) \cdot \nabla u_t + f(\cdot, a_t) \Big] \Big\rangle \, dt \\
&+ \int_0^T \Big\langle p_t, f(\cdot, a_t) \Big\rangle \, dt \,.
\end{aligned}$$

From this and (6.29) we have that

$$
\begin{aligned}
J(a') - J(a) &= \int_0^T \langle f_t(\cdot, a'_t), p'_t\rangle - \langle f_t(\cdot, a_t), p_t\rangle \, dt + \langle g, p'_T - p_T\rangle \, dt \\
&= \int_0^T \langle f_t(\cdot, a'_t), p'_t\rangle + \left\langle \tfrac{1}{2}(\sigma\sigma^\top)(\cdot, a'_t)\nabla^2 u_t, p'_t \right\rangle + \left\langle \nabla u_t \cdot b(\cdot, a'_t), p'_t \right\rangle \\
&\qquad - \left\langle p'_t, \left[\tfrac{1}{2}\mathrm{tr}(\sigma\sigma^\top(\cdot, a_t)\nabla^2 u_t) + b(\cdot, a_t)\cdot \nabla u + f(\cdot, a_t)\right] \right\rangle dt \\
&= \int_0^T \Big\langle p'_t, \tfrac{1}{2}\mathrm{tr}((\sigma\sigma^\top)(\cdot, a'_t)\nabla^2 u_t) + \nabla u_t \cdot b(\cdot, a'_t) + f_t(\cdot, a'_t) \\
&\qquad\qquad - \tfrac{1}{2}\mathrm{tr}(\sigma\sigma^\top(\cdot, a_t)\nabla^2 u_t) - b(\cdot, a_t)\cdot\nabla u - f(\cdot, a_t) \Big\rangle \, dt \, .
\end{aligned}
$$

Using the definition of the Hamiltonian (6.28) we get

$$
J(a') - J(a) = \int_0^T \left\langle p'_t, H(\cdot, \nabla u_t, \nabla^2 u_t, a'_t) - H(\cdot, \nabla u_t, \nabla^2 u_t, a_t) \right\rangle \, dt \, .
$$

Assuming the Hamiltonian is concave in a for every x, y, z we have

$$
H(x, y, z, a') - H(x, y, z, a) \le (\nabla_a H)(x, y, z, a)(a' - a)
$$

and so

$$
J(a') - J(a) \le \int_0^T \left\langle p'_t, (\nabla_a H)(\cdot, \nabla u_t, \nabla^2 u_t, a_t)\cdot (a' - a) \right\rangle \, dt \, . \tag{6.30}
$$

Now assume that

$$
a_t(x) \in \mathrm{argmax}_{b\in\mathbb{R}^p} H(x, \nabla u_t, \nabla^2 u_t, b)
$$

then $(\nabla_a H)(\cdot, \nabla u_t, \nabla^2 u_t, a_t)\cdot(a' - a) = 0$ and so $J(a') - J(a) \le 0$.

6.6 Exercises

Exercise 6.13 (To complement Example 6.12)**.** Show that, under the assumptions of Example 6.12, the set $\{\mathrm{Var}(X_T^\alpha) : \mathbb{E}X_T^\alpha = m\}$ is nonempty.

Exercise 6.14 (Merton's problem with exponential utility, no consumption, using Pontryagin's Maximum Principle)**.** Consider a model with a risky asset $(S_t)_{t\in[0,T]}$ and a risk-free asset $(B_t)_{t\in[0,T]}$ given by

$$
\begin{aligned}
dS_t &= \mu S_t \, dt + \sigma S_t \, dW_t \;\; t \in [0, T] \, , S_0 = S \, , \\
dB_t &= r B_t \, dt \;\; t \in [0, T] \, , S_0 = S \, , B_0 = 1 \, ,
\end{aligned}
$$

where $\mu, r \in \mathbb{R}$ and $\sigma > 0$ are given constants. Let $(X_t)_{t\in[0,T]}$ denote the value of a self-financing investment portfolio with $X_0 = x > 0$ and let α_t denote the fraction of the portfolio value X_t invested in the risky asset. We note that X_t depends on the investment strategy α_t and so we write $X_t = X_t^\alpha$. We will only consider α that are real-valued, adapted and such that $\mathbb{E}\int_0^T \alpha_t^2 \, dt < \infty$, denoting such strategies $\mathcal{A}$ and calling them admissable.

Our aim is to find the investment strategy $\hat\alpha$ which maximizes, over $\alpha \in \mathcal{A}$,

$$
J(\alpha) = \mathbb{E}\left[-\exp(-\gamma X_T^\alpha)\right] \, ,
$$

for some $\gamma > 0$.

i) Use the definition of a self-financing portfolio to derive the equation for the portfolio value:

$$dX_t = X_t \left[\alpha_t(\mu - r) + r\right] dt + X_t\alpha_t\sigma\, dW_t\,.$$

ii) Write down the Hamiltonian for the problem and the adjoint BSDE for the optimal portfolio (use $\hat\alpha$ to denote the optimal control, $(\hat Y, \hat Z)$ to denote the BSDE).

iii) Explain how Pontryagin's maximum principle implies that

$$\hat Z_t = -\frac{\mu - r}{\sigma}\hat Y_t\,.$$

iv) Noting that $\hat Y_T = \gamma e^{-\gamma \hat X_T}$ use the "ansatz" $\hat Y_t = \phi_t e^{-\psi_t \hat X_t}$ with some $\phi, \psi \in C^1([0,T])$ such that $\phi_T = \gamma$ and $\psi_T = \gamma$. Hence show that

$$\hat X_t \hat\alpha_t = e^{-r(T-t)}\frac{\mu - r}{\gamma\sigma^2}\,.$$

6.7 Solutions to the exercises

Solution (to Exercise 6.13)**.** We start by solving (6.17) for some $\alpha_t = a$ constant. Note that (with $X = X^\alpha$)

$$d(e^{-rt}X_t) = e^{-rt}\left[dX_t - rX_t\, dt\right] = e^{-rt}\left[a(\mu - r)\, dt + \sigma a\, dW_t\right]\,.$$

Thus

$$e^{-rT}X_T = x + \int_0^T e^{-rt}a(\mu - r)\, dt + \int_0^T \sigma a e^{-rt}\, dW_t\,.$$

Since the stochastic integral is a true martingale

$$\mathbb{E}X_T = e^{rT}x + e^{rT}a(\mu - r)\int_0^T e^{-rt}\, dt = e^{rT}x + a(\mu - r)\frac{1}{r}\left(e^{rT} - 1\right)\,.$$

Thus with

$$a = r\frac{m - e^{rT}x}{(\mu - r)(e^{rT} - 1)}$$

we see that $\mathbb{E}X_T = m$ and so the set is non-empty.

Solution (to Exercise 6.14)**.** i) We have

$$dX_t = \frac{\alpha_t X_t}{S_t}\, dS_t + \frac{X_t - \alpha_t X_t}{B_t}\, dB_t = \alpha_t X_t \mu\, dt + \alpha_t X_t \sigma\, dW_t + X_t r\, dt - \alpha_t X_t r\, dt$$

so

$$dX_t = X_t\left[\alpha_t(\mu - r) + r\right] dt + X_t\alpha_t\sigma\, dW_t\,.$$

ii) Let us write down the Hamiltonian:

$$H_t(x, a, y, z) = x[a(\mu - r) + r]y + x\, a\, \sigma\, z$$

so

$$\partial_x H_t(x, a, y, z) = [a(\mu - r) + r]y + a\, \sigma\, z\,.$$

The adjoint BSDE for the optimal portfolio $\hat X$, which we denote $(\hat Y, \hat Z)$ then is

$$d\hat Y_t = -[\hat\alpha_t(\mu - r) + r]\hat Y_t\, dt - \hat\alpha_t\sigma\hat Z_t\, dt + \hat Z_t\, dW_t \quad t \in [0,T]\,, \hat Y_T = \gamma\exp(-\gamma\hat X_T)\,. \qquad (6.31)$$

We can show that $\hat X_t > 0$ since $x > 0$. Hence $|\hat Y_T|^2 = \gamma^2\exp(-2\gamma\hat X_T) \le \gamma^2$ and so $\hat Y_T \in L^2(\mathcal{F}_T)$. The above affine BSDE thus has a unique solution $(\hat Y, \hat Z)$ and we may proceed.

iii) We note that the terminal reward function $g(x) = -e^{-\gamma x}$ is concave. We can check that the Hamiltonian is concave in x as well as in a but not in (x, a), so the optimality principle as a sufficient condition doesn't apply.

Nevertheless, according to the optimality principle the optimal control $\hat\alpha$ must satisfy

$$H_t(\hat X_t, \hat\alpha_t, \hat Y_t, \hat Z_t) = \max_{a\in\mathbb{R}}\left[\hat X_t\big(a(\mu - r) - r\big)\hat Y_t + \hat X_t a\sigma\hat Z_t\right]\,.$$

We need the Hamiltonian to be finite which in turns means that it must hold that

$$\hat{X}_t \hat{Y}_t(\mu - r) + \hat{X}_t \hat{Z}_t \sigma = 0\,.$$

Hence

$$\hat{Z}_t = -\frac{\mu - r}{\sigma} \hat{Y}_t\,. \tag{6.32}$$

iv) We will use the "ansatz" $\hat{Y}_t = \phi_t e^{-\psi_t \hat{X}_t}$ with some $\phi, \psi \in C^1([0,T])$ such that $\phi_T = \gamma$ and $\psi_T = \gamma$. We note that

$$d(-\psi_t \hat{X}_t) = -\psi_t \hat{X}_t[\hat{\alpha}_t(\mu - r) + r]\,dt - \psi_t \hat{X}_t \hat{\alpha}_t \sigma\,dW_t - \psi_t' \hat{X}_t\,dt$$

so that

$$\begin{aligned}
d\hat{Y}_t &= \phi_t d(e^{-\psi_t \hat{X}_t}) + e^{-\psi_t \hat{X}_t} d\phi_t \\
&= e^{-\psi_t \hat{X}_t}\Big[\phi_t d(-\psi_t \hat{X}_t) + \frac{1}{2}\phi_t d(-\psi_t \hat{X}_t) d(-\psi_t \hat{X}_t) + d\phi_t\Big] \\
&= e^{-\psi_t \hat{X}_t}\Big[- \psi_t \phi_t d\hat{X}_t - \hat{X}_t \phi_t d\psi_t + \frac{1}{2}\phi_t \psi_t^2 d\hat{X}_t d\hat{X}_t + d\phi_t\Big] \\
&= e^{-\psi_t \hat{X}_t}\Big[- \psi_t \phi_t \hat{X}_t\big[\,(\hat{\alpha}_t(\mu - r) + r)\ dt + \hat{\alpha}_t \sigma\,dW_t\big] - \hat{X}_t \phi_t \psi_t'\,dt + \frac{1}{2}\phi_t \psi_t^2 \hat{\alpha}_t^2 \hat{X}_t^2 \sigma^2\,dt + \phi_t'\,dt\Big]\,.
\end{aligned}$$

If we now go to the adjoint BSDE (6.31) and substitute for $\hat{Z}_t$ from (6.32) we see that we must also have

$$d\hat{Y}_t = -r\hat{Y}_t dt - \frac{\mu - r}{\sigma}\hat{Y}_t\,dW_t\,.$$

Equating the "dW terms" leads to

$$\frac{\mu - r}{\sigma^2} = \psi_t \hat{X}_t \hat{\alpha}_t \implies \hat{X}_t \hat{\alpha}_t = \frac{\mu - r}{\sigma^2 \psi_t}\,.$$

Equating the "dt terms" will let us identify ψ and ϕ. Indeed we get

$$-r\phi_t e^{-\psi_t \hat{X}_t} = e^{-\psi_t \hat{X}_t}\Big[- \psi_t \phi_t \hat{X}_t\big[\,(\hat{\alpha}_t(\mu - r) + r)\,\big] - \hat{X}_t \phi_t \psi_t' + \frac{1}{2}\phi_t \psi_t^2 \hat{\alpha}_t^2 \hat{X}_t^2 \sigma^2 + \phi_t'\Big]\,.$$

Substituting the control and dividing by the exponential term leads to:

$$-r\phi_t = -\phi_t \left(\tfrac{(\mu-r)^2}{\sigma^2} + \psi_t \hat{X}_t r\right) - \hat{X}_t \phi_t \psi_t' + \tfrac{1}{2}\phi_t \tfrac{(\mu-r)^2}{\sigma^2} + \phi_t'\,.$$

This simplifies to

$$-r\phi_t = -\psi_t \phi_t \hat{X}_t r - \tfrac{1}{2}\phi_t \tfrac{(\mu-r)^2}{\sigma^2} - \hat{X}_t \phi_t \psi_t' + \phi_t'\,.$$

From this we get (equating the terms with X_t and without):

$$\begin{aligned}
\phi_t' &= \left(\tfrac{1}{2}\tfrac{(\mu-r)^2}{\sigma^2} - r\right)\phi_t\,, \quad \phi_T = \gamma \\
\psi_t' &= -r\psi_t\,, \quad \psi_T = \gamma\,.
\end{aligned}$$

Hence

$$\begin{aligned}
\phi_t &= \gamma \exp\left((T - t)\left(\tfrac{1}{2}\tfrac{(\mu-r)^2}{\sigma^2} - r\right)\right)\,, \quad t \in [0,T]\,, \\
\psi_t &= \gamma \exp(r(T - t))\,, \quad t \in [0,T]\,.
\end{aligned}$$

So finally the optimal control is:

$$\hat{X}_t \hat{\alpha}_t = e^{-r(T-t)}\frac{\mu - r}{\gamma \sigma^2}\,.$$

A Appendix

A.1 Basic notation and useful review of analysis concepts

Here we set the main notation for the rest of the course. These pages serve as an easy reference.

General For any two real numbers x, y,

$$x \wedge y = \min\{x, y\}, \quad x \vee y = \max\{x, y\}, \quad x^+ = \max\{x, 0\}, \quad x^- = \max\{-x, 0\}.$$

Sets, metrics and matrices $\mathbb{N}$ is the set of strictly positive integers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

$\mathbb{R}^d$ denotes the d-dimensional Euclidean space of real numbers. For any $x = (x_1, \cdots, x_d)$, $y = (y_1, \cdots, y_d)$ in $\mathbb{R}^d$, we denote the inner product by xy and by $|\cdot|$ the Euclidean norm i.e.

$$xy := \sum_{i=1}^{d} x_i y_i \quad \text{and} \quad |x| := \left(\sum_{i=1}^{d} x_i^2\right)^{\frac{1}{2}}$$

$\mathbb{R}^{d\times n}$ denotes the set of real valued $d \times n$-matrices; I_n denotes the $n \times n$-identity matrix. For any $\sigma \in \mathbb{R}^{n\times d}$, $\sigma = (\sigma_{ij})_{1\le i\le n, 1\le j\le d}$ we write the transpose of σ as $\sigma^\top = (\sigma_{ji})_{1\le j\le d, 1\le i\le n} \in \mathbb{R}^{d\times n}$. We write the trace operator of an $n \times n$-matrix σ as $\mathrm{Tr}(\sigma) = \sum_{i=1}^n \sigma_{ii}$. For a matrices we will use the norm $|\sigma| := \big(\mathrm{Tr}(\sigma\sigma^\top)\big)^{1/2}$.

Definition A.1 (Supremum/Infimum)**.** Given a set $S \subset \mathbb{R}$, we say that μ is the supremum of S if *(i)* $\mu \ge x$ for each $x \in S$ and if *(ii)* for every $\varepsilon > 0$ there exists an element $y \in S$ such that $y > \mu - \varepsilon$. We write $\mu = \sup S$.

The infimum is defined symmetrically as follows: λ is the infimum if *(i)* $\lambda \le x$ for each $x \in S$ and if *(ii)* for every $\varepsilon > 0$ there exists an element $y \in S$ such that $y < \lambda + \varepsilon$. We write $\lambda = \inf S$.

Note that supremum is the *least upper bound*, i.e. the smallest real number greater than or equal to all the elements of the set S. Infimum is the *greatest lower bound*, i.e. the largest number smaller than or equal to all the elements of the set S. It is also important to note that the infimum (or supremum) do not necessarily have to belong to the set S.

Functions, derivatives For any set A, the indicator function of A is

$$\mathbb{1}_A(x) = 1 \text{ if } x \in A, \quad \text{otherwise } \mathbb{1}_A(x) = 0 \text{ if } x \notin A.$$

We write $C^k(A)$ is the space of all real-valued continuous functions on A with continuous derivatives up to order $k \in \mathbb{N}_0$, $A \subset \mathbb{R}^n$. In particular $C^0(A)$ is the space of real-valued functions on A that are continuous.

For a real-valued function functions $f = f(t, x)$ defined $I \times A$ we write $\partial_t f$, $\partial_{x_i} f$ and $\partial_{x_i x_j} f$ for $1 \le i, j \le n$ for its partial derivatives. By Df we denote the gradient vector of f and by $D^2 f$ the Hessian matrix of f (whose entries $1 \le i, j \le d$ are given by $\partial_{x_i x_j} f(t, x)$).

Consider an interval I (and think of I as a time interval $I = [0, T]$ or $I = [0, \infty)$). Then $C^{1,2}(I \times A)$ is the set of real valued functions $f = f(t, x)$ on $I \times A$ whose partial derivatives $\partial_t f$, $\partial_{x_i} f$ and $\partial_{x_i x_j} f$ for $1 \le i, j \le n$ exist and are continuous on $I \times A$.

Integration and probability We use $(\Omega, \mathcal{F}, \mathbb{P})$ to denote a probability space with $\mathbb{P}$ being the probability measure and $\mathcal{F}$ the σ-algebra.

"$\mathbb{P}$-a.s." denotes "almost surely for the probability measure $\mathbb{P}$" (we often omit the reference to $\mathbb{P}$). "μ-a.e." denotes "almost everywhere for the measure μ"; here μ will not be a probability measure. This means is that a statement Z made about $\omega \in \Omega$ holds $\mathbb{P}$-a.s. if there is a set $E \in \mathcal{F}$ such that $\mathbb{P}(E) = 0$ and Z is true for *all* $\omega \in E^c = \Omega \setminus E$.

$\mathcal{B}(U)$ is the Borel σ-algebra generated by the open sets of the topological space U.

$\mathbb{E}[X]$ is the expectation of the random variable X with respect to a probability $\mathbb{P}$. $\mathbb{E}[X|\mathcal{G}]$ is the conditional expectation of X given $\mathcal{G}$. The variance of the random variable X, possibly vector valued, is denoted by $\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}(X))(X - \mathbb{E}(X))^\top]$.

Since we may define different measures on the same σ-algebra we must sometimes distinguish which measure is used for expectation, conditional expectation or variance. We thus sometimes write $\mathbb{E}^\mathbb{Q}[X]$, $\mathbb{E}^\mathbb{Q}[X|\mathcal{G}]$ or $\mathrm{Var}^\mathbb{Q}$ to show which measure was used.

General analysis definitions and inequalities

Definition A.2 (Convex function)**.** A function $f : \mathbb{R} \to (-\infty, \infty]$ is called convex if

$$\forall\, \lambda \in [0,1]\ \forall\, x, y \in \mathbb{R} \qquad f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda) f(y).$$

If a function f is convex then it is differentiable a.e. and (with f'_- denoting its left-derivative, f'_+ its right-derivative) and we have

$$f'_+(x) := \lim_{y \searrow x} \frac{f(y) - f(x)}{y - x} = \inf_{y > x} \frac{f(y) - f(x)}{y - x}\,,$$

$$f'_-(x) := \lim_{y \nearrow x} \frac{f(y) - f(x)}{y - x} = \sup_{y < x} \frac{f(y) - f(x)}{y - x}\,.$$

So, from the expression with infimum we see that,

$$\text{if } y > x \text{ then } f'_+(x) \leq \frac{f(y) - f(x)}{y - x} \text{ which implies } f(y) \geq f(x) + f'_+(x)(y-x) \text{ for } y > x.$$

Moreover, from the expression with supremum we see that[17],

$$\text{if } y < x \text{ then } f'_-(x) \geq \frac{f(y) - f(x)}{y - x} \text{ which implies } f(y) \geq f(x) + f'_-(x)(y-x) \text{ for } y < x.$$

We review a few standard analysis inequalities, some not named and some others named: Cauchy-Schwarz, Holder, Young and Gronwall's inequality.

$$\begin{aligned} \forall\, x \in \mathbb{R} &\qquad x \leq 1 + x^2 \\ \forall a, b \in \mathbb{R} &\qquad 2ab \leq a^2 + b^2 \\ \forall\, n \in \mathbb{N}\ \forall a, b \in \mathbb{R} &\qquad |a + b|^n \leq 2^{n-1}\big(|a|^n + |b|^n\big) \end{aligned}$$

Lemma A.3 (Cauchy–Schwarz inequality)**.** *Let H be a Hilbert space with inner product $(\cdot,\cdot)$ and norm $|\cdot|_H$. If $x, y \in H$ then $(x, y) \leq |x|_H |y|_H$.*

Example A.4. i) If $x, y \in \mathbb{R}^d$ then $xy < |x||y|$.

[17] As $y < x$ we multiply by negative number, flipping the inequality.

ii) We can check that $L^2(\Omega)$ with inner product given by $\mathbb{E}[XY]$ for $X, Y \in L^2(\Omega)$ is a Hilbert space. Hence the Cauchy–Schwarz inequality is $\mathbb{E}[XY] \le \big(\mathbb{E}[X^2]\big)^{1/2}\big(\mathbb{E}[Y^2]\big)^{1/2}$.

Lemma A.5 (Young's inequality). *Let $a, b \in \mathbb{R}$. Then for any $\varepsilon \in (0,\infty)$ for any $p, q \in (1,\infty)$ such that $1/p + 1/q = 1$ it holds that*

$$ab \le \varepsilon \frac{|a|^p}{p} + \frac{1}{\varepsilon}\frac{|b|^q}{q}.$$

The above inequality is not the original Young's inequality, that is for the choice $\varepsilon = 1$. The one here is the original Young's inequality with the choice $(ab) = (\varepsilon a)(b/\varepsilon)$.

Lemma A.6 (Gronwall's lemma / inequality). *Let $\lambda = \lambda(t) \ge 0$, $a = a(t)$, $b = b(t)$ and $y = y(t)$ be locally integrable, real valued functions defined on I (with $I = [0,T]$ or $I = [0,\infty)$) such that λy is also locally integrable and for almost all $t \in [0,T]$*

$$y(t) + a(t) \le b(t) + \int_0^t \lambda(s) y(s)\, ds.$$

Then

$$y(t) + a(t) \le b(t) + \int_0^t \lambda(s) e^{\int_s^t \lambda(r)dr}(b(s) - a(s))\, ds \quad \text{for almost all } t \in I.$$

Furthermore, if b is monotone increasing and a is non-negative, then

$$y(t) + a(t) \le b(t) e^{\int_0^t \lambda(r)\, dr}, \quad \text{for almost all } t \in I.$$

If the function y in Gronwall's lemma is continuous then the conclusions hold for all $t \in I$. For proof see Exercise 1.9.

Some fundamental probability results

(Following the notation established in SAF) we define lim inf and lim sup.

Definition A.7 (limsup & liminf). Let $(a_n)_{n\in\mathbb{N}}$ be any sequence in $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$

$$\liminf_{n\to\infty} a_n := \lim_{n\to\infty}\lim_{k\to\infty} \min\{a_n, a_{n+1}, a_{n+2}, \ldots, a_k\} = \inf_n \sup_{k\ge n} a_k,$$
$$\limsup_{n\to\infty} a_n := \lim_{n\to\infty}\lim_{k\to\infty} \max\{a_n, a_{n+1}, a_{n+2}, \ldots, a_k\} = \sup_n \inf_{k\ge n} a_k.$$

Clearly $\liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n$ and if $\lim_{n\to\infty} a_n =: a$ exists, then $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n = a$. On the other hand, if $\liminf_{n\to\infty} a_n \ge \limsup_{n\to\infty} a_n$, then $\lim_{n\to\infty} a_n = a$ exists.

Exercise A.8 (lim sup and lim inf of RV are RV). Show that $\liminf_{n\to\infty} X_n$ and $\limsup_{n\to\infty} X_n$ are random variables for any sequence of random variables X_n.

Lemma A.9 (Fatou's lemma). *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of non-negative random variables. Then*

$$\mathbb{E}\left[\liminf_{n\to\infty} X_n\right] \le \liminf_{n\to\infty} \mathbb{E}[X_n].$$

Moreover,

i) If there exists a r.v. Y such that $\mathbb{E}[\,|Y|\,] < \infty$ and $Y \leq X_n$ $\forall n$ (allows $X_n < 0$), then

$$\mathbb{E}\left[\liminf_{n\to\infty} X_n\right] \leq \liminf_{n\to\infty} \mathbb{E}[X_n].$$

ii) If there exists a r.v. Y such that $\mathbb{E}[\,|Y|\,] < \infty$ and $Y \geq X_n$ $\forall n$, then

$$\mathbb{E}\left[\limsup_{n\to\infty} X_n\right] \geq \limsup_{n\to\infty} \mathbb{E}\left[X_n\right].$$

The first part of the above lemma does not require integrability of the sequence of $(X_n)_{n\in\mathbb{N}}$ due to the use of the Monotone Convergence Theorem in its proof. The enumerated statements follow as a corollary of the first statement. Of course, a version of Fatou's lemma using conditional expectations also exists (simply replace $\mathbb{E}[\cdot]$ with $\mathbb{E}[\cdot|\mathcal{F}_t]$).

Lemma A.10 (Hölder's inequality)**.** *Let $(X, \mathcal{X}, \mu)$ be a measure space (i.e. X is a set, $\mathcal{X}$ a σ-algebra and μ a measure). Let $p, q > 1$ be real numbers s.t. $1/p + 1/q = 1$ or let $p = 1, q = \infty$. Let $f \in L^p(X, \mu)$, $g \in L^q(X, \mu)$. Then*

$$\int_X |fg|\, d\mu \leq \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}} \left(\int_X |g|^q d\mu\right)^{\frac{1}{q}}$$

In particular if p, q are such that $1/p+1/q = 1$ and $X \in L^p(\Omega)$, $Y \in L^q(\Omega)$ are random variables then

$$\mathbb{E}[\,|XY|\,] \leq \mathbb{E}[\,|X|^p\,]^{\frac{1}{p}} \mathbb{E}[\,|Y|^q\,]^{\frac{1}{q}}.$$

Lemma A.11 (Minkowski's inequality or triangle inequality)**.** *Let $(X, \mathcal{X}, \mu)$ be a measure space (i.e. X is a set, $\mathcal{X}$ a σ-algebra and μ a measure). For any $p \in [1, \infty]$ and $f, g \in L^p(X, \mu)$*

$$\left(\int_X |f+g|^p\, d\mu\right)^{\frac{1}{p}} \leq \left(\int_X |f|^p\, d\mu\right)^{\frac{1}{p}} + \left(\int_X |g|^p\, d\mu\right)^{\frac{1}{p}}.$$

Lemma A.12 (Jensen's inequality)**.** *Let f be a convex function and X be any random variable with $\mathbb{E}[|X|] < \infty$. Then*

$$f\big(\mathbb{E}[X]\big) \leq \mathbb{E}\big[f(x)\big].$$

A.2 Some useful results from stochastic analysis

For convenience we state some results from stochastic analysis. Proofs can be found for example in Stochastic Analysis for Finance lecture notes, in [15], [2] or [10].

Probability Space

Let us always assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a fixed probability space. We assume that $\mathcal{F}$ is complete which means that all the subsets of sets with probability zero are included in $\mathcal{F}$. We assume there is a filtration $(\mathcal{F}_t)_{t\in[0,T]}$ (which means $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$) such that $\mathcal{F}_0$ contains all the sets of probability zero.

Stochastic Processes, Martingales

A stochastic process $X = (X_t)_{t\geq 0}$ is a collection of random variables X_t which take values in $\mathbb{R}^d$.

We will always assume that stochastic processes are *measurable*. This means that $(\omega, t) \mapsto X(\omega)_t$ taken as a function from $\Omega \times [0,\infty)$ to $\mathbb{R}^d$ is measurable with respect to σ-algebra $\mathcal{F} \otimes \mathcal{B}([0,\infty))$.[18] This product is defined as the σ-algebra generated by sets $E \times B$ such that $E \in \mathcal{F}$ and $B \in \mathcal{B}([0,\infty))$. From Theorem A.35 we then get that

$$t \mapsto X_t(\omega) \text{ is measurable for all } \omega \in \Omega.$$

We say X is $(\mathcal{F}_t)_{t\geq 0}$ *adapted* if for all $t \geq 0$ we have that X_t is $\mathcal{F}_t$-measurable.

Definition A.13. Let X be a stochastic process that is adapted to $(\mathcal{F}_t)_{t\geq 0}$ and such that for every $t \geq 0$ we have $\mathbb{E}[|X_t|] < \infty$. If for every $0 \leq s < t \leq T$ we have

i) $\mathbb{E}[X_t|\mathcal{F}_s] \geq X_s$ a.s.then the process is called *submartingale.*

ii) $\mathbb{E}[X_t|\mathcal{F}_s] \leq X_s$ a.s.then the process is called *supermartingale.*

iii) $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ a.s.then the process is called *martingale.*

For submartingales we have Doob's maximal inequality:

Theorem A.14 (Doob's submartingale inequality)**.** *Let $X \geq 0$ be an $(\mathcal{F}_t)_{t\in[0,T]}$-submartingale and $p > 1$ be given. Assume $\mathbb{E}\left[X_T^p\right] < \infty$. Then*

$$\mathbb{E}\Big[\sup_{0\leq t\leq T} X_t^p\Big] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[X_T^p\right].$$

Definition A.15 (Local Martingale)**.** A stochastic process X is called a *local martingale* if is there exists a sequence of stopping time $(\tau_n)_{n\in\mathbb{N}}$ such that $\tau_n \leq \tau_{n+1}$ and $\tau_n \to \infty$ as $n \to \infty$ and if the *stopped process* $(X(t \wedge \tau_n))_{t\geq 0}$ is a martingale for every n.

Lemma A.16 (Bounded from below local martingales are supermartingales)**.** *Let $(M_t)_{t\in[0,T]}$ be a local Martingale and assume it is positive or more generally bounded from below. Then M is a super-martingale.*

Proof. The proof makes use of Fatou's Lemma A.9 above. Since M is a local Martingale then there exists a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ increasing to infinity a.s. such that the stopped process $M_t^n := M_{t\wedge\tau_n}$ is a Martingale. We have then, using Fatou's lemma for any $0 \leq s \leq t \leq T$

$$\mathbb{E}[M_t|\mathcal{F}_s] = \mathbb{E}[\liminf_{n\to\infty} M_t^n|\mathcal{F}_s] \leq \liminf_{n\to\infty} \mathbb{E}[M_t^n|\mathcal{F}_s] = \liminf_{n\to\infty} M_s^n = M_s,$$

and hence M is a supermartingale. □

Exercise A.17 (Submartingale)**.** In view of the previous lemma, is a bounded from above local martingale a submartingale?

[18] If the process is almost surely continuous i.e. if the map $[0,\infty) \ni t \mapsto X_t(\omega) \in \mathbb{R}^d$ is continuous for almost all $\omega \in \Omega$ then $\Omega \times [0,\infty) \ni (\omega, t) \mapsto X_{(}\omega) \in \mathbb{R}^d$ is a so-called Carathéodory map the stochastic process will be measurable due to e.g. Aliprantis and Border [1, Lemma 4.51].

Integration Classes and Itô's Formula

Definition A.18. By $\mathcal{H}$ we mean all $\mathbb{R}$-valued and adapted processes g such that for any $T > 0$ we have

$$\|g\|^2_{\mathcal{H}_T} := \mathbb{E}\left[\int_0^T |g_s|^2 ds\right] < \infty.$$

By $\mathcal{S}$ we mean all $\mathbb{R}$-valued and adapted processes g such that for any $T > 0$ we have

$$\mathbb{P}\left[\int_0^T |g_s|^2 ds < \infty\right] = 1.$$

The importance of these two classes is that stochastic integral with respect to W is defined for all integrands in class $\mathcal{S}$ and this stochastic integral is a continous *local* martingale. For the class $\mathcal{H}$ the stochastic integral with respect to W is a martingale.

Definition A.19. By $\mathcal{A}$ we denote $\mathbb{R}$-valued and adapted processes g such that for any $T > 0$ we have

$$\mathbb{P}\left[\int_0^T |g_s| ds < \infty\right] = 1.$$

By $\mathcal{H}^{d\times n}$, $\mathcal{S}^{d\times n}$ we denote processes taking values the space of $d\times n$-matrices such that each component of the matrix is in $\mathcal{H}$ or $\mathcal{S}$ respectively. By $\mathcal{A}^d$ we denote processes taking values in $\mathbb{R}^d$ such that each component is in $\mathcal{A}$

Itô processes and Itô Formula

We will need the multi-dimensional version of the Itô's formula. Let W be an n-dimensional Wiener martingale with respect to $(\mathcal{F})_{t\geq 0}$. Let $\sigma \in \mathcal{S}^{m\times d}$ and let $b \in \mathcal{A}^m$. We say that the d-dimensional process X has the stochastic differential

$$dX_t = b_t\, dt + \sigma_t\, dW_t \tag{A.1}$$

for $t \in [0, T]$, if

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW(s).$$

Such a process is also called an *Itô process.*

The Itô formula or chain rule for stochastic processes Before we go into the main result, let us go over an example from classic analysis. Take three functions, $u = u(t, x)$, $g = g(t)$ and $h = h(t)$ given by $h(t) := u\big(t, g(t)\big)$. Let us compute $\frac{d}{dt}h(t)$.

Since h is given as a composition of functions, we use here is the standard chain for functions of several variables (this takes into account that the variation of h arising from changes in t comes from the variation of g and also from the first component in u). Thus we have

$$\frac{d}{dt}h(t) = \big(\partial_t u\big)\big(t, g(t)\big) + \big(\partial_x u\big)\big(t, g(t)\big)\frac{d}{dt}g(t).$$

We want to see the contrast with Itô formula, which has to be written in integral form (since W has almost everywhere non-differentiable paths). To that end, we integrate

$$\int_0^t \frac{d}{dt}h(s)\, ds = \int_0^t \big(\partial_t u\big)\big(s, g(s)\big)\, ds + \int_0^t \big(\partial_x u\big)\big(s, g(s)\big)\frac{d}{dt}g(s)\, ds$$

and use the Fundamental theorem of calculus

$$h(t) - h(0) = \int_0^t \big(\partial_t u\big)\big(s, g(s)\big)\, ds + \int_0^t \big(\partial_x u\big)\big(s, g(s)\big)\, dg(s)$$

which can be written in the differential notation as

$$dh(t) = \partial_t f\big(t, g(t)\big)\, dt + \partial_x f\big(t, g(t)\big)\, dg(t). \tag{A.2}$$

Compare (A.2) with (A.3) below. You see a fundamental difference: the second derivative term! It appears there exactly because the Wiener process has non-differentiable paths and hence a correction to (A.2) is needed.

We have then the following important result.

Theorem A.20 (Multi-dimensional Itô formula)**.** *Let X be a m-dimensional Itô process given by* (A.1)*. Let $u \in C^{1,2}([0,T] \times \mathbb{R}^m)$. Then the process given by $u(t, X_t)$ has the stochastic differential*

$$\begin{aligned} du\big(t, X_t\big) = \partial_t u\big(t, X_t\big)\, dt &+ \sum_{i=1}^{d} \partial_{x_i} u\big(t, X_t\big)\, dX_t^i \\ &+ \frac{1}{2} \sum_{i,j=1}^{d} \partial_{x_i x_j} u\big(t, X_t\big)\, dX_t^i\, dX_t^j, \end{aligned} \tag{A.3}$$

where for $i, j = 1, \ldots, m$

$$dt\, dt = dt\, dW_t^i = 0, \quad dW_t^i\, dW_t^j = \delta_{ij}\, dt.$$

We now consider a very useful special case. Let X and Y be $\mathbb{R}$-valued Itô processes. We will apply to above theorem with $f(x, y) = xy$. Then $\partial_x f = y$, $\partial_y f = x$, $\partial_{xx} f = \partial_{yy} f = 0$ and $\partial_{xy} f = \partial_{yx} f = 1$. Hence from the multi-dimensional Itô formula we have

$$df\big(X_t, Y_t\big) = Y_t\, dX_t + X_t\, dY_t + \frac{1}{2}\, dY_t\, dX_t + \frac{1}{2}\, dX_t\, dY_t.$$

Hence we have the following corollary

Corollary A.21 (Itô's product rule)**.** *Let X and Y be $\mathbb{R}$-valued Itô processes. Then*

$$d\big(X_t Y_t\big) = X_t\, dY_t + Y_t\, dX_t + dX_t\, dY_t.$$

Martingale Representation Formula and Girsanov's theorem

Theorem A.22 (Lévy characterization)**.** *Let $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration. Let $X = (X_t)_{t\in[0,T]}$ be a continuous m-dimensional local martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ such that $X_0 = 0$ and $dX_t^i\, dX_t^j = \delta_{ij}\, dt$ for $i, j = 1, \ldots, d$. Then X is a Wiener martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.*

So essentially any continuous local martingale with the right quadratic variation is a Wiener process.

Theorem A.23 (Girsanov)**.** *Let $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration. Let $W = (W_t)_{t\in[0,T]}$ be a d-dimensional Wiener martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$. Let $\varphi = (\varphi_t)_{t\in[0,T]}$ be a d-dimensional process adapted to $(\mathcal{F}_t)_{t\in[0,T]}$ such that*

$$\mathbb{E}\Big[\int_0^T |\varphi_s|^2\, ds\Big] < \infty.$$

Let

$$L_t := \exp\left\{-\int_0^t \varphi_s^\top\, dW(s) - \frac{1}{2}\int_0^t |\varphi_s|^2\, ds\right\} \tag{A.4}$$

and assume that $\mathbb{E}[L_T] = 1$. Let $\mathbb{Q}$ be a new measure on $\mathcal{F}_T$ given by the Radon-Nikodym derivative $d\mathbb{Q} = L(T)\, d\mathbb{P}$. Then

$$W_t^{\mathbb{Q}} := W_t + \int_0^t \varphi_s\, ds$$

is a $\mathbb{Q}$-Wiener martingale.

We don't give proof but only make some useful observations.

1. Clearly $L_0 = 1$.
2. The Novikov condition is a useful way of establishing that $\mathbb{E}[L_T] = 1$: if
$$\mathbb{E}\left[e^{\frac{1}{2}\int_0^T |\varphi_t|^2\, dt}\right] < \infty$$
then L is a martingale (and hence $\mathbb{E}[L_T] = \mathbb{E}[L_0] = 1$).
3. Applying Itô's formula to $f(x) = \exp(x)$ and
$$dX_t = -\varphi_t^\top\, dW_t - \frac{1}{2}|\varphi_t|^2\, dt$$
yields
$$dL_t = -L_t\varphi_t^\top\, dW_t.$$

Theorem A.24 (Martingale representation)**.** *Let $W = (W_t)_{t\in[0,T]}$ be a d-dimensional Wiener martingale and let $(\mathcal{F}_t)_{t\in[0,T]}$ be generated by W. Let $M = (M_t)_{t\in[0,T]}$ be a continuous real valued martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.*

Then there exists unique adapted d-dimensional process $h = (h_t)_{t\in[0,T]}$ such that for $t \in [0,T]$ we have

$$M_t = M_0 + \sum_{i=1}^d \int_0^t h_s^i\, dW_s^i.$$

If the martingale M is square integrable then h is in $\mathcal{H}$.

Essentially what the theorem is saying is that we can write continuous martingales as stochastic integrals with respect to some process as long as they're adapted to the filtration generated by the process.

Theorem A.25 (Kolmogorov continuity criteria)**.** *Let $(X_t)_{t\in[0,T]}$ be a stochastic process taking values in a separable Banach space $(E, \|\cdot\|)$ such that for some $C > 0$, $\alpha > 0$, $\beta > 1$ we have that $t, s \in [0,T]$ it holds that*

$$\mathbb{E}\|X_t - X_s\|^\beta \le C|t-s|^{1+\alpha}.$$

Then there is a version of X such that for any $\delta \in (0, \alpha/\beta)$ and for almost all $\omega \in \Omega$ there is a random variable $c = c(\omega)_{\delta,T}$ such that

$$|X_t(\omega) - X_s(\omega)| \leq c(\omega)|t-s|^\delta .$$

That is, the sample paths are δ-Hölder continuous.

A.2.1 PDEs and Feynman–Kac Formula

(This section can be traced back to either [15] or SAF notes (Section 16).)

In the case of deterministic maps b and σ in (3.1), the so-called *diffusion SDE*, we can give the following definition of Infinitesimal generator.

Definition A.26 (Infinitesimal generator (associated to an SDE))**.** Let b and σ be deterministic functions in (3.1). For all $t \in [0,T]$, the following second order differential operator $\mathcal{L}$ is called the *infinitesimal generator associated to the diffusion* (3.1),

$$\mathcal{L}\varphi(t,x) = b(t,x)D\varphi(t,x) + \frac{1}{2}\mathrm{Tr}(\sigma\sigma^\top D^2\varphi)(t,x), \qquad \varphi \in C^{0,2}([0,T]\times\mathbb{R}^m).$$

Although the above definition does seems weird and unfamiliar, the operator $\mathcal{L}$ appears every time one uses the Itô formula to $\varphi(t,X_t)$ where the process $(X_t)_{t\in[0,T]}$ is the solution to (3.1).

Exercise A.27. Let $(X_t)_{t\in[0,T]}$ be the solution to (3.1).

Show that for $\varphi \in C^{1,2}([0,T]\times\mathbb{R})$, we have

$$d\varphi(t,X_t) = \big(\partial_t\varphi + \mathcal{L}\varphi\big)(t,X_t)\,dt + \big(\partial_x\varphi\sigma\big)(t,X_t)\,dW_t.$$

It is possible to see that expectations of functions of stochastic process satisfy certain PDEs.

Exercise A.28. For any $(t,x) \in [0,T]\times\mathbb{R}$, define the stochastic process $(X_s^{t,x})_{s\in[t,T]}$ as

$$X_s^{t,x} = x + W_s - W_t .$$

Let $\mathbb{E}^{t,x}[\,\cdot\,] := \mathbb{E}[\,\cdot\,|X_t = x]$. Define a function $v = v(t,x)$ as

$$v(t,x) = \mathbb{E}^{t,x}\big[\,g\big(X_T\big)\,\big] \qquad \forall\,(t,x) \in [0,T)\times\mathbb{R}\,.$$

Assume that $v \in C^{1,2}([0,T)\times\mathbb{R})$ and that $\big(\partial_x v(s,X_s)\big)_{s\in[t,T]} \in L^2([0,T]\times\Omega;\mathbb{R})$. Show that

$$\begin{aligned} \partial_t v + \frac{1}{2}\partial_{xx}v &= 0 \;\text{ on }\; [0,T)\times\mathbb{R}\,, \\ v(T,\cdot) &= g \;\text{ on }\; \mathbb{R}\,. \end{aligned}$$

Hints:

i) Apply Itô formula to the function v and process $X^{t,x}$ between t and some $\tau \geq t$ a stopping time.

ii) Use the Markov and flow properties to show that $(v(s, X_s^{t,x}))_{s\in[t,T]}$ is a martingale. Use fact that the stochastic integral is a martingale (the condition $\big(\partial_x v(s, X_s)\big)_{s\in[t,T]} \in L^2([0,T]\times\mathbb{R})$ ensures stochastic integral is a martingale). Conclude that the "ds" integral terms from the Itô formula must be 0.

iii) Convince yourself that the PDE must hold; you can use the conclusion of Exercise 4.9.

Solution (to Exercise A.28)**.** First we trivially observe that $v(T,x) = \mathbb{E}[g(X_T)|X_T = x] = g(x)$ and so the terminal conditon holds. It remains to verify the PDE holds.

Let us treat (t,x) as fixed and we omit it from notation for a moment. We have $dX_s = dW_s$ and with Itô formula we get

$$
\begin{aligned}
dv(s, X_s) &= \partial_t v(s, X_s)\, ds + \partial_x v(s, X_s)\, dX_s + \frac{1}{2}\partial_{xx} v(s, X_s)\, dX_s dX_s \\
&= \Big(\partial_t + \frac{1}{2}\partial_{xx}\Big) v(s, X_s)\, ds + \partial_x v(s, X_s)\, dW_s\,.
\end{aligned}
$$

Integrating from t to an arbitrary stopping time $\tau \geq t$ we get

$$
v(\tau, X_\tau^{t,x}) - v(t,x) = \int_t^\tau \Big(\partial_t + \frac{1}{2}\partial_{xx}\Big) v(s, X_s^{t,x})\, ds + \int_t^\tau \partial_x v(s, X_s)\, dW_s\,.
$$

Next we show that $(v(s, X_s^{t,x}))_{s\in[t,T]}$ is a martingale. Take $s \geq s' \geq t$ and observe that

$$
\mathbb{E}[v(s, X_s^{t,x})|\mathcal{F}_{s'}] = \mathbb{E}[v(s, X_s^{t,x})|X_{s'}^{t,x}] = v(s, X_s^{s', X_{s'}^{t,x}}) = v(s, X_s^{t,x})
$$

where the first equality holds by Markov property, second equality is just what the conditional expectation is while third equality is the flow property. Thus $(v(s, X_s^{t,x}))_{s\in[t,T]}$ is a martingale.

We note that $v(\tau, X_\tau^{t,x}) = \mathbb{E}\Big[g(X_T^{\tau, X_\tau^{t,x}})\Big] = \mathbb{E}[g(X_T^{t,x})] = v(t,x)$ again due to the flow property. Hence

$$
0 = v(\tau, X_\tau^{t,x}) - v(t,x) = \int_t^\tau \Big(\partial_t + \frac{1}{2}\partial_{xx}\Big) v(s, X_s^{t,x})\, ds + \int_t^\tau \partial_x v(s, X_s)\, dW_s\,.
$$

Moreover, since we're assuming $\big(\partial_x v(s, X_s)\big)_{s\in[t,T]} \in L^2([0,T]\times\mathbb{R})$ we know that the stochastic integral is a martingale i.e. almost surely

$$
0 = \int_t^\tau \Big(\partial_t + \frac{1}{2}\partial_{xx}\Big) v(s, X_s^{t,x})\, ds
$$

with $\tau \geq t$ an arbitrary stopping time. This can only be true if $\partial_t v + \frac{1}{2}\partial_{xx} v = 0$, which we prove in Exercise 4.9.

It is possible, for certain classes of SDE and differential equations, to write the solution to a PDE as an expectation of (a function of) the solution to the SDE associated to the differential operator appearing in the PDE; it is not surprising that the PDE differential operator must be the infinitesimal generator. That is the core message of the next result.

Theorem A.29 (Feynman-Kac formula in 1-dim)**.** *Assume that the function $v : [0,T] \times \mathbb{R} \to \mathbb{R}$ belongs to $C^{1,2}([0,T) \times \mathbb{R}) \cap C^0([0,T] \times \mathbb{R})$ and is a solution to the following boundary value problem*

$$\partial_t v(t,x) + b(t,x)\partial_x v(t,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx} v(t,x) - rv(t,x) = 0, \tag{A.5}$$

$$v(T,x) = h(x), \tag{A.6}$$

where b and σ are deterministic functions.

For any $(t,x) \in [0,T] \times \mathbb{R}$, define the stochastic process $(X_s)_{s\in[t,T]}$ as the solution to the SDE

$$dX_s = b(s,X_s)\,ds + \sigma(s,X_s)\,dW_s, \quad \forall\, s \in [t,T], \qquad X_t = x. \tag{A.7}$$

Assume that the stochastic process $\big(e^{-rs}\sigma(s,X_s)\partial_x v(s,X_s)\big)_{s\in[t,T]} \in L^2([0,T] \times \mathbb{R})$.

Then the solution v of (A.5)-(A.6) *can be expressed as (with $\mathbb{E}_{t,x}[\cdot] = \mathbb{E}[\cdot|X_t = x]$)*

$$v(t,x) = e^{-r(T-t)}\mathbb{E}_{t,x}\big[\, h\big(X_T\big)\,\big] \qquad \forall\,(t,x) \in [0, T \times \mathbb{R}.$$

Proof. The proof is rather straightforward and is based on a direct application of Itô's formula.

Define the process $(Y_s)_{s\in[t,T]}$ as $Y_s = e^{-rs}v(s,X_s)$ where X is given by (A.7). Applying Itô's formula to Y, i.e. computing dY_s gives

$$\begin{aligned}
dY_s =&\ d\Big(e^{-rs}v(s,X_s)\Big)\\
&= (-r)e^{-rs}v\,ds + e^{-rs}\partial_s v\,ds + e^{-rs}\partial_x v\,dX_s + \frac{1}{2}e^{-rs}\partial_{xx}v\big(\,dX_s\big)^2\\
&= e^{-rs}\big[-rv + \partial_t v + b\partial_x v + \frac{1}{2}\sigma^2\partial_{xx}v\big]\,ds + e^{-rs}\big[\sigma\partial_x v\big]\,dW_s,
\end{aligned}$$

where the v function is evaluated in point (s,X_s). Using the equality given by (A.5) we see that the ds term disappears completely leaving

$$dY_s = d\Big(e^{-rs}v(s,X_s)\Big) = e^{-rs}\big[\ \sigma(s,X_s)\partial_x v(s,X_s)\ \big]\,dW_s.$$

Integrating both sides from $s = t$ to $s = T$ gives

$$\begin{aligned}
\big[e^{-rs}v(s,X_s)\big]\Big|_{s=t}^{s=T} &= \int_t^T e^{-ru}\sigma(u,X_u)\partial_x v(u,X_u)\,dW_u\\
\Leftrightarrow e^{-rt}v(t,X_t) &= e^{-rT}v(T,X_T) - \int_t^T e^{-ru}\sigma(u,X_u)\partial_x v(u,X_u)\,dW_u,\\
\Leftrightarrow v(t,X_t) &= e^{-r(T-t)}v(T,X_T) - \int_t^T e^{-r(u-t)}\sigma(u,X_u)\partial_x v(u,X_u)\,dW_u.
\end{aligned}$$

Taking expectations $\mathbb{E}_{(t,x)}[\cdot]$ on both sides (recall that the process X starts at time t in position x; this is the meaning of the subscript (t,x) in the expectation sign),

$$v(t,X_t) = e^{-r(T-t)}\mathbb{E}_{t,x}\big[v(T,X_T)\big] = e^{-r(T-t)}\mathbb{E}_{t,x}\big[h(X_T)\big],$$

where the expectation of the stochastic integral disappears due to the properties of the stochastic integral, since by assumption we have $\big(e^{-rs}\sigma(s,X_s)\partial_x v(s,X_s)\big)_{s\in[t,T]} \in L^2([0,T] \times \mathbb{R})$. □

Exercise A.30 (Two extensions of the Feynman-Kac formula)**.** a) Redo the previous proof when the constant r is replaced by a function $r : [0, T] \times \mathbb{R} \to \mathbb{R}$; assume r to be bounded and continuous. Hint instead of e^{-rs}, use $\exp\{-\int_t^s r(u, X_u)\, du\}$.

b) Redo the previous proof when the PDE (A.5) is equal to some $f(t, x)$ instead of being equal to zero.

A.2.2 Fokker–Planck–Kolmogorov equation

Let us fix a reference probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Let $B(\mathbb{R}^d)$ denote the set of all $\mathbb{R}$-valued, bounded measurable functions on $\mathbb{R}^d$. Let $\mathcal{P}(\mathbb{R}^d)$ denote the set of all probability measures on $\mathbb{R}^d$. For any $\mu \in \mathcal{P}(\mathbb{R}^d)$ and any $\phi \in B(\mathbb{R}^d)$ let

$$\langle \mu, \phi \rangle := \int_{\mathbb{R}^d} \phi(x)\mu(dx)\,.$$

Recall that if $\xi : \Omega \to \mathbb{R}^d$ then $\mathrm{Law}(\xi) \in \mathcal{P}(\mathbb{R}^d)$ is given by $\mathrm{Law}(\xi)(B) = \mathbb{P}(\xi \in B)$ for all $B \in \mathcal{B}(\mathbb{R}^d)$. Note that for any $\phi \in B(\mathbb{R}^d)$ we have

$$\langle \mathrm{Law}(\xi), \phi \rangle = \int_{\mathbb{R}^d} \phi(x)\mathrm{Law}(\xi)(dx) = \mathbb{E}[\phi(\xi)]\,.$$

Let us introduce $L : C_b^2(\mathbb{R}^d) \to B(\mathbb{R}^d)$ as

$$L\phi := b \cdot \nabla\phi + \frac{1}{2}\mathrm{tr}[\sigma\sigma^\top \nabla^2\phi]\,, \tag{A.8}$$

where $\nabla\phi$ is the gradient vector of all the first order partial derivatives of ϕ and $\nabla^2\phi$ is the Hessian matrix of all the second order partial derivatives of ϕ.

Theorem A.31. *Assume that there is K such that $|b(x)| + |\sigma(x)| \le K(1 + |x|)$. Let X be a solution to*

$$X_t = \xi + \int_0^t b(X_s)\, ds + \int_0^t \sigma(X_s)\, dW_s\,,\ \ t \in I\,,\ \ \xi \ \text{ a given } \mathbb{R}^d\text{-valued r.v.} \tag{A.9}$$

on I such that for any $T > 0$ there is $C_T > 0$ so that $\sup_{t\le T} \mathbb{E}|X_t|^2 \le C_T\mathbb{E}|\xi|^2$. Let $\nu_t :=$ Law(X_t). Then ν_t is a solution to

$$\langle \nu_t, \phi \rangle = \langle \nu_0, \phi \rangle + \int_0^t \langle \nu_s, L\phi \rangle\, ds \ \ \forall \phi \in C_b^2(\mathbb{R}^d)\,,\ \ t \in I\,. \tag{A.10}$$

Proof. Let us fix $t \in I$. Let us apply Itô's formula to the process X and the function ϕ. We get

$$\begin{aligned}
\phi(X_t) &= \phi(X_0) + \int_0^t b(X_s) \cdot \nabla\phi(X_s)\, ds + \int_0^t (\nabla\phi(X_s))^\top \sigma(X_s)\, dW_s \\
&\quad + \int_0^t \frac{1}{2}\mathrm{tr}[\sigma(X_s)\sigma(X_s)^\top \nabla^2\phi(X_s)]\, ds \\
&= \phi(X_0) + \int_0^t (L\phi)(X_s)\, ds + \int_0^t (\nabla\phi(X_s))^\top \sigma(X_s)\, dW_s\,.
\end{aligned}$$

Due to our assumptions

$$\mathbb{E}\int_0^t |\sigma(X_s)|^2\, ds \le \mathbb{E}\int_0^t K^2(1 + |X_s|)^2\, ds \le 2K^2t + 2K^2t\sup_{s\le t}\mathbb{E}|X_s|^2 < \infty$$

and so the stochastic integral above is a martingale. Thus taking expectation we get

$$\mathbb{E}[\phi(X_t)] = \mathbb{E}[\phi(X_0)] + \int_0^t \mathbb{E}[(L\phi)(X_s)]\,ds\,.$$

Since $\nu_t = \mathrm{Law}(X_t)$ we get

$$\langle \nu_t, \phi\rangle = \langle \nu_0, \phi\rangle + \int_0^t \langle \nu_s, L\phi\rangle\,ds\,.$$

Noting that $t \in I$ and $\phi \in C_b^2(\mathbb{R}^d)$ were arbitrary concludes the proof. $\square$

We give the following result without proof.

Theorem A.32. *Assume that there is $K > 0$ such that for all i, j we have $|\partial_{x_i} b_i|_{L^\infty(\mathbb{R}^d)} \leq K$ and $|\partial_{x_i}\partial_{x_j} a_{i,j}|_{L^\infty(\mathbb{R}^d)} \leq K$. Show that if ν is a measure valued solution to* (A.10) *such that there is $u_0 \in L^2(\mathbb{R}^d)$ and $u_0(x)\,dx = \nu_0(dx)$ then there is $u : \mathbb{R}^d \times I \to \mathbb{R}$ such that $u(t,x)\,dx = \nu_t(dx)$ and for any $t' > 0$ we have $u \in L^\infty(0,t';L^2(\mathbb{R}^d))$.*

It is further possible to show that (under further assumptions) this density has more regularity (e.g. by showing it belongs to appropriate Sobolev space and then using Sobolev embedding [5, Ch. 7]). We won't do that here but just note that in that case the FPKE equation has a classical solution.

Alternatively it is possible to develop a strong solution theory for these equations using what is sometimes referred to as the "paramterix" method, see [6]; in fact the book covers both forward and backward equations.

Corollary A.33. *Let $\nu = (\nu_t)_{t\in I}$ be as in Theorem A.31. If there exists $u \in C^{2,1}(\mathbb{R}^d \times I)$ such that $u(t,\cdot)$ is the density of ν_t w.r.t. the Lebesgue measure (i.e. we have $\nu_t(dx) = u(x,t)\,dx$) for every $t \in I \setminus \{0\}$ then u solves*

$$\partial_t u = L^* u \quad \text{on } \mathbb{R}^d \times (I \setminus \{0\})\,,$$

where

$$L^*\psi = \nabla \cdot \Big(-\psi b + \frac{1}{2}\nabla \cdot \big(\sigma\sigma^\top \psi\big)\Big)\,.$$

Proof. In this proof all integrals without explicitly specified domain of integration are over $\mathbb{R}^d$ and all summation indices run from 1 to d. From Theorem A.31 we get

$$\int u(t,\cdot)\phi\,dx = \int u(0,\cdot)\phi\,dx + \int_0^t \int u(s,\cdot)L\phi\,dx\,.$$

From integration by parts (noting ϕ has compact support) we get

$$\int u(s,\cdot) b\cdot\nabla\phi\,dx = -\int \nabla\cdot\big(u(s,\cdot)b\big)\phi\,dx$$

and

$$\int u(s,\cdot)\mathrm{tr}[\sigma\sigma^\top\nabla^2\phi]\,dx = \sum_{i,j}\int u(s,\cdot)\sigma_{ij}\partial_{x_i}\partial_{x_j}\phi\,dx$$
$$= \sum_{i,j}\int \partial_{x_i}\partial_{x_j}\big[\sigma_{ij}u(s,\cdot)\big]\phi\,dx = \int \nabla\nabla\cdot\big[\sigma_{ij}u(s,\cdot)\big]\phi\,dx\,.$$

Hence for all $\phi \in C_0^2(\mathbb{R}^d)$ we have

$$\int u(t,\cdot)\phi\,dx = \int u(0,\cdot)\phi\,dx + \int_0^t \int L^* u(s,\cdot)\phi\,dx\,.$$

From this and differentiablity in time of u the conclusion follows. $\square$

A.3 Other useful results

The aim of this section is to collect, mostly without proofs, results that are needed or useful for this course but that cannot be covered in the lectures i.e. prerequisites. You are expected to be able to use the results given here.

A.3.1 Linear Algebra

The inverse of a square real matrix A exists if and only if $\det(A) \neq 0$.

The inverse of square real matricies A and B exists if and only if the inverse of AB exists and moreover $(AB)^{-1} = B^{-1}A^{-1}$.

The inverse of a square real matrix A exists if and only if the inverse of A^T exists and $(A^T)^{-1} = (A^{-1})^T$.

If x is a vector in $\mathbb{R}^d$ then $\mathrm{diag}(x)$ denotes the matrix in $\mathbb{R}^{d\times d}$ with the entries of x on its diagonal and zeros everywhere else. The inverse of $\mathrm{diag}(x)$ exists if and only if $x_i \neq 0$ for all $i = 1,\ldots,d$ and moreover

$$\mathrm{diag}(x)^{-1} = \mathrm{diag}(1/x_1, 1/x_2, \ldots, 1/x_d).$$

A.3.2 Real Analysis and Measure Theory

Let $(X, \mathcal{X}, \mu)$ be a measure space (i.e. X is a set, $\mathcal{X}$ a σ-algebra and μ a measure).

Lemma A.34 (Fatou's Lemma)**.** *Let $f_1, f_2, \ldots$ be a sequence of non-negative and measurable functions. Then the function defined point-wise as*

$$f(x) := \liminf_{k\to\infty} f_k(x)$$

is $\mathcal{X}$-measurable and

$$\int_X f\,d\mu \leq \liminf_{k\to\infty} \int_X f_k\,d\mu.$$

Consider sets X and Y with σ-algebras $\mathcal{X}$ and $\mathcal{Y}$. By $\mathcal{X}\times\mathcal{Y}$ we denote the collection of sets $C = A\times B$ where $A\in\mathcal{X}$ and $B\in\mathcal{Y}$. By $\mathcal{X}\otimes\mathcal{Y} = \sigma(\mathcal{X}\times\mathcal{Y})$, which is the σ-algebra generated by $\mathcal{X}\times\mathcal{Y}$.

Theorem A.35. *Let $f : X\times Y \to \mathbb{R}$ be a measurable function, i.e. measurable with respect to the σ-algebras $\mathcal{X}\otimes\mathcal{Y}$ and $\mathcal{B}(\mathbb{R})$. Then for each $x\in X$ the function $y\mapsto f(x,y)$ is measurable with respecto to $\mathcal{Y}$ and $\mathcal{B}(\mathbb{R})$. Similarly for each $y\in Y$ the function $x\mapsto f(x,y)$ is measurable with respecto to $\mathcal{X}$ and $\mathcal{B}(\mathbb{R})$.*

The proof is short and so it's easiest to just include it here.

Proof. We first consider functions of the form $f = \mathbb{1}_C$ with $C\in\mathcal{X}\otimes\mathcal{Y}$. Let

$$\mathcal{H} = \{C\in\mathcal{X}\otimes\mathcal{Y} : y\mapsto \mathbb{1}_C(x,y) \text{ is } \mathcal{F}-\text{measurable for each fixed } x\in E\}.$$

It is easy to check that $\mathcal{H}$ is a σ-algebra. Moreover if $C = A\times B$ with $A\in\mathcal{X}$ and $B\in\mathcal{Y}$ then

$$y\mapsto \mathbb{1}_C(x,y) = \mathbb{1}_A(x)\mathbb{1}_B(y).$$

As x is fixed $\mathbb{1}_A(x)$ is just a constant and since $B \in \mathcal{Y}$ the function $y \mapsto \mathbb{1}_A(x)\mathbb{1}_B(y)$ must be measurable. Hence $\mathcal{X} \times \mathcal{Y} \subseteq \mathcal{H}$ and thus $\mathcal{X} \otimes \mathcal{Y} \subseteq \mathcal{H}$. But $\mathcal{H} \subseteq \mathcal{X} \otimes \mathcal{Y}$ and so $\mathcal{H} = \mathcal{X} \otimes \mathcal{Y}$. Hence if f is a simple function then the conclusion of the theorem holds.

Now consider $f \geq 0$ and let f_n be a sequence of simple functions increasing to f. Then for a fixed x the function $y \mapsto g_n(y) = f_n(x, y)$ is measurable. Moreover since $g(y) = \lim_{n\to\infty} g_n(y) = f(x, y)$ and since the limit of measurable functions is measurable we get the result for $f \geq 0$. For general $f = f^+ - f^-$ the result follows using the result for $f^+ \geq 0, f^- \geq 0$ and noting that the difference of measurable functions is measurable. $\square$

Consider measure spaces $(X, \mathcal{X}, \mu_x)$, $(Y, \mathcal{Y}, \mu_y)$. That is, X and Y are sets, $\mathcal{X}$ and $\mathcal{Y}$ are σ-algebras and μ_x and μ_y are measures on $\mathcal{X}$ and $\mathcal{Y}$ respectively. For all details on Fubini's Theorem we refer to Kolmogorov and Fomin [13].

Theorem A.36 (Fubini)**.** *Let μ be the Lebesgue extension of $\mu_x \otimes \mu_y$. Let $A \in \mathcal{X} \otimes \mathcal{Y}$. and let $f : A \to \mathbb{R}$ be a measurable function (considering $\mathcal{B}(\mathbb{R})$, the Borel σ-algebra on $\mathbb{R}$). If f is integrable i.e. if*

$$\int_A |f(x, y)| d\mu < \infty$$

then

$$\int_A f(x, y) d\mu = \int_X \left[\int_{A_x} f(x, y) d\mu_y \right] d\mu_x = \int_Y \left[\int_{A_y} f(x, y) d\mu_x \right] d\mu_y,$$

where $A_x := \{y \in Y : (x, y) \in A\}$ and $A_y := \{x \in X : (x, y) \in A\}$.

Remark A.37. The conclusion of Fubini's theorem implies that for μ_x-almost all x the integral $\int_{A_x} f(x, y) d\mu_y$ exists which in turn implies that the function $f(x, \cdot) : A_x \to \mathbb{R}$ must be measurable. This statement also holds if we exchange x for y.

A.3.3 Conditional Expectation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be given.

Theorem A.38. *Let X be an integrable, $\mathcal{F}$-measurable, random variable. If $\mathcal{G} \subseteq \mathcal{F}$ is a σ-algebra then there exists a unique $\mathcal{G}$ measurable random variable $\hat{Y}$ such that*

$$\forall G \in \mathcal{G} \quad \int_G X d\mathbb{P} = \int_G \hat{Y} d\mathbb{P}.$$

The proof can be found in xxxx xxxx but an outline is as follows. Let us write $L^2(\mathcal{G})$ for the space of all $\mathcal{G}$-measurable r.v.s Y such that $\|Y\|_2 := \left(\mathbb{E}[|Y|^2]\right)^{1/2} < \infty$. Then we have the following:

Lemma A.39. *Let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra and let $X \in L^2(\mathcal{F})$. Let $\alpha = \inf_{Y \in L^2(\mathcal{G})} \|X - Y\|_2$. Then there exists $\hat{Y} \in L^2(\mathcal{G})$ such that $\|X - \hat{Y}\|_2 = \alpha$.*

Proof. xxxx xxxx $\square$

Corollary A.40 (Orthogonal projection)**.** *Let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra and let $X \in L^2(\mathcal{F})$. Let $\hat{Y}$ be given by Lemma A.39. Then for all $Z \in L^2(\mathcal{G})$ we have*

$$\mathbb{E}\big[(X - \hat{Y})Z\big] = 0\,.$$

Proof. Let α be also given by Lemma A.39. Fix $Z \in L^2(\mathcal{G})$. If $Z = 0$ a.s. then the conclusion is true so we can assume that $Z \neq 0$ a.s. Then for any $t \in \mathbb{R}$ we have

$$\begin{aligned}\alpha^2 \leq \|X - (\hat{Y} + tZ)\|_2^2 &= \|X - \hat{Y}\|_2^2 - 2\|(X - \hat{Y})tZ\|_2^2 + \|tZ\|_2^2 \\ &= \alpha^2 - 2t\mathbb{E}\big[(X - \hat{Y})Z\big] + t^2\mathbb{E}\big[Z^2\big]\,.\end{aligned}$$

Since this must hold for all $t \in \mathbb{R}$ it must also hold for the one minimizing the right hand side. Thus we must have

$$\alpha^2 \leq \alpha^2 - \frac{\Big(\mathbb{E}\big[(X - \hat{Y})Z\big]\Big)^2}{\mathbb{E}\big[Z^2\big]}\,.$$

But this can only be true if $\mathbb{E}\big[(X - \hat{Y})Z\big] = 0$. □

Now take $G \in \mathcal{G}$ and apply Corollary A.40 with $Z = \mathbb{1}_G$. Then

$$\mathbb{E}[X\mathbb{1}_G] = \mathbb{E}[\hat{Y}\mathbb{1}_G]\,.$$

This proves Theorem A.38 for $X \in L^2(\mathcal{F})$. This can be extended to $X \in L^1(\mathcal{F})$ by approximation (consider $X \geq 0$ first and take $X_n = X \wedge n \in L^2(\mathcal{F})$.

Definition A.41. Let X be an integrable random variable. If $\mathcal{G} \subseteq \mathcal{F}$ is a σ-algebra then $\mathcal{G}$-measurable random variable from Theorem A.38 is called the *conditional expectation* of X given $\mathcal{G}$ and write $\mathbb{E}(X|\mathcal{G}) := \hat{Y}$.

Conditional expectations are rather abstract notion so two examples might help.

Example A.42. Consider $\mathcal{G} := \{\emptyset, \Omega\}$. So $\mathcal{G}$ is just the trivial σ-algebra.

First note that any random variable Z measurable w.r.t. $\mathcal{G}$ must be a constant. Indeed if Z is $\mathcal{G}$-measurable and non-constant, then it assumes at least two values c_1 and c_2. The set $Z^{-1}(\{c_1\})$ must be in $\mathcal{G}$ (as $\{c_1\}$ is a closed set) but this set is non-empty (as c_1 is a value of Z) and not Ω (as the points ω where Z assumes the value c_2 are not in it). So this set cannot be in $\mathcal{G}$ leading to a contradiction.

For a random variable X we then have, by definition, that Z is the conditional expectation (denoted $\mathbb{E}[X|\mathcal{G}]$), if and only if

$$\int_\Omega Z d\mathbb{P} = \int_\Omega X d\mathbb{P}.$$

The right hand side of the above expression is in fact just $\mathbb{E}X$ and so the equality would be satisfied if we set $Z = \mathbb{E}X$ (just a constant). Indeed then (going right to left)

$$\mathbb{E}X = \int_\Omega X d\mathbb{P} = \int_\Omega Z d\mathbb{P} = \int_\Omega \mathbb{E}X d\mathbb{P} = \mathbb{E}X \int_\Omega d\mathbb{P} = \mathbb{E}X.$$

Example A.43. Let $X \sim N(0,1)$. Let $\mathcal{G} = \{\emptyset, \{X \leq 0\}, \{X > 0\}, \Omega\}$. One can (and should) check that this is a σ-algebra. By definition the conditional expectation is a unique random variable that satisfies

$$\begin{aligned}\int_\Omega \mathbb{1}_{\{X>0\}} Z d\mathbb{P} &= \int_\Omega \mathbb{1}_{\{X>0\}} X d\mathbb{P}, \\ \int_\Omega \mathbb{1}_{\{X\leq 0\}} Z d\mathbb{P} &= \int_\Omega \mathbb{1}_{\{X\leq 0\}} X d\mathbb{P}, \\ \int_\Omega Z d\mathbb{P} &= \int_\Omega X d\mathbb{P}\,.\end{aligned} \tag{A.11}$$

It is a matter of integrating with respect to normal density to find out that

$$\int_\Omega \mathbb{1}_{\{X>0\}} X d\mathbb{P} = \int_0^\infty x\phi(x)dx = \frac{1}{2}\sqrt{\frac{2}{\pi}}, \quad \int_\Omega \mathbb{1}_{\{X\le 0\}} X d\mathbb{P} = -\frac{1}{2}\sqrt{\frac{2}{\pi}}. \tag{A.12}$$

Since Z must be $\mathcal{G}$ measurable it can only take two values:

$$Z = \begin{cases} z_1 & \text{on} \quad \{X > 0\}, \\ z_2 & \text{on} \quad \{X \le 0\}, \end{cases}$$

for some real constants z_1 and z_2 to be yet determined. But (A.11) and (A.12) taken together imply that

$$\frac{1}{2}\sqrt{\frac{2}{\pi}} = \int_\Omega \mathbb{1}_{\{X>0\}} Z d\mathbb{P} = \int_\Omega \mathbb{1}_{\{X>0\}} z_1 d\mathbb{P} = z_1 \mathbb{P}(X > 0) = \frac{1}{2} z_1.$$

Hence $z_1 = \sqrt{2/\pi}$. Similarly we calculate that $z_2 = -\sqrt{2/\pi}$. Finally we check that the third equation in (A.11) holds. Thus

$$\mathbb{E}[X|\mathcal{G}] = Z = \begin{cases} \sqrt{\frac{2}{\pi}} & \text{on} \quad \{X > 0\}, \\ -\sqrt{\frac{2}{\pi}} & \text{on} \quad \{X \le 0\}. \end{cases}$$

Here are some further important properties of conditional expectations which we present without proof.

Theorem A.44 (Properties of conditional expectations)**.** *Let X and Y be random variables. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$.*

1. *If $\mathcal{G} = \{\emptyset, \Omega\}$ then $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}X$.*
2. *If $X = x$ a. s. for some constant $x \in \mathbb{R}$ then $\mathbb{E}(X|\mathcal{G}) = x$ a.s..*
3. *For any $\alpha, \beta \in \mathbb{R}$*

   $$\mathbb{E}(\alpha X + \beta Y|\mathcal{G}) = \alpha\mathbb{E}(X|\mathcal{G}) + \beta\mathbb{E}(Y|\mathcal{G}).$$

   This is called linearity*.*
4. *If $X \le Y$ almost surely then $\mathbb{E}(X|\mathcal{G}) \le \mathbb{E}(Y|\mathcal{G})$ a.s..*
5. *$|\mathbb{E}(X|\mathcal{G})| \le \mathbb{E}(|X|\,|\mathcal{G})$.*
6. *If $X_n \to X$ a. s. and $|X_n| \le Z$ for some integrable Z then $\mathbb{E}(X_n|\mathcal{G}) \to \mathbb{E}(X|\mathcal{G})$ a. s. . This is the "dominated convergence theorem for conditional expectation".*
7. *If Y is $\mathcal{G}$ measurable then $\mathbb{E}(XY|\mathcal{G}) = Y\mathbb{E}(X|\mathcal{G})$.*
8. *Let $\mathcal{H}$ be a sub-σ-algebra of $\mathcal{G}$. Then*

   $$\mathbb{E}(X|\mathcal{H}) = \mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H}).$$

   This is called the tower property*. A special case is $\mathbb{E}X = \mathbb{E}(\mathbb{E}(X|\mathcal{G}))$.*
9. *If $\sigma(X)$ is independent of $\mathcal{G}$ then $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}X$.*

Example A.45. Let X and Z be $\mathcal{F}$-measurable, integrable and independent. Let

$$Y = f(X) + g(Z)$$

for some deterministic functions f and g s.t. $\mathbb{E}[|f(X)|] + \mathbb{E}[|g(Z)|] < \infty$. We would like to calculate $\mathbb{E}[Y|X] = \mathbb{E}[Y|\sigma(X)]$. Clearly

$$\mathbb{E}[Y|X] = \mathbb{E}[f(X) + g(Z)|X] = f(X) + \mathbb{E}[g(Z)]\,,$$

due to points 7. and 9. of Theorem A.44 .

Definition A.46. Let X and Y be two random variables. The *conditional expectation of X given Y* is defined as $\mathbb{E}(X|Y) := \mathbb{E}(X|\sigma(Y))$, that is, it is the conditional expectation of X given the σ-algebra generated by Y.

Definition A.47. Let X a random variables and $A \in \mathcal{F}$ an event. The *conditional expectation of X given A* is defined as $\mathbb{E}(X|A) := \mathbb{E}(X|\sigma(A))$. This means it is the conditional expectation of X given the sigma algebra generated by A i.e. $\mathbb{E}(X|A) := \mathbb{E}(X|\{\emptyset, A, A^c, \Omega\})$.

We can immediately see that $\mathbb{E}(X|A) = \mathbb{E}(X|\mathbb{1}_A)$.

Recall that if X and Y are jointly continuous random variables with joint density $(x, y) \mapsto f(x, y)$ then for any measurable function $\rho : \mathbb{R}^2 \to \mathbb{R}$ such that $\mathbb{E}|\rho(X, Y)| < \infty$ we have

$$\mathbb{E}\rho(X, Y) = \int_{\mathbb{R}} \int_{\mathbb{R}} \rho(x, y) f(x, y) dy dx.$$

Moreover the marginal density of X is

$$g(x) = \int_{\mathbb{R}} f(x, y) dy$$

while the marginal density of Y is

$$h(y) = \int_{\mathbb{R}} f(x, y) dx.$$

Theorem A.48. *Let X and Y be jointly continuous random variables with joint density $(x, y) \mapsto f(x, y)$. Then for any measurable function $\varphi : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}|\varphi(Y)| < \infty$ the conditional expectation of $\varphi(Y)$ given X is*

$$\mathbb{E}(\varphi(Y)|X) = \psi(X)$$

where $\psi : \mathbb{R} \to \mathbb{R}$ is given by

$$\psi(x) = \mathbb{1}_{\{g(x)>0\}} \frac{\int_{\mathbb{R}} \varphi(y) f(x, y) dy}{g(x)}.$$

Proof. Every A in $\sigma(X)$ must be of the form $A = \{\omega \in \Omega : X(\omega) \in B\}$ for some B in $\mathcal{B}(\mathbb{R})$. We need to show that for any such A

$$\int_A \psi(X) d\mathbb{P} = \int_A \varphi(Y) d\mathbb{P}.$$

But since $\mathbb{E}|\varphi(Y)| < \infty$ we can use Fubini's theorem to show that

$$
\begin{aligned}
\int_A \psi(X) d\mathbb{P} &= \mathbb{E}\mathbb{1}_A \psi(X) = \mathbb{E}\mathbb{1}_{\{X\in B\}}\psi(X) = \int_B \psi(x) g(x) dx \\
&= \int_B \int_{\mathbb{R}} \varphi(y) f(x,y) dy dx = \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}_B(x)\varphi(y) f(x,y) dx dy \\
&= \mathbb{E}\mathbb{1}_{\{X\in B\}}\varphi(Y) = \int_A \varphi(Y) d\mathbb{P}.
\end{aligned}
$$

□

Let on $(\Omega, \mathcal{F})$ be a measurable space. Recall that we say that a measure $\mathbb{Q}$ is absolutely continuous with respect to a measure $\mathbb{P}$ if $\mathbb{P}(E) = 0$ implies that $\mathbb{Q}(E) = 0$. We write $\mathbb{Q} << \mathbb{P}$.

Proposition A.49. *Take two probability measures $\mathbb{P}$ and $\mathbb{Q}$ such that $\mathbb{Q} << \mathbb{P}$ with*

$$d\mathbb{Q} = \Lambda d\mathbb{P}.$$

Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Then $\mathbb{Q}$ almost surely $\mathbb{E}[\Lambda|\mathcal{G}] > 0$. Moreover for any $\mathcal{F}$-random variable X we have

$$\mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}] = \frac{\mathbb{E}[X\Lambda|\mathcal{G}]}{\mathbb{E}[\Lambda|\mathcal{G}]}. \tag{A.13}$$

Proof. Let $S := \{\omega : \mathbb{E}[\Lambda|\mathcal{G}](\omega) = 0\}$. Then $S \in \mathcal{G}$ and so by definition of conditional expectation

$$\mathbb{Q}(S) = \int_S d\mathbb{Q} = \int_S \Lambda d\mathbb{P} = \int_S \mathbb{E}[\Lambda|\mathcal{G}] d\mathbb{P} = \int_S 0\, d\mathbb{P} = 0.$$

Thus $\mathbb{Q}$-a.s. we have $\mathbb{E}[\Lambda|\mathcal{G}](\omega) > 0$.

To prove the second claim assume first that $X \geq 0$. We note that by definition of conditional expectation, for all $G \in \mathcal{G}$:

$$\int_G \mathbb{E}[X\Lambda|\mathcal{G}] d\mathbb{P} = \int_G X\Lambda d\mathbb{P} = \int_G X d\mathbb{Q} = \int_G \mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}] d\mathbb{Q} = \int_G \mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\Lambda d\mathbb{P}.$$

Now we use the definition of conditional expectation to take *another* conditional expectation with respect to $\mathcal{G}$. Since $G \in \mathcal{G}$:

$$\int_G \mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\Lambda d\mathbb{P} = \int_G \mathbb{E}\left[\mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\Lambda|\mathcal{G}\right] d\mathbb{P}.$$

But $\mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable and so

$$\int_G \mathbb{E}\left[\mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\Lambda|\mathcal{G}\right] d\mathbb{P} = \int_G \mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\mathbb{E}\left[\Lambda|\mathcal{G}\right] d\mathbb{P}.$$

Thus, since in particular $\Omega \in \mathcal{G}$, we get

$$\int_\Omega \mathbb{E}[X\Lambda|\mathcal{G}] d\mathbb{P} = \int_\Omega \mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\mathbb{E}\left[\Lambda|\mathcal{G}\right] d\mathbb{P}.$$

Since $X \geq 0$ (and $\Lambda \geq 0$) this means that $\mathbb{P}$-a.s. and hence $\mathbb{Q}$-a.s. we have (A.13).

$$\mathbb{E}[X\Lambda|\mathcal{G}] = \mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\mathbb{E}\left[\Lambda|\mathcal{G}\right].$$

For a general X write $X = X^+ - X^-$, where $X^+ = \mathbb{1}_{\{X\geq 0\}}X \geq 0$ and $X^- = -\mathbb{1}_{\{X<0\}}X \geq 0$. Then

$$\mathbb{E}^{\mathbb{Q}}[X^+ - X^-|\mathcal{G}] = \frac{\mathbb{E}[X^+\Lambda|\mathcal{G}]}{\mathbb{E}[\Lambda|\mathcal{G}]} - \frac{\mathbb{E}[X^-\Lambda|\mathcal{G}]}{\mathbb{E}[\Lambda|\mathcal{G}]} = \frac{\mathbb{E}[X^+ - X^-\Lambda|\mathcal{G}]}{\mathbb{E}[\Lambda|\mathcal{G}]}.$$

□

A.3.4 Multivariate normal distribution

There are a number of ways how to define a multivariate normal distribution. See e.g. [7, Chapter 5] for a more definite treatment. We will define a multivariate normal distribution as follows. Let $\mu \in \mathbb{R}^d$ be given and let Σ be a given symmetric, invertible, positive definite $d\times d$ matrix (it is also possible to consider positive semi-definite matrix Σ but for simplicity we ignore that situation here).

A matrix is positive definite if, for any $x \in \mathbb{R}^d$ such that $x \neq 0$, the inequality $x^T\Sigma x > 0$ holds. From linear algebra we know that this is equivalent to:

1. The eigenvalues of the matrix Σ are all positive.
2. There is a unique (up to multiplication by -1) matrix B such that $BB^T = \Sigma$.

Let B be a $d \times k$ matrix such that $BB^T = \Sigma$.

Let $(X_i)_{i=1}^d$ be independent random variables with $N(0,1)$ distribution. Let $X = (X_1, \ldots, X_d)^T$ and $Z := \mu + BX$. We then say $Z \sim N(\mu, \Sigma)$ and call Σ the covariance matrix of Z.

Exercise A.50. Show that $\text{Cov}(Z_i, Z_j) = \mathbb{E}((Z_i - \mathbb{E}Z_i)(Z_j - \mathbb{E}Z_j)) = \Sigma_{ij}$. This justifies the name "covariance matrix" for Σ.

It is possible to show that the density function of $N(\mu, \Sigma)$ is

$$f(x) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}((x-\mu)^T\Sigma^{-1}(x-\mu))\right). \tag{A.14}$$

Note that if Σ is symmetric and invertible then Σ^{-1} is also symmetric.

Exercise A.51. You will show that $Z = BX$ defined above has the density f given by (A.14) if $\mu = 0$.

i) Show that the characteristic function of $Y \sim N(0,1)$ is $t \mapsto \exp(-t^2/2)$. In other words, show that $\mathbb{E}(e^{itY}) = \exp(-t^2/2)$. *Hint.* complete the squares.

ii) Show that the characteristic function of a random variable Y with density f given by (A.14) is

$$\mathbb{E}\left(e^{i(\Sigma^{-1}\xi)^TY}\right) = \exp\left(-\frac{1}{2}\xi^T\Sigma^{-1}\xi\right).$$

By taking $y = \Sigma^{-1}\xi$ conclude that

$$\mathbb{E}\left(e^{iy^TY}\right) = \exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right).$$

Hint. use a similar trick to completing squares. You can use the fact that since Σ^{-1} is symmetric $\xi^T\Sigma^{-1}x = (\Sigma^{-1}\xi)^Tx$.

iii) Recall that two distributions are identiacal if and only if their characteristic functions are identical. Compute $\mathbb{E}\left(e^{iy^T Z}\right)$ for $Z = BX$ and $X = (X_1, \ldots, X_d)^T$ with $(X_i)_{i=1}^d$ independent random variables such that $X_i \sim N(0,1)$. Hence conclude that Z has density given by (A.14) with $\mu = 0$.

You can now also try to show that all this works with $\mu \neq 0$.

A.3.5 Stochastic Analysis Details

The aim of this section is to collect technical details in stochastic analysis needed to make the main part of the notes correct but perhaps too technical to be of interest to many readers.

Definition A.52. We say that a process X is called *progressively measurable* if the function $(\omega, t) \mapsto X(\omega, t)$ is measurable with respect to $\mathcal{F}_t \otimes \mathcal{B}([0,t])$ for all $t \in [0,T]$.

We will use Prog_T to denote the σ-algebra generated by all the progressively measurable processes on $\Omega \times [0,T]$.

If X is progressively measurable then the processes $\left(\int_0^t X(s)ds\right)_{t\in[0,T]}$ and $(X(t \wedge \tau))_{t\in[0,T]}$ are adapted (provided the paths of X are Lebesgue integrable and provided τ is a stopping time). The important thing for us is that any left (or right) continuous adapted process is progressively measurable.

A.3.6 More Exercises

Exercise A.53. Say $g : \mathbb{R} \to \mathbb{R}$ is bounded, measurable with compact support and $W = (W_t)_{t\in[0,T]}$ is a Wiener process.

i) Show that if $X, Y \in \mathcal{H}$ then

$$\mathbb{E}\int_0^T X_s Y_s\,ds = \mathbb{E}\left[\left(\int_0^T X_s\,dW_s\right)\left(\int_0^T Y_s\,dW_s\right)\right].$$

ii) Assume that $u \in C_b^{1,2}([0,T)\times\mathbb{R})$, so u, $\partial_t u$, $\partial_x u$, Δu are continuous and bounded, satisfies

$$\begin{aligned} \partial_t u + \frac{1}{2}\Delta u = 0 \ \text{ on } \ [0,T)\times\mathbb{R} \\ u(T,\cdot) = g \ \text{ on } \ \mathbb{R}\,. \end{aligned}$$

Show that

$$\partial_x u(t,x) = \mathbb{E}\left[\frac{g(x+W_{T-t})}{T-t}\int_t^T dW_s\right].$$

iii) Hence calculate

$$\mathbb{E}\left[g(W_T)W_T\right].$$

A.4 Solutions to the exercises

Solution (to Exercise A.53). i) Due to Itô's isometry we have

$$
\begin{aligned}
&\mathbb{E}\Big[\int_0^T X_s^2\,ds + 2\int_0^T X_sY_s\,ds + \int_0^T Y_s^2\,ds\Big] = \mathbb{E}\Big[\int_0^T (X_s+Y_s)^2\,ds\Big]\\
&= \mathbb{E}\Big[\Big(\int_0^T (X_s+Y_s)\,dW_s\Big)^2\Big]\\
&= \mathbb{E}\Big[\Big(\int_0^T X_s\,dW_s\Big)^2 + 2\Big(\int_0^T X_s\,dW_s\Big)\Big(\int_0^T Y_s\,dW_s\Big) + \Big(\int_0^T Y_s\,dW_s\Big)^2\Big].
\end{aligned}
$$

Due to Itô's isometry again we have $\mathbb{E}\big[\big(\int_0^T X_s\,dW_s\big)^2\big] = \mathbb{E}\big[\int_0^T X_s^2\,ds\big]$ and similar identity with Y. Hence, cancelling equal terms above, we get

$$
\mathbb{E}\Big[2\int_0^T X_sY_s\,ds\Big] = \mathbb{E}\Big[2\Big(\int_0^T X_s\,dW_s\Big)\Big(\int_0^T Y_s\,dW_s\Big)\Big].
$$

ii) Due to Itô's formula applied to u and the process $X_s^{t,x} = x + W_{s-t}$ we have for any $t \le t' \le T$ that

$$
\begin{aligned}
u(T, X_T^{t,x}) &= u(t', X_{t'}^{t,x}) + \int_{t'}^T \Big[\partial_t u + \frac{1}{2}\Delta u\Big](s, X_s^{t,x})\,ds + \int_{t'}^T \partial_x u(s, X_s^{t,x})\,dW_s\\
&= u(t', X_{t'}^{t,x}) + \int_{t'}^T \partial_x u(s, X_s^{t,x})\,dW_s\,.
\end{aligned}
\tag{A.15}
$$

Since the derivative is bounded the stochastic integral is a martingale and so

$$
\mathbb{E}u(T, X_T^{t,x}) = \mathbb{E}u(t', x_{t'}^{t,x})\,,\;\; t \le t' \le T\,.
$$

Hence $u(t,x) = \mathbb{E}[g(X_T^{t,x})]$ and moreover

$$
\partial_x u(t,x) = \partial_x \mathbb{E}u(T, X_T^{t,x}) = \partial_x \mathbb{E}u(t', X_{t'}^{t,x})\,,\;\; t \le t' \le T\,.
$$

So $\partial_x \mathbb{E}u(t', X_{t'}^{t,x})$ is constant in $t' \in [t,T]$. Hence (assuming we can interchage derivative and expectation) we have

$$
\nabla_x u(t,x) = \frac{1}{T-t}\int_t^T \mathbb{E}\Big[(\partial_x u)(t', X_{t'}^{t,x})\partial_x X_{t'}^{t,x}\Big]\,dt'\,.
$$

But here $\partial_x X_{t'}^{t,x} = \partial_x\big[x + W_{t-s}\big] = 1$. Then, due to part i), we get

$$
\nabla_x u(t,x) = \frac{1}{T-t}\mathbb{E}\Big[\Big(\int_t^T (\partial_x u)(t', X_{t'}^{t,x})\,dW_{t'}\Big)\Big(\int_t^T 1\,dW_{t'}\Big)\Big].
$$

From (A.15) we then get

$$
\begin{aligned}
\nabla_x u(t,x) &= \frac{1}{T-t}\mathbb{E}\Big[\big(u(T, X_T^{t,x}) - u(t,x)\big)\Big(\int_t^T 1\,dW_{t'}\Big)\Big]\\
&= \mathbb{E}\Big[\frac{u(T, X_T^{t,x})}{T-t}(W_T - W_t)\Big].
\end{aligned}
$$

iii) We see that

$$
\mathbb{E}\Big[g(W_T)W_T\Big] = T\partial_x u(0,0)
$$

and that $u(0,x) = \mathbb{E}\Big[g(x+W_T)\Big]$. We know that $x + W_T \sim N(x,T)$ and so

$$
u(0,x) = \int_{\mathbb{R}} g(y)\frac{1}{\sqrt{2\pi T}}e^{-\frac{|x-y|^2}{2\sqrt{T}}}\,dy\,.
$$

Differentiating we get

$$
\partial_x u(0,x) = -\int_{\mathbb{R}} g(y)\frac{x}{\sqrt{T}}\frac{1}{\sqrt{2\pi T}}e^{-\frac{|x-y|^2}{2\sqrt{T}}}\,dy = -\frac{x}{\sqrt{T}}u(0,x)\,.
$$

So finally

$$
\mathbb{E}\Big[g(W_T)W_T\Big] = T\partial_x u(0,0) = 0\,.
$$

References

[1] C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis.* Springer, Berlin, 2006.

[2] T. Björk. *Arbitrage theory in continuous time.* Oxford University Press, 2009.

[3] R. Carmona and F. Delarue. *Probabilistic theory of mean field games with applications.* Springer, 2018.

[4] A. Cartea, S. Jaimungal, and J. Penalva. *Algorithmic and High-Frequency Trading.* Cambridge University Press, 2015.

[5] L. C. Evans. *Partial differential equations.* American Mathematical Society, 1998.

[6] A. Friedman. *Partial differential equations of parabolic type.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964.

[7] A. Gut. *An intermediate course in probability.* Springer, 1995.

[8] I. Gyöngy and D. Šiška. On randomized stopping. *Bernoulli*, 14(2):352–361, 2008.

[9] J. Harter and A. Richou. A stability approach for solving multidimensional quadratic bsdes. *Electron. J. Probab.*, 24:1–51, 2019.

[10] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus.* Springer, 1991.

[11] B. Kerimkulov, D. Šiška, and L. Szpruch. Exponential convergence and stability of Howards's policy improvement algorithm for controlled diffusions. *SIAM J. Control Optim.*, 58:1314–1340, 2020.

[12] B. Kerimkulov, D. Šiška, and L. Szpruch. A modified MSA for stochastic control problems. *arXiv:2007.05209*, 2020.

[13] A. N. Kolmogorov and S. V. Fomīn. *Introductory real analysis.* Dover Publications, Inc., New York, 1975.

[14] N. V. Krylov. *Controlled Diffusion Processes.* Springer, 1980.

[15] H. Pham. *Continuous-time stochastic control and optimization with financial applications.* Springer, 2009.

[16] M. L. Puterman. *Markov decision processes.* Wiley, 1994.

[17] A. Seierstad. *Stochastic control in discrete and continuous time.* Springer, 2009.

[18] R. S. Sutton and A. G. Barto. *Reinforcement learning.* MIT Press, 2nd edition, 2018.

[19] N. Touzi. Optimal stochastic control, stochastic target problems and backward sde. `http://www.cmap.polytechnique.fr/~touzi/Fields-LN.pdf`, 2010.

[20] C. J. C. H. Watkins and P. Dayan. Q-learning. *Machine Learning*, 8(3–4):279–292, 1992.