Optimal Control

Arash Fahim • •

Last Revision: May 7, 2025

Contents

Notations

$(\Omega, \mathcal{F}, \mathbb{P})$	a probability space with a σ-field $\mathcal{F}$ and probability $\mathcal{P}$
$\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$	a filtration in a probability space
$\mathbb{F}^X$	filtration generated by process X
$\mathbb{F}^{X+}$	$\cup_{s>t}\mathcal{F}_s^X$
$\bar{\mathbb{F}}$	filtration augmented by adding all the $\mathbb{P}$-null events
$(\Omega, \mathbb{F}, \mathbb{P})$	a filtered probability space with right-continuous augmented filtration (usual conditions) that hosts a Brownian motion adapted to the filtration $\mathbb{F}$
$O \subseteq \mathbb{R}^d$	open set
USC(O) (LSC(O))	lower semicontinuous (upper semicontinuous) functions on O
$\mathrm{C}^k(O)$	k times continuously differentiable functions on O
$\mathrm{C}^k(\bar{O})$	functions on $\bar{O}$ k times continuously differentiable on all variables over on O with derivatives continuously extendable to $\bar{O}$
$\mathrm{C}^{k,l,\cdots}(O)$	functions on O k times continuously differentiable on first variable, l times on second variable, ...
$\mathrm{C}^{k,l,\cdots}(\bar{O})$	functions on $\bar{O}$ times continuously differentiable on first variable, l times on second variable, ... on O with derivatives continuously extendable to $\bar{O}$
$M(n, m)$	the set of all real n by m matrices
$A \cdot B := \mathrm{Tr}[AB^T]$	the inner product of two matrices A and B in $M(n, m)$
$\|A\| := A \cdot A = \sum_{i=1}^n \sum_{j=1}^m a_{ij}^2$	l^2-norm of matrices in $M(n, m)$
Ω	the sample space of a random event
ω	is reserved for the members of Ω
$\partial_t V$, $\partial_x V$, $\partial_{xx} V$	Partial derivatives of a function $V : [0, T] \times \mathbb{R}$ once wrt t, once wrt x and twice wrt x
$\partial_t V$, ∇V, $D^2 V$	Partial derivative of a function $V : [0, T] \times \mathbb{R}^2$ once wrt t, gradient of V wrt x and Hessian of V wrt x

Chapter 1

Preliminaries

1.1 Optimization Versus Control

In this chapter, we provide a brief overview of the aspect in which optimization and control are different. We start the chapter with some examples.

Example 1. *We start by a quadratic problem. Let* $\alpha : [0, T] \to\in \mathbb{R}$ *be given.*

$$\inf \left\{ \int_0^T \left(x_t^2 - \alpha_t x_t \right) \mathrm{d}t \right\} \tag{1.1.1}$$

where the infimum is over all functions $x : [0, T] \to\in \mathbb{R}$*. As* x *can be any mapping, one can solve the following problem for each* x_t *separately to obtain* $x_t^* = \alpha_t/2$*.*

$$\inf_{x \in \mathbb{R}^d} \left\{ x^2 - \alpha_t x \right\} \tag{1.1.2}$$

The above problem is a dynamic optimization problem.

Example 2. *We make the above example more complicated by specifying a dynamics for* x_t*:*

$$\inf \left\{ \int_0^T \left(x_t^2 - \alpha_t x_t \right) \mathrm{d}t \right\} \tag{1.1.3}$$

where the infimum is over all functions $x : [0, T] \to\in \mathbb{R}$ *such that for some (Borel measurable) function* $u : [0, T] \to \mathbb{R}$

$$\mathrm{d}x_t = -\beta x_t \mathrm{d}t + u_t \mathrm{d}t, \quad x_0 = x \tag{1.1.4}$$

Note that the existence of the term $u_t\mathrm{d}t$ *is necessary for the meaning of infimum. This problem is not a simple dynamic optimization problem because one can only choose* x *such that the dynamic equation (1.1.4) holds for some function* u*.* ***Therefore,*** *we can only choose* u *and indirectly modify* x_t *to minimize (1.1.3).*

Exercise 1. *Solve (1.1.4) for* x_t *in terms of* u*. Hint:* $\mathrm{d}(e^{\beta t} x_t) = e^{\beta t} u_t \mathrm{d}t$*.*

Exercise 2. *Show that if there exists a function* u *such that* $\alpha_t = 2 \int_0^t e^{\beta(s-t)} u_s \mathrm{d}s$*, then* u *minimizes (1.1.3).*

Exercise 3. *If we modify an optimal control* u *in a countable number of points described in the exercise above, does it remain an optimal control? Does the initial value* $x_0 = x$ *play a role in the problem?*

Exercise 4. *Assume that there exists no function* u *such that* $\alpha_t = 2 \int_0^t e^{-\beta(s-t)} u_s \mathrm{d}s$*. What is the minimum value of (1.1.3)?*

The problem (1.1.3) is a simple control problem. However, after doing the above exercises, you note that it can simply be reduced to an optimization problem. Such a solution for control problems are called myopic solutions and do not necessarily exist for more interesting control problems. Here is an example which does not allow for a myopic solution.

Example 3. *Consider the control problem:*

$$\inf \int_0^T \left(x_t^2 - \alpha_t x_t + u_t^2\right) \mathrm{d}t \tag{1.1.5}$$

where the infimum is over all (Riemann integrable) functions $u : [0,T] \to \mathbb{R}$ *and* $x : [0,T] \to \mathbb{R}$ *and* $u : [0,T] \to \mathbb{R}$ *satisfy (1.1.4). In the above problem, the cost function for the control problem* $C(t,x,u) = x^2 - \alpha_t x + u^2$ *depends on the control* u*. Unlike the previous example, the free choice of* u *induces a new cost that no longer makes* $x_t = \alpha_t/2$ *optimal. Such problems do not have a myopic solution.*

A general control problem is described as

$$\inf_{u \in \mathcal{U}} \int_0^T C(t, x_t, u_t) \mathrm{d}t + g(x_T) \tag{1.1.6}$$

where $\mathrm{d}x_t = f(x_t, u_t)\mathrm{d}t$. The function $C : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$ is called the running cost and $g : \mathbb{R}^d \to \mathbb{R}$ is called the terminal cost. The set $\mathcal{U}$ is the set of all functions $u : [0,T] \to \mathbb{R}^n$, control variable, which we are interested.

Remark 1.1.1. *Important remark on the set of all controls It is crucial to choose the set* $\mathcal{U}$ *of all control variables properly. For instance, if* $\mathcal{U}$ *is the set of all function* $u : [0,T] \to \mathbb{R}$*, then*

$$\inf_{u \in \mathcal{U}} \int_0^T (x_t - u_t^2)\mathrm{d}t, \quad \mathrm{d}x_t = (x_t - u_t)\mathrm{d}t = -\infty \tag{1.1.7}$$

by simply choosing $u_t = -x_t$*. However, if we restrict the set* $\mathcal{U}$ *to function* $u : [0,T] \to [-1,1]$*, then*

$$\inf_{u \in \mathcal{U}} \int_0^T (x_t - u_t^2)\mathrm{d}t, \quad \mathrm{d}x_t = (x_t - u_t)\mathrm{d}t > -\infty. \tag{1.1.8}$$

The suitable set of controls chosen for a specific problem is referred to as admissible control. We denote this set by $\mathcal{U}$*.*

Note that infinite horizon is accommodated by $T = \infty$.

Exercise 5. *Write the following problem as a generic control problems by associating the horizon* T*, the running cost* $C(t,x,u)$ *and terminal cost* $g(x)$ *in (1.1.6):*

(Shortest time to exit a bounded domain) Given a bounded domain $D \subset \mathbb{R}^d$*, find*

$$\inf_u \{t \geq 0 \; : \; x_t \notin D\} \tag{1.1.9}$$

where $\mathrm{d}x_t = u_t \mathrm{d}t$ *with control* $|u_t| \leq 1$ *and* $u_t \in \mathbb{R}^d$ *and initial position* $x_0 = x \in D$*.*

We give a quick review of some of the most important optimization results, including the Lagrange multiplier.

1.2 Lagrange multiplier

If we add a constraint to a simple minimization problem such as $\min_x f(x)$, the Lagrange multiplier method is the way to proceed. In a nutshell, the Lagrange multiplier method turns a constrained optimization problem into a saddle point problem without constraints by adding more variables.

Consider the constrained problem below:

$$\inf_x f(x) \quad \text{subject to} \quad g(x) = 0 \tag{1.2.1}$$

Define the Lagrangian:

$$L(x, \lambda) := f(x) - \lambda \cdot g(x) \tag{1.2.2}$$

Then, under proper conditions, the following saddle point problem yields the solution to (1.2.1).

$$\sup_\lambda \inf_x L(x, \lambda) \tag{1.2.3}$$

The function $f^*(\lambda) = \inf_x L(x, \lambda)$ is called the dual of f, and the Lagrange multiplier λ is called the dual variable. The problem

$$\sup_\lambda f^*(\lambda) \tag{1.2.4}$$

is called the dual problem for the primal problem (1.2.1).

The key to the success of the Lagrange multiplier method is the strong duality.

$$\sup_\lambda \inf_x L(x, \lambda) = \inf_x \sup_\lambda L(x, \lambda) \tag{1.2.5}$$

It is not always true that the strong duality holds. However, Karush-Kuhn-Tucker conditions (KKT) provide some necessary and sufficient conditions for strong duality.

Theorem 1.2.1. *Assume the differentiability of f and g. If x^* solves (1.2.1) and λ^* solves (1.2.4) such that*

$$\begin{cases} \nabla f(x^*) - \lambda^* \cdot \nabla g(x^*) = 0 \\ g(x^*) = 0 \end{cases} \tag{1.2.6}$$

then, strong duality holds and (x^*, λ^*) is a saddle point for $\sup_\lambda \inf_x L(x, \lambda)$.

Conversely, if strong duality holds, then any saddle point (x^*, λ^*) satisfies (1.2.6). In particular, x^* solve (1.2.1).

A geometric interpretation of λ^* in KKT conditions is explained in See Figure 1.2.1. Note that (x^*, λ^*) constitutes a saddle point for $L(x, \lambda)$.

When the constraint is given by some inequalities and qualities, i.e.,

$$\inf_x f(x) \quad \text{subject to} \quad g(x) = 0, \quad h(x) \geq 0 \tag{1.2.7}$$

the Lagrangian is similar, but the dual variable in the Lagrangian is different.

$$L(x, \lambda, \mu) := f(x) - \lambda \cdot g(x) - \mu \cdot h(x) \tag{1.2.8}$$

$$\sup_{\mu \geq 0} \sup_\lambda \inf_x L(x, \lambda, \mu) \tag{1.2.9}$$

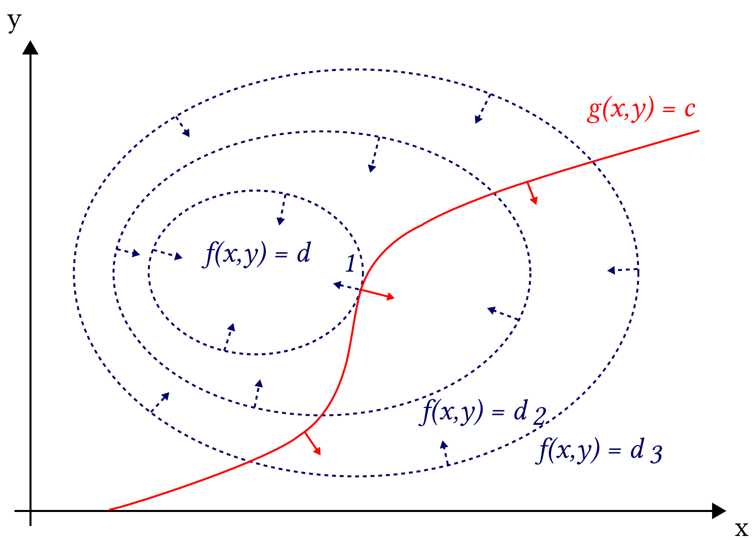

Figure 1.2.1: KKT conditions for $g(x) = 0$. Source WIKIPEDIA

The reason for such modification can formally be seen after switching $\inf_x$ and $\sup_{\mu\geq 0}$ in the above

$$\sup_{\mu\geq 0}\inf_x L(x,\lambda,\mu) = \inf_x \sup_{\mu\geq 0} L(x,\lambda,\mu) \tag{1.2.10}$$

If x_1 is such that $h(x_1) < 0$, then

$$\sup_{\mu\geq 0} L(x_1,\lambda,\mu) = f(x_1) - \lambda g(x_1) - h(x_1)\sup_{\mu\geq 0}\mu = \infty$$

Therefore, $\inf_x \sup_{\mu\geq 0} L(x,\lambda,\mu)$ is not attained at x_1. Thus, any saddle point (x^*,λ^*,μ^*) with $\mu^* > 0$ must satisfy $h(x^*) \geq 0$. More general KKT conditions guarantee the strong duality in this case:

$$\begin{cases} \nabla f(x^*) - \lambda^*\nabla g(x^*) - \mu^*\nabla h(x^*) = 0 \\ g(x^*) = 0 \\ h(x^*) \geq 0 \\ \mu^* \geq 0 \\ \mu^*\cdot h(x^*) = 0 \end{cases} \tag{1.2.11}$$

The last equality emphasizes that either $h(x^*) > 0$ holds, in which case $\mu^* = 0$, or $h^*(x^*) = 0$, in which case μ^* is irrelevant. See Figure 1.2.2.

1.2.1 Lagrange multiplier and constrained optimal control

Let us put the Lagrange multiplier in the context of an optimization problem:

$$\inf\left\{\int_0^T \left(x_t^2 - \alpha_t x_t\right)\mathrm{d}t\right\} \tag{1.2.12}$$

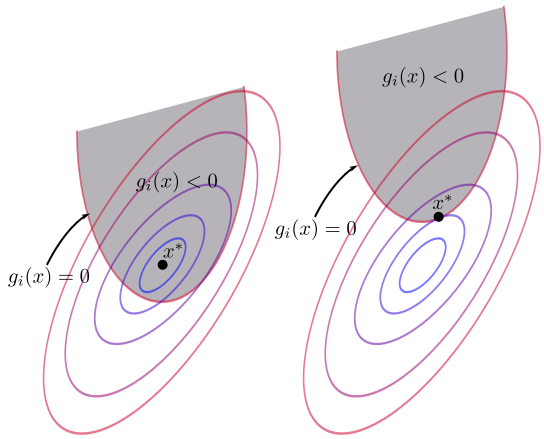

Figure 1.2.2: KKT conditions for $h(x) \geq 0$. The figure uses g for h. Source WIKIPEDIA

subject to $x_T = 0$. The Lagrangian simple is

$$\int_0^T \Big(x_t^2 - \alpha_t x_t\Big)\mathrm{d}t + \lambda x_T \tag{1.2.13}$$

and KKT condition suggests that $x_t^* = 2\alpha_t$ for $t < T$ and $x_T^* = 0$ solves the problem. However, this is not an interesting problem. One can simply argue that changing x at T does not change the value of the integral, and therefore it is not really a constraint. Even if we impose a more restricted constraint such as $x_t \geq 0$ on all t, the solution will simply be $x_t^* = \max\{\alpha_t, 0\}/2$.

However, if the constraint is something as $\int_0^T x_t \mathrm{d}t = 0$, then the optimization problem becomes more interesting.

$$\int_0^T \Big(x_t^2 - \alpha_t x_t + \lambda x_t\Big)\mathrm{d}t \sup_\lambda \inf_x \tag{1.2.14}$$

KKT condition becomes $x_t^* = (\alpha_t - \lambda^*)/2$ and $\int_0^T x_t^* = 0$. This implies that $\lambda^* = \frac{1}{T}\int_0^T \alpha_t \mathrm{d}t$, and therefore $x_t^* = (\alpha_t - \frac{1}{T}\int_0^T \alpha_t \mathrm{d}t)/2$.

Adding restrictions to a control problem is slightly more subtle due to the dynamics of x. We postpone the study of such problems to the future endeavors. However, a simple control problem can be described as a constrained optimization and can be solved via the Lagrange multiplier. We discuss this approach in the next section.

1.3 Solution methods for deterministic optimal control

In this section, we propose methods for the optimal control problems. There are two groups of numerical methods, first group are based on the dynamic programming principle (DPP). DPP provides a backward recursive method to solve an optimal control problem. The second group completely avoids DPP and formulates problem as an optimization.

1.3.1 Dynamic programming principle (DPP)

Consider the optimal control problem (1.1.6):

$$\inf_u \int_0^T C(t, x_t, u_t)\mathrm{d}t + g(x_T), \quad \mathrm{d}x_t = f(x_t, u_t)\mathrm{d}t \tag{1.3.1}$$

To explain DPP, we first define the value function for (1.1.6).

Definition 1.3.1. *Let $x_t = x$. Then, the value function of (1.1.6) is defined by*

$$V(t, x) := \inf_{u \in \mathcal{U}} \int_t^T C(s, x_s, u_s)\mathrm{d}s + g(x_T), \quad \mathrm{d}x_s = f(x_s, u_s)\mathrm{d}s \tag{1.3.2}$$

and u belongs to the set of controls restricted to interval $[t, T]$.

1.3.2 Pontryagin principle

Consider the simple control problem in Section 1.1:

$$\inf_u \left\{ \int_0^T \left(x_t^2 - \alpha_t x_t \right) \mathrm{d}t \right\} \tag{1.3.3}$$

where $\mathrm{d}x_t = -\beta x_t \mathrm{d}t + u_t \mathrm{d}t$, $x_0 = x$. One can consider the dynamics of x as a constraint. Therefore, we can formally write Lagrangian by

$$\sup_\lambda \inf_{u,x} \left\{ \int_0^T (x_t^2 - \alpha_t x_t)\mathrm{d}t - \int_0^T \lambda_t (\mathrm{d}x_t + (\beta x_t - u_t)\mathrm{d}t) \right\} \tag{1.3.4}$$

Here, there are two important remarks. First, since the constraint is given by a differential equation for each t, the dual variable λ is a function of t. Second, the problem is now unconstrained, that is, the state variable x and the control variable u are now both variables in an optimization problem.

The main trick to solve this optimization problem is integration by part formula

$$\int_0^T \lambda_t \mathrm{d}x_t = \lambda_T x_T - \lambda_0 x_0 - \int_0^T x_t \mathrm{d}\lambda_t \tag{1.3.5}$$

to write (1.3.4) as

$$\sup_\lambda \inf_{u,x} \left\{ \int_0^T \left(x_t \mathrm{d}\lambda_t - \lambda_t(\beta x_t - u_t) + x_t^2 - \alpha_t x_t \right) \mathrm{d}t - \lambda_T x_T + \lambda_0 x_0 \right\} \tag{1.3.6}$$

Note that optimization on x_T is independent of x_t for $t < T$. The KKT conditions for the strong duality in the saddle point problem (1.3.6) were discovered by Lev Pontryagin in 1952. Define the Hamiltonian by

$$H(t, x, \lambda, u) := -\lambda(\beta x - u) + x^2 - \alpha_t x \tag{1.3.7}$$

Thus, (1.3.6) is written as a saddle point problem with free variables x_t, u_t, and λ_t:

$$\sup_\lambda \inf_{u,x} \left\{ \int_0^T \left(x_t \mathrm{d}\lambda_t + H(t, x_t, \lambda_t, u_t) \right) \mathrm{d}t - \lambda_T x_T + \lambda_0 x_0 \right\} \tag{1.3.8}$$

$$\begin{cases} \mathrm{d}\lambda_t^* + \partial_x H(t, x_t^*, \lambda_t^*, u_t^*)\mathrm{d}t = \mathrm{d}\lambda_t^* + (-\lambda_t^*\beta + 2x_t^* - \alpha_t)\mathrm{d}t = 0 \text{ (minimize integrand wrt } x) \\ \lambda_T^* = 0 \text{ (minimize terminal wrt } x_T) \\ H(t, x_t^*, \lambda_t^*, u_t^*) \leq H(t, x_t^*, \lambda_t^*, u) \text{ for all } u \text{ (minimize integrand wrt } u) \\ \mathrm{d}x_t^* = (-\beta x_t^* + u_t^*)\mathrm{d}t \text{ (constraint)} \end{cases} \tag{1.3.9}$$

In the above, the first equation is obtained from taking derivative with respect to x_t, second equality corresponds to derivative with respect to x_T in $\lambda_T x_T$, the third line guarantees the optimality of u^*, and the last equation is the constraint of the problem which is the dynamic of the state variable in the control.

Exercise 6. *Show that $\lambda_t^* = 0$, $x_t^* = \alpha_t/2$, and u^* with $\alpha_t = 2\int_0^t e^{-\beta(s-t)} u_s^* \mathrm{d}s$ satisfy Pontryagin principle for the above problem.*

Following the steps of in the above example, the Pontryagin principle for a generic deterministic control problem, 1.1.6 is described as the following saddle point problem.

$$\sup_{\lambda} \inf_{u,x} \left\{ \int_0^T C(t, x_t, u_t)\mathrm{d}t + g(x_T) - \int_0^T \lambda_t \big(\mathrm{d}x_t + (\beta x_t - u_t)\mathrm{d}t\big) \right\} \tag{1.3.10}$$

By applying (1.3.5), we obtain

$$\sup_{\lambda} \inf_{u,x} \left\{ \int_0^T \Big(x_t \mathrm{d}\lambda_t + H(t, x_t, \lambda_t, u_t)\Big)\mathrm{d}t + g(x_T) - \lambda_T x_T + \lambda_0 x_0 \right\} \tag{1.3.11}$$

where the Hamiltonian is given by

$$H(t, x, \lambda, u) := \lambda f(t, x, u) + C(t, x, u) \tag{1.3.12}$$

Theorem 1.3.1. *[Pontryagin principle] Assume that there exists $(x^*, u^*, \lambda^*) : [0, T] \to \mathbb{R}^d \times \mathbb{R}^n \times \mathbb{R}^d$ such that*

$$\begin{cases} \mathrm{d}\lambda_t^* + \partial_x H(t, x_t^*, \lambda_t^*, u_t^*)\mathrm{d}t = 0 \text{ (minimize integrand wrt } x) \\ \lambda_T^* = \nabla g(x_T^*) \text{ (minimize terminal wrt } x_T) \\ H(t, x_t^*, \lambda_t^*, u_t^*) \leq H(t, x_t^*, \lambda_t^*, u) \text{ for all } u \text{ (minimize integrand wrt } u) \\ \mathrm{d}x_t^* = f(t, x_t^*, u_t^*)\mathrm{d}t \text{ (constraint)} \end{cases} \tag{1.3.13}$$

Then, u^* is an optimal control for (1.1.6).

The function λ_t^*, described by 1.3.1 is called the adjoint process. It is important to distinguish between the two ODEs in (1.3.13), namely,

$$\mathrm{d}x_t^* = f(t, x_t^*, u_t^*)\mathrm{d}t \text{ and } \mathrm{d}\lambda_t^* + \partial_x H(t, x_t^*, \lambda_t^*, u_t^*)\mathrm{d}t = 0 \tag{1.3.14}$$

The first one has an initial condition $x_0^* = x_0$, the initial position of the state process. However, the adjoint equation comes with a terminal condition, $\lambda_T^* = \nabla g(x_T^*)$. This makes solving Theorem **??** challenging. We shall see later how Pontryagin principle is applied to some specific examples such as linear quadratic linear control problem, where solving (1.3.13) is simpler to solve.

Remark 1.3.1. *If u_t^* is an interior minimizer of $H(t, x_t^*, \lambda_t^*, u)$, then we can write $\partial_u H(t, x_t^*, \lambda_t^*, u_t^*) = 0$. However, if there are constraint on u, e.g., $u \geq 0$, we shall stick to the inequality above.*

1.4 Linear quadratic linear control problem (LQC)

In LQC, the running cost and the terminal cost are quadratic functions of state and control and the ODE is linear in the state and control. We start with an example of a LQC problem.

Example 4. *Consider the control problem with $C(x,u) = x^2 + u^2$, $g(x) = x^2 + x$, and $f(x,u) = x + u$:*

$$\inf_u \int_0^T (x_t^2 + u_t^2)\mathrm{d}t + x_T^2 + x_T, \quad \textit{subject to } \mathrm{d}x_t = (x_t + u_t)\mathrm{d}t \tag{1.4.1}$$

LQC problems can be solved via Pontryagin principle. More precisely, for the above example, the Hamiltonian is given by

$$H(x,\lambda,u) = x^2 + u^2 + \lambda(x+u)$$

and (1.3.13) is

$$\begin{cases} \mathrm{d}\lambda_t^* + (2x_t^* + \lambda_t^*)\mathrm{d}t = 0 \\ \lambda_T^* = 2x_T^* + 1 \\ H(t, x_t^*, \lambda_t^*, u_t^*) \leq H(t, x_t^*, \lambda_t^*, u) \\ \mathrm{d}x_t^* = (x_t^* + u_t^*)\mathrm{d}t \end{cases} \tag{1.4.2}$$

Minimizing $H(x,\lambda,u)$ in u suggests that $H_u(x, \lambda_t^*, u_t^*) = 2u_t^* + \lambda_t^* = 0$, equivalently, $u_t^* = -\frac{1}{2}\lambda_t^*$.

Exercise 7. *(1) Use $u_t^* = -\frac{1}{2}\lambda_t^*$ to write the system of ODEs (1.4.2) for λ^* and x^* by*

$$\begin{cases} \mathrm{d}\lambda_t^* = (-\lambda_t^* - 2x_t^*)\mathrm{d}t \\ \mathrm{d}x_t^* = (-\frac{1}{2}\lambda_t^* + x_t^*)\mathrm{d}t \end{cases} \tag{1.4.3}$$

(2) Find the general solution the system of ODEs in terms of x_0^* and λ_0^*. Hint: Use denationalization of the matrix

$$\begin{bmatrix} -1 & -2 \\ -\frac{1}{2} & 1 \end{bmatrix} = \begin{bmatrix} 2(\sqrt{2}+1) & 1 \\ 1 & 2(\sqrt{2}-1) \end{bmatrix} \begin{bmatrix} -\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{2}{3}(\sqrt{2}-1) & -1 \\ -1 & \frac{2}{3}(\sqrt{2}+1) \end{bmatrix} \tag{1.4.4}$$

and write the ODEs as

$$\begin{bmatrix} \mathrm{d}\lambda^* \\ \mathrm{d}x^* \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \lambda^* \\ x^* \end{bmatrix} = \begin{bmatrix} 2(\sqrt{2}+1) & 1 \\ 1 & 2(\sqrt{2}-1) \end{bmatrix} \begin{bmatrix} -\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{2}{3}(\sqrt{2}-1) & -1 \\ -1 & \frac{2}{3}(\sqrt{2}+1) \end{bmatrix} \begin{bmatrix} \lambda^* \\ x^* \end{bmatrix} \tag{1.4.5}$$

(3) Consider x_0^* given. Use $\lambda_T^* = 2x_T^* + 1$ to find λ_0^*, hence a special solution for the system of ODEs from Pontryagin principle as a function x_0^* and x_T^*.

The generic linear quadratic control problem (LQC) is given by

$$\begin{aligned} \inf_u \int_0^T (x_t \cdot A_t x_t + x_t \cdot C_t u_t + u_t \cdot B_t u_t + a_t \cdot x_t + b_t \cdot u_t)\mathrm{d}t + x_T \cdot A_T x_T + a_T \cdot x_T \\ dx_t = (M_t x_t + N_t u_t)dt \end{aligned} \tag{1.4.6}$$

subject to $\mathrm{d}x_t = (C_t \cdot x_t + E_t \cdot u_t)\mathrm{d}t$ where $A : [0,T] \to \mathbb{M}(d,d)$, $B : [0,T] \to \mathbb{M}(m,m)$, $C : [0,T] \to \mathbb{M}(d,m)$, $a : [0,T] \to \mathbb{R}^d$, $b : [0,T] \to \mathbb{R}^n$, $M : [0,T] \to \mathbb{M}(d,d)$, and $N : [0,T] \to \mathbb{M}(d,m)$ are all given functions. Both the running cost and the terminal cost have to be convex functions, otherwise, the value of the LQC problem is simply $-\infty$.

Exercise 8. *In Example 4, change the running cost function to $C(x, u) = x^2 - u^2$. Make some effort to find a solution to this modified problem. Of course, since the cost function is not convex, your effort fails. Explain why the optimal control problem is ill-defined.*

Exercise 9. *Assume that the cost functions of the general LQC problem (1.4.6) are convex. Write Pontryagin principle for the general LQC and suggest a solution method.*

Chapter 2

Stochastic Control

We explain the components of a stochastic control problem.

2.1 State process

Given a progressively measurable process[1] u_t taking values in a set $\mathbf{U} \subseteq \mathbb{R}^m$, which will be clarified later, the state process is given by the SDE

$$\begin{cases} \mathrm{d}x_t^u = \mu(t, X_t^u, u_t)\mathrm{d}t + \sigma(t, X_t^u, u_t)\mathrm{d}B_t \\ X_0^u = x \in \mathbb{R}^d \end{cases} \tag{2.1.1}$$

Remark 2.1.1. *If we assume that $\mu(t,x,u)$ and $\sigma(t,x,u)$ are Lipschitz continuous in (t,x) uniformly in u, i.e.,*

$$\sup_{u\in\mathbf{U}}\{|\mu(t,x,u)-\mu(s,y,u)| + |\sigma(t,x,u)-\sigma(s,y,u)|\} \leq K(|t-s|+|x-y|) \tag{2.1.2}$$

Then for any progressively measurable u_t, (2.1.1) has a strong solution.

Remark 2.1.2. *In the deterministic control, $\sigma \equiv 0$.*

Example 5. *Take $u \equiv 0$ in $\mathrm{d}X_t^u = u_t X_t^u \mathrm{d}t$. Then, $X_t^0 = X_0$. For $u \equiv 1$, $X_t^1 = X_0 e^t$ and For $u_t = -t$, $X_t^u = X_0 e^{-t^2/2}$. In general, $X_t^u = X_0 e^{\int_0^t u_s \mathrm{d}s}$.*

Example 6. *In Black-Scholes model, the price of an asset satisfies $\mathrm{d}s_t = S_t(\mu \mathrm{d}t + \sigma \mathrm{d}B_t)$. If the interest rate $r = 0$ and u_t is the amount of money invested in the asset, the wealth process from a self-financing portfolio satisfies*

$$\mathrm{d}X_t^u = u_t(\mu\mathrm{d}t + \sigma\mathrm{d}B_t) \tag{2.1.3}$$

Then, $X_t^u = X_0 \exp\left(\sigma \int_0^t u_s \mathrm{d}B_s + (\mu - \sigma^2/2)\int_0^t u_s^2 \mathrm{d}s\right)$, where X_0 is the initial wealth.

2.2 Objective function

The goal is to minimize (or maximize) an objective function of the following form. Set $X_0^u = x$ for all u and define

$$J(x;u) = \mathbb{E}\left[\int_0^T C(s, X_s^u, u_s)\mathrm{d}s + g(X_T^u)\right]. \tag{2.2.1}$$

[1] A process is called progressively measurable if for any t, the restriction to $[0,t]\times\Omega$ is $\mathcal{B}[0,t]\otimes\mathcal{F}_t$-measurable.

where $C : [0, T] \times \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}$ is called running cost and $g : \mathbb{R}^d \to \mathbb{R}$ is the terminal cost. We assume that $C(t, x, u)$ is Lipschitz in (t, x) uniformly in u and g is Lipschitz. Here, we only focus on minimization as minimization can be obtained by a negative sign.

The value of the problem is

$$V := \inf_u J(x; u). \tag{2.2.2}$$

If a process u^* exists such that $V = J(x; u^*)$, we call u^* and optimal control.

2.3 Admissibility

It is easy to model stochastic control problems from real applications. However, if not carefully done, the stochastic control problem becomes degenerate.

Exercise 10 (St. Peterburg paradox)**.** *Consider the wealth process as described by* (6) *and assume that asset price S is a martingale by setting $\mu = 0$. In this case, $X_t = X_0 + \sigma \int_0^t u_s \mathrm{d}B_s$. Further, assume $X_0 = 0$. Let's try to maximize expected value of wealth by applying an investment strategy u:*

$$V = \sup_u \mathbb{E}[X_T^u] = \sigma \sup_u \mathbb{E}\left[\int_0^T u_s \mathrm{d}B_s\right]. \tag{2.3.1}$$

Is the stochastic integral above zero? Why?

Now, consider the strategy that chooses $u_t = u > 0$, where u is large and liquidate the investment as soon as $X_t^u = M$. If X_t^u never hits M, the investment yields X_T^u.

Note that for $X_t^u = \sigma u B_t$, probability of hitting M before T is

$$\mathbb{P}\big(\max_{t \le T} X_t^u \ge M\big) = \mathbb{P}\left(\max_{t \le T} B_t \ge \frac{M}{\sigma u}\right) = 2\mathbb{P}\left(B_T \ge \frac{M}{\sigma u}\right) \tag{2.3.2}$$

The last equality is coming from the Schwartz reflection principle for Brownian motion. If we send $u \to \infty$ and $M \to \infty$ such that $\frac{M}{u} \to 0$, we obtain $\mathbb{P}(B_T \ge \frac{M}{\sigma u}) \to \frac{1}{2}$.

On the other hand, from the strategy described above, we have

$$\mathbb{E}[X_T^u] = \mathbb{E}[M 1_{\{\max_{t \le T} X_t^u \ge M\}}] + \mathbb{E}[X_T^u 1_{\{\max_{t \le T} X_t^u < M\}}] = 2M\mathbb{P}\left(B_T \ge \frac{M}{\sigma u}\right) + \mathbb{E}[X_T^u 1_{\{\max_{t \le T} X_t^u < M\}}] \tag{2.3.3}$$

Now, answer the following questions:

1. *Is it true that $\mathbb{E}[X_t^u 1_{\{\max_{t \le T} X_t^u < M\}}] \to 0$ as $u \to \infty$ and $M \to \infty$ such that $\frac{M}{u} \to 0$?*
2. *What is the limit of $2M\mathbb{P}(B_T \ge \frac{M}{\sigma u})$? $V = \infty$?*
3. *If the stochastic integral in* (2.3.1) *is zero, then $V = 0$, why did we also get $V = \infty$?*

The above example shows that if someone can have unlimited borrowing power, the value functions is infinite. From the practical point of view, this is not possible. If we define a set admissible controls to be all progressively measurable u such that $c_t \ge 0$ and $X_T^u \ge -C$ for some credit limit $C \ge 0$[2] $\mathbb{P}$-a.s., then the strategies in Example 10 are not admissible.

Another consideration is that the choice of control u should be made such that the state process in (2.1.1) admits a solution for which $J(x; u)$ is defined and finite. For instance, in Example 10, there are choices

[2] $C \ge 0$ is no-short-selling condition.

for θ_t such that $\mathrm{d}X_t = \sigma\theta_t \mathrm{d}B_t$ does not have a (strong) solution, see [10] for an example when strong solution does not exist. The set of admissible controls should be chosen such that it includes the (unknown) optimal control we are looking for. If too restricted, we would not get optimality; if two wide, we get bizaar solutions such as in St. Petersburg paradox, 10.

Based on the above discussion, we define the set of admissible controls for 10 can be

$$\mathcal{A}_C = \Big\{u\ : \mathbb{P}\text{-a.s., } u \text{ is progressively measurable, SDE } \mathrm{d}X_t^u = \sigma u_t \mathrm{d}B_t \text{ has a "strong" solution } X_t^u, \\ \text{and } X_T^u \geq -C\ \mathbb{P}\text{-a.s.}\Big\}. \tag{2.3.4}$$

In the above, $C \geq 0$ is a given constant. Because $u \equiv 0$ is an optimal control for the St. Petersburg problem, we can also choose $\mathcal{A}_0$ as the set of admissible strategies.

Next, we provide a more practical problem.

Exercise 11 (Merton problem)**.** *In the same setting as 10, consider*

$$V(x) = \sup_u \mathbb{E}[U(X_T^u)|X_0^u = x] \tag{2.3.5}$$

where $U : \mathbb{R} \to \mathbb{R}$ is a differentiable strictly concave strictly increasing function such that $U'(\infty) = 0$, e.g., $U(x) = 1 - e^{-\alpha x}$, $U(x) = \frac{\ln \alpha x}{\alpha}$, or $U(x) = \frac{x^{1-\alpha}}{1-\alpha}$ for $\alpha \in (0, 1)$. Does the same paradox as in St. Petersburg occurs here? Does the set of admissible controls $\mathcal{A}_C$ solves the issue?

2.4 Value function

Assume that we have a generic control problem

$$\inf_{u\in\mathcal{A}} J(u), \quad J(u) := \mathbb{E}\left[\int_0^T C(s, X_s^u, u_s)\mathrm{d}s + g(X_T^u)\right]. \tag{2.4.1}$$

where and X_t^u satisfies (2.1.1) and $\mathcal{A}$ is a set of admissible controls, which we assume given.

One way to solve this problem is to define a value function and obtain a dynamic programming equation on the value function.

Definition 2.4.1 (Value function)**.** *The value function of the control problem is defined by*

$$V(t, \omega) = \mathit{essinf}_{u\in\mathcal{A}_{t,T}} \mathbb{E}\left[\int_t^T C(s, X_s^u, u_s)\mathrm{d}s + g(X_T^u)\Big|\mathcal{F}_t\right] \tag{2.4.2}$$

In the above, $\mathcal{A}_{t,T}$ the restriction of the set of admissible controls on the time interval $[t, T]$. [3]

Note that $V(t, \omega)$ is a $\mathcal{F}_t$-measurable random variable and in this form it is not useful. However, the following result gives us a lifeline. This well-celebrated result was independently found by [6] and [4] by using Markov selection theorem from [7].

[3] Essential infimum. Using essential infimum is necessary because the supremum of random variables in not necessarily measurable (a random variable). For a family of random variables $\{\chi_u\}_{u\in\mathcal{A}_{t,T}}$, $\mathrm{essinf}_{u\in\mathcal{A}_{t,T}}\chi_u$ is $\mathbb{P}$-a.s unique and is a by the smallest random variable that is larger or equal to χ_u for all $u \in \mathcal{A}_{t,T}$. In our case, $\chi_u = \mathbb{E}\left[\int_t^T C(s, X_s^u, u_s)\mathrm{d}s + g(X_T^u)\Big|\mathcal{F}_t\right]$.

Theorem 2.4.1. *For any well-defined (proper admissibility condition and finite value function) stochastic control problem, the value function does not change if we reduce the set of controls to* ***Markovian controls****, i.e.,* $u_t := \phi(t, X_t^u)$.

Markovian controls are also called feedback controls, especially if the problem is deterministic.

Corollary 2.4.1. *For the value function we have* $V(t,\omega) = V(t, X_t^u)$ *where*

$$\begin{aligned} V(t,x) &= \inf_{u\in\mathcal{A}_{t,T}} \mathbb{E}\left[\int_t^T C(s, X_s^u, u_s)\mathrm{d}s + g(X_T^u)\Big| X_t^u = x\right] \\ &= \inf_{u\in\mathcal{A}^f_{t,T}} \mathbb{E}\left[\int_t^T C(s, X_s^u, u_s)\mathrm{d}s + g(X_T^u)\Big| X_t^u = x\right] \end{aligned} \tag{2.4.3}$$

where $\mathcal{A}^f_{t,T}$ *is he set of all Markovian controls* $u_t = \phi(t, X_t^u)$*; in particular the following SDE has a strong solution for all* x

$$\begin{cases} \mathrm{d}X_s = \mu(s, X_s, \phi(s, X_s))\mathrm{d}s + \sigma(s, X_s, \phi(s, X_s))\mathrm{d}B_s \\ X_t = x \end{cases} \tag{2.4.4}$$

Remark 2.4.1 (Discussion of strong and week solution)**.** *The requirement of existence of strong solution for* (2.4.4) *is for the sake of simplicity. We recall that an SDE has a strong solution when for all filtered probability space hosting a Brownian motion, the SDE has a solution. An SDE has a weak solution if there exists a filtered probability space hosting a Brownian motion in which the SDE has a solution. For more discussion and example of nonexistence of strong solutions see [10] and the discussion in [8].*

It is worth mentioning there are three type of controls:

1. Open-loop controls which do not care about the current state of the system. In our case, an open-loop control is a deterministic control. Control $u_t = \alpha_t' + \beta\alpha_t$ in Example1.1.6 is open-loop because it does not depend on the state of system x_t.

2. A closed-loop or feedback control is a control that depends on the path of the state process, i.e., $u_t = \phi(t, X_{\cdot\wedge t}) = \phi(t, X_s : s \le t)$ or equivalently u_t is adapted to the filtration generated by X, $\{\mathcal{F}_t^X\}_t$, measurable. Note that a closed-loop control is automatically adapter to the filtration generated by the Brownian motion.

3. A Markovian control is a special case of feedback control that only depends on the latest value of the state variable, $u_t = \phi(t, X_t)$. In this course, when we discuss the existence of optimal control, we mean Markovian optimal control unless otherwise is specified.

2.5 Discrete-time dynamic programming principle (DPP)

Recall from Corollary 2.4.1 that the value function satisfies

$$V(t,x) = \inf_{u\in\mathcal{A}_{t,T}} \mathbb{E}\left[\int_t^T C(s, X_s^u, u_s)\mathrm{d}s + g(X_T^u)\Big| X_t^u = x\right] \tag{2.5.1}$$

Dynamic programming principle can be better understood in discrete-time setting. So, here we spend some time to explain a stochastic control problem in discrete time. It is natural to expect that, under suitable conditions, any continuous-time problem can be approximated by a discrete-time problem. For instance, (2.4.1) can be approximated by

$$\inf_u \mathbb{E}\left[\sum_{t=0}^{T-1} C(t, \hat{X}_t^u, u_t)\Delta t + g(\hat{X}_T^u)\right] \tag{2.5.2}$$

where

$$\hat{X}_{t+1}^u = \hat{X}_t^u + \mu(t, \hat{X}_t^u, u_t)\Delta t + \sigma(t, \hat{X}_t^u, u_t)\Delta B_{t+1} \tag{2.5.3}$$

where $\Delta t = T/N$ and $\Delta B_{t+1} = B_{t+\Delta t} - B_t$.

Without loss of generality, we can drop Δt and replace ΔB_{t+1} by a standard i.i.d random variable.

$$\inf_u \mathbb{E}\left[\sum_{t=0}^{T-1} C(t, X_t^u, u_t) + g(X_T^u)\right] \tag{2.5.4}$$

where the infimum is over all stochastic process $u : [0, T] \times \Omega \to \mathbb{R}^m$ and

$$X_{t+1}^u = X_t^u + \mu(t, X_t^u, u_t) + \sigma(t, X_t^u, u_t)\xi_{t+1} \tag{2.5.5}$$

and $\{\xi_t\}_{t=1}^T$ is a sequence of i.i.d. random variables with mean 0 and variance 1.

Note that we can think about the process as a sequence of random variables $u_0, \ldots, u_{T-1}$, which can be chosen based on the discretion of the controller up to the adaptedness condition. This allows us to write the control problem as

$$\inf_{u_0} \cdots \inf_{u_{T-1}} \mathbb{E}\left[\sum_{t=0}^{T-1} C(t, X_t^u, u_t) + g(X_T^u)\right] \tag{2.5.6}$$

The value function for this problem is written as

$$V(t, x) = \inf_{u_t} \cdots \inf_{u_{T-1}} \mathbb{E}\left[\sum_{s=t}^{T-1} C(s, X_s^u, u_s) + g(X_T^u)\middle| X_t^u = x\right] \tag{2.5.7}$$

Since $X_t^u = x$, u_t is chosen over all real numbers and, therefore, $C(t, X_t^u, u_t)$ is deterministic. One can write

$$V(t, x) = \inf_{u_t} C(t, x, u_t) + \inf_{u_{t+1}} \cdots \inf_{u_{T-1}} \mathbb{E}\left[\sum_{s=t+1}^{T-1} C(s, X_s^u, u_s) + g(X_T^u)\middle| X_t^u = x\right] \tag{2.5.8}$$

Tower property of conditional expectation implies that

$$\mathbb{E}\left[\sum_{s=t+1}^{T-1} C(s, X_s^u, u_s) + g(X_T^u)\middle| X_t^u = x\right] = \mathbb{E}\left[\mathbb{E}\left[\sum_{s=t+1}^{T-1} C(s, X_s^u, u_s) + g(X_T^u)\middle| X_{t+1}^u\right]\middle| X_t^u = x\right] \tag{2.5.9}$$

Exercise 12. *Show that one can take the infimum inside the first expectation, i.e.,*

$$\begin{aligned}
&\inf_{u_{t+1},\ldots,u_{T-1}} \mathbb{E}\left[\mathbb{E}\left[\sum_{s=t+1}^{T-1} C(s, X_s^u, u_s) + g(X_T^u)\middle| X_{t+1}^u\right]\middle| X_t^u = x\right] \\
&\quad = \mathbb{E}\left[\inf_{u_{t+1},\ldots,u_{T-1}} \mathbb{E}\left[\sum_{s=t+1}^{T-1} C(s, X_s^u, u_s) + g(X_T^u)\middle| X_{t+1}^u\right]\middle| X_t^u = x\right]
\end{aligned} \tag{2.5.10}$$

By definition of value function,

$$\inf_{u_{t+1},\dots,u_{T-1}} \mathbb{E}\Big[\sum_{s=t+1}^{T-1} C(s, X_s^u, u_s) + g(X_T^u)\Big| X_{t+1}^u\Big] = V(t+1, X_{t+1}^u) \tag{2.5.11}$$

Therefore,

$$V(t,x) = \inf_{u_t} C(t,x,u_t) + \mathbb{E}[V(t+1, X_{t+1}^u)|X_t = x] \tag{2.5.12}$$

This provides us with a recursive formula to solve a discrete-time control problem.

$$\begin{cases} V(t,x) = \inf_{u_t} C(t,x,u_t) + \mathbb{E}[V(t+1, X_{t+1}^u)|X_t = x] \\ V(T,x) = g(x) \end{cases} \tag{2.5.13}$$

2.6 Dynamic programming principle in continuous time

In continuous time, dynamic programming principle is more complicated. Early results used measurable selection theorems to overcome these complications. Later, quantization methods were used to avoid measurable selection theorems. For more information, see [9] and references therein.

Theorem 2.6.1. *If the value function is continuous and have linear growth, then for any stopping time τ, we have*

$$V(t,x) = \inf_u \mathbb{E}_{t,x}\Big[\int_t^\tau C(s, X_s^u, u_s)\mathrm{d}s + V(\tau, X_\tau^u)\Big] \tag{2.6.1}$$

2.7 Hamilton-Jacobi-Bellman equation

In Theorem 2.6.1, for fixed $h, \epsilon > 0$ take $\tau = \tau^h = \inf\{s > 0, |X_s^u - x| \geq \epsilon\} \wedge (t+h)$. Assume that the value function is $C^{1,2}$, once continuously differentiable on t and twice continuously differentiable on x.By applying Itô's formula on $V(\tau^h, X_{\tau^h}^u)$, we obtain

$$\begin{aligned} V(\tau^h, X_{\tau^h}^u) = & V(t,x) \\ & + \int_t^{\tau^h} \Big(\partial_t V(s, X_s^u) + \frac{1}{2} D^2 V(s, X_s^u)\cdot a(s, X_s^u, u_s) + \nabla V(s, X_s^u)\cdot \mu(s, X_s^u, u_s)\Big)\mathrm{d}s \\ & + \int_t^{\tau^h} \sigma(s, X_s^u, u_s)\mathrm{d}B_s \end{aligned} \tag{2.7.1}$$

where $a = \sigma^\intercal \sigma$. Recall for two matrices $A \cdot B = \mathrm{Tr}[A^\intercal B]$. Assuming $\mathbb{E}_{t,x}\Big[\int_0^{\tau^h} \sigma(s, X_s^u, u_s)\mathrm{d}B_s\Big] = 0$[4], (2.6.1) can be written as

$$\begin{aligned} 0 = \inf_u \mathbb{E}_{t,x}\Big[\int_t^{\tau^h} & \Big(C(s, X_s^u, u_s)\mathrm{d}s \\ & \partial_t V(s, X_s^u) + \frac{1}{2} D^2 V(s, X_s^u)\cdot a(s, X_s^u, u_s) + \nabla V(s, X_s^u)\cdot \mu(s, X_s^u, u_s)\Big)\mathrm{d}s\Big] \end{aligned} \tag{2.7.2}$$

[4]One can always choose ϵ in the definition of τ^h such that the stochastic integral has zero expected value.

Note that $\tau^h = O(h)$. Therefore, we divide by h and send $h \to 0$:

$$
\begin{aligned}
0 = &\lim_{h\to 0} \inf_u \mathbb{E}_{t,x}\Big[\frac{1}{h}\int_t^{\tau^h} \Big(C(s, X_s^u, u_s)\mathrm{d}s \\
&\partial_t V(s, X_s^u) + \frac{1}{2}D^2 V(s, X_s^u)\cdot a(s, X_s^u, u_s) + \nabla V(s, X_s^u)\cdot \mu(s, X_s^u, u_s)\Big)\mathrm{d}s\Big] \\
= &\inf_u \lim_{h\to 0} \mathbb{E}_{t,x}\Big[\frac{1}{h}\int_t^{\tau^h} \Big(C(s, X_s^u, u_s)\mathrm{d}s \\
&\partial_t V(s, X_s^u) + \frac{1}{2}D^2 V(s, X_s^u)\cdot a(s, X_s^u, u_s) + \nabla V(s, X_s^u)\cdot \mu(s, X_s^u, u_s)\Big)\mathrm{d}s\Big] \\
= &\inf_u \mathbb{E}_{t,x}\Big[\lim_{h\to 0}\frac{1}{h}\int_t^{\tau^h} \Big(C(s, X_s^u, u_s)\mathrm{d}s \\
&\partial_t V(s, X_s^u) + \frac{1}{2}D^2 V(s, X_s^u)\cdot a(s, X_s^u, u_s) + \nabla V(s, X_s^u)\cdot \mu(s, X_s^u, u_s)\Big)\mathrm{d}s\Big] \\
= &\partial_t V(t,x) + \inf_u\Big\{C(t,x,u) + \frac{1}{2}D^2V(t,x)\cdot a(t,x,u) + \nabla V(t,x)\cdot\mu(t,x,u)\Big\}
\end{aligned} \tag{2.7.3}
$$

The HJB is given by the following PDE:

$$
\begin{cases}
0 = \partial_t V(t,x) + \inf_u \Big\{C(t,x,u) + \frac{1}{2}D^2V(t,x)\cdot a(t,x,u) + \nabla V(t,x)\cdot\mu(t,x,u)\Big\} \\
V(T,x) = g(x)
\end{cases} \tag{2.7.4}
$$

Remark 2.7.1. *Rigorously, the proof is more complected. First, we need to justify* $\lim_{h\to 0}\inf_u = \inf_u \lim_{h\to 0}$*. To show this, we have to show that the expected value as a function of* h *is continuous uniformly on* u*. This can be obtained by continuity of* C*,* V*, and first and second derivatives of* V*. Second, justification of* $\lim_{h\to 0}\mathbb{E} = \mathbb{E}\lim_{h\to 0}$ *requires dominated convergence theorem, which allows to change the order of expected value and limit. This also requires use of mean value theorem. For more information, see [9].*

Example 7. *Consider the stochastic linear quadratic problem below:*

$$
\inf_u \mathbb{E}\Big[\int_0^T \Big(aX_t^2 + bX_t + Au_t^2 + Bu_t\Big)\mathrm{d}t + \alpha X_T^2 + \beta X_T\Big] \tag{2.7.5}
$$

with

$$
\mathrm{d}X_t = (cX_t + du_t)\mathrm{d}t + \sigma \mathrm{d}B_t \tag{2.7.6}
$$

where where $a, A, \alpha > 0$ *and* $b, B, c,$ *and* d *are constants. The HJB is given by*

$$
\begin{cases}
0 = \partial_t V(t,x) + \frac{\sigma^2}{2}\partial_x^2 V(t,x) + cx\partial_x V(t,x) + ax^2 + bx + \inf_u\Big\{Au^2 + \big(B + d\partial_x V(t,x)\big)u\Big\} \\
V(T,x) = \alpha x^2 + \beta x
\end{cases} \tag{2.7.7}
$$

$u = -\frac{B + d\partial_x V(t,x)}{2A}$

$\inf_u\{Au^2 + \big(B + d\partial_x V(t,x)\big)u\} = -\frac{(B + d\partial_x V(t,x))^2}{4A^2}$

$$
\begin{cases}
0 = \partial_t V(t,x) + \frac{\sigma^2}{2}\partial_x^2 V(t,x) + cx\partial_x V(t,x) + ax^2 + bx - \frac{(B + d\partial_x V(t,x))^2}{4A^2} \\
V(T,x) = \alpha x^2 + \beta x
\end{cases} \tag{2.7.8}
$$

We anticipate that $V(t,x) = f(t)x^2 + h(t)x + k(t)$*, which is a separation of variable technique. If we plug* *in $V(t,x)$ into the HJB, we obtain*

$$\begin{cases} f' = -2cf(t) + \frac{f^2(t)}{A^2} - a & f(T) = \alpha \\ h'(t) = (\frac{f(t)}{A^2} - c)h(t) + \frac{Bd}{A^2} f(t) - b & h(T) = \beta \\ k'(t) = \frac{PB^2}{4A} + \frac{bd}{2A^2} h(t) + \frac{1}{4A^2} h^2(t) + f(t) & k(T) = 0 \end{cases} \tag{2.7.9}$$

Note that the above system of ODEs is fully solvable. The first ODE is Riccati equation. We shall show in Section 2.9 that the solution of the PDE is indeed the value function of the linear quadratic optimal control problem.

Exercise 13. *Solve modified version of Exercise 7 with modification*

$$\mathrm{d}X_t = (cX_t + du_t)\mathrm{d}t + (eX_t + fu_t)\mathrm{d}B_t \tag{2.7.10}$$

Exercise 14. *Extend the stochastic linear quadratic problem in Exercise 13 to higher dimension and write the HJB.*

Example 8 (Merton optimal investment problem)**.** *Remember a self-financing portfolio with a Black-Scholes risky asset, $\mathrm{d}S_t = S_t(\mu \mathrm{d}t + \sigma \mathrm{d}B_t)$, under zero interest rate satisfies*

$$\mathrm{d}X_t = \theta_t(\mu \mathrm{d}t + \sigma \mathrm{d}B_t) \tag{2.7.11}$$

where θ is the amount of money invested in the risky asset. Merton problem is to maximize the expected utility of wealth at a time horizon T:

$$\sup_\theta \mathbb{E}[U(X_T^\theta)] \tag{2.7.12}$$

The HJB for this problem is given by

$$\begin{cases} 0 = \partial_t V + \sup_\theta \left\{ \frac{\sigma^2}{2} \partial_x^2 V + \mu \partial_x V \right\} \\ V(T,x) = U(x) \end{cases} \tag{2.7.13}$$

After simplification, we have

$$\begin{cases} 0 = \partial_t V - \frac{\mu^2 (\partial_x V)^2}{2\sigma^2 \partial_x^2 V} \\ V(T,x) = U(x) \end{cases} \tag{2.7.14}$$

We can find a closed form solution for the following cases of utility by using separation of variables, $V(t,x) = f(t)U(x)$:

1. *Exponential utility $U(x) = 1 - e^{-\lambda x}$. The separation of variables is $V(t,x) = -f(t)e^{-\lambda x}$.*
2. *HARA utility $U(x) = \frac{x^{1-\lambda}}{1-\lambda}$, with $\lambda \in (0,1)$.*

Exercise 15 (Merton optimal consumption problem)**.** *It is similar to the Merton optimal investment problem except, the investor is consuming from the account and what matters is the utility of consumption.*

$$\mathrm{d}X_t = \theta_t(\mu \mathrm{d}t + \sigma \mathrm{d}B_t) - c_t \mathrm{d}t \tag{2.7.15}$$

where θ is the amount of money invested in the risky asset and the consumption c satisfies $c_t \geq 0$. Merton problem is

$$\sup_{c \geq 0,\ \theta} \mathbb{E}\Big[\int_0^\infty e^{-\gamma t} U(c_t) \mathrm{d}t\Big], \quad \gamma > 0 \tag{2.7.16}$$

Write the HJB for this problem.

Remark 2.7.2. $\gamma > 0$ *in the above example represents preference of current consumption over future consumption.*

2.7.1 Validity of HJB equation

In the previous section, we showed that if the value function is in $C^{1,2}$, then it satisfies (2.7.4). Therefore, HJB can be used to find the value function. If we know that the HJB has a unique $C^{1,2}$ solution in a suitable class of function which includes the value function, we are done. However, it is not always easy to obtain existence and uniqueness results for nonlinear PDEs, which HJB equations are.

Example 9. *Recall from Exercise 5 the fastest exit problem:*

$$\inf_{|u_t|\le 1} \int_0^\infty 1_{\{X_t^u \in O\}} dt \tag{2.7.17}$$

with $\mathrm{d}X_t^u = u_t dt$. *The HJB for this problem is* $0 = \inf_{|u|\le 1}\{u \cdot \nabla V(x)\} + 1_{\{x\in O\}}$, *which is equivalent to the following boundary value problem*

$$\begin{cases} 0 = -|\nabla V(x)| + 1 & x \in O \\ 0 = V(x) & x \in \partial O \end{cases} \tag{2.7.18}$$

The optimal feedback (closed-loop) exit strategy is give by $u_t^*(x) = -\nabla V(x)$, *which is equivalent to gradient descent.*

When $O = [-1, 1] \subset \mathbb{R}$, *the equation is* $|V'(x)| = 1$. *On the other hand, this one-dimensional simplification has an obvious answer: run as fast as you can to the nearest exit point. This yields* $V(x) = 1 - |x|$. *The value function satisfies the HJB except at* $x = 0$, *which is the sole point with two optimal exit strategies.*

Assume that a robot only knows how to solve HJB equations and it ignores finite number of point of irregularities of the solution. For the robot the function $v(x)$ *given below is as a good of a solution as* $V(x) = 1 = |x|$.

$$v(x) = \begin{cases} 1 - |x| & \frac{1}{2} \le |x| \le 1 \\ |x| & |x| \le \frac{1}{2} \end{cases} \tag{2.7.19}$$

However, if the robot comes of with v *instead of* V, *it suggests to move towards* 0 *for* $|x| \le \frac{1}{2}$, *which is obviously incorrect.*

The above example questions the validity of HJB approach. There are two ways to address the validity issue for the solution of HJB equations. One is the notion of viscosity solution, which singles out $V(x) = 1 - |x|$, as the unique viscosity solution of the HJB. For more information of the viscosity solution approach, see [2]. There are simple criteria that the robot can check to see whether his approximate solution to the HJB equation converges to the unique viscosity solution. See for example, [1]. However, here we closed this discussion by giving a glimpse of what it means to be a viscosity solution in the following example.

Example 10. *Consider the exit time problem with noise:*

$$\inf_{|u_t|\le 1} \mathbb{E}\Big[\int_0^\infty 1_{\{X_t^u \in O\}} dt\Big] \tag{2.7.20}$$

with $\mathrm{d}X_t^u = u_t \mathrm{d}t + \epsilon \mathrm{d}B_t$. *The HJB is given by* $0 = \frac{\epsilon^2}{2}V''(x) + \inf_{|u|\le 1}\{u \cdot \nabla V(x)\} + 1_{\{x\in O\}}$[5], *which is*

[5] My best guess is that the term viscosity comes a fluid mechanics term for the second derivative $\frac{\epsilon^2}{2}V''$.

equivalent to the following boundary value problem

$$\begin{cases} 0 = \frac{\epsilon^2}{2} V''(x) - |\nabla V(x)| + 1 & x \in O \\ 0 = V(x) & x \in \partial O \end{cases} \tag{2.7.21}$$

Back to the simple case where $O = [-1, 1]$, there exists a unique bounded solution, $V^\epsilon(x)$, in C^2. One can find this solution in closed form and verify that as $\epsilon \to 0$, $V^\epsilon(x) \to 1 - |x|$. This clarifies why $v(x)$ is not a solution.

Exercise 16. *In the previous exercise, show that as $\epsilon \to 0$, $V^\epsilon(x) \to 1 - |x|$.*

2.8 Viscosity solutions to the HJB equations

Before moving to the notion of viscosity solutions, we point out an important remark about partial differential equations (PDE) such as the HJBs. The solution to the simplest equations may be not exists. For instance, the backward heat equation given by

$$\begin{cases} 0 = \partial_t V + \frac{\sigma^2}{2} \Delta V - rV + f(t, x) \\ V(0, x) = g(x) \end{cases} \tag{2.8.1}$$

does not have any solution unless f and g have sufficient regularity. However, the forward (regular) heat equation

$$\begin{cases} 0 = \partial_t V + \frac{\sigma^2}{2} \Delta V - rV + f(t, x) \\ V(T, x) = g(x) \end{cases} \tag{2.8.2}$$

has a classical solution regardless of the choice of g and f. The forward heat equation is a parabolic equation while the backward equation is not. Roughly speaking, in a parabolic equation, $\partial_t V$ and $\frac{\sigma^2}{2}\Delta V$ show up with the same sign. In the HJB equation (2.7.4), the coefficient of $\partial_t V$ and $\frac{\sigma^2}{2}\Delta V$, 1 and $\frac{\sigma^2}{2}$ respectively, have the same sign. More rigorously, a nonlinear equation of the form

$$\begin{cases} 0 = \partial_t V(t, x) + F\big(t, x, V(t, x), \nabla V(t, x), D^2 V(t, x)\big) \\ V(T, x) = g(x) \end{cases} \tag{2.8.3}$$

is parabolic if the nonlinerity, $F(t, x, \varrho, \Pi, \Gamma)$, is nondecreasing in the component Γ, which represents the Hessian matrix in the order induced by nonnegative-definite matrices. I.e., if for two matrices Γ_1 and Γ_2, $\Gamma_2 - \Gamma_1$ is a nonnegative-definite matrix, then $F(t, x, \varrho, \Pi, \Gamma_2) \geq F(t, x, \varrho, \Pi, \Gamma_1)$. One can check that

$$F(t, x, \varrho, \Pi, \Gamma) := \sup_{u \in \mathbf{U}} \left\{ C(t, x, u) + \mu(t, x, u)\Pi + \frac{1}{2}\sigma^2(t, x, u)\Gamma - r(t, x, u)\varrho \right\} \tag{2.8.4}$$

is nondecreasing in Γ.

Example 11. *Show that the HJB equation (2.9.16) is a parabolic equation.*

On the other hand, even an equation as simple as the heat equation may have multiple solution if we do not restrict the solution to a proper class of functions. For instance, if $g(x)$ and $f(t, x)$ have linear growth in x, then the heat equation has a unique solution in the class of function with linear growth, but not in the class of functions with super exponential growth.

A linear partial differential equation such as the heat equation has classical solution. Nonlinear equations, however, do not necessary have a classical solution. A classical solution is a solution that is continuously differentiable as much as the equation requires. For instance, if in the heat equation above, $f(t,x)$ is a continuous function, then a solution $V(t,x)$ is a classical solution if $V \in \mathrm{C}^{1,2}$. But, as seen in Exercise 5, $V(x) = 1 - |x|$ is the value function that we expect to satisfy the HJB equation. However, it is not differentiable as $x = 0$.

A suitable notion of weak solution for the HJB equations, which are nonlinear equations, is viscosity solutions. To understand the viscosity solutions, we first need to provide candidates for the weak derivatives of a function at the points of nondifferentiability. The first derivative can always be represented by the slope of a tangent plane. Locally speaking, the tangent plane is a plane that hits the graph of the function at at least one point locally and keeps the graph on one side. For example, for the function$V(x) = 1 - |x|$ at point $x = 0$, the line $y = 1 + mx$ is a tangent line by $x = 0$ for $m \in [-1, 1]$ (red line) but $y = 1 + mx$ for $m \notin [-1, 1]$ is not (green line). See Figure 2.8.1. Therefore, we have

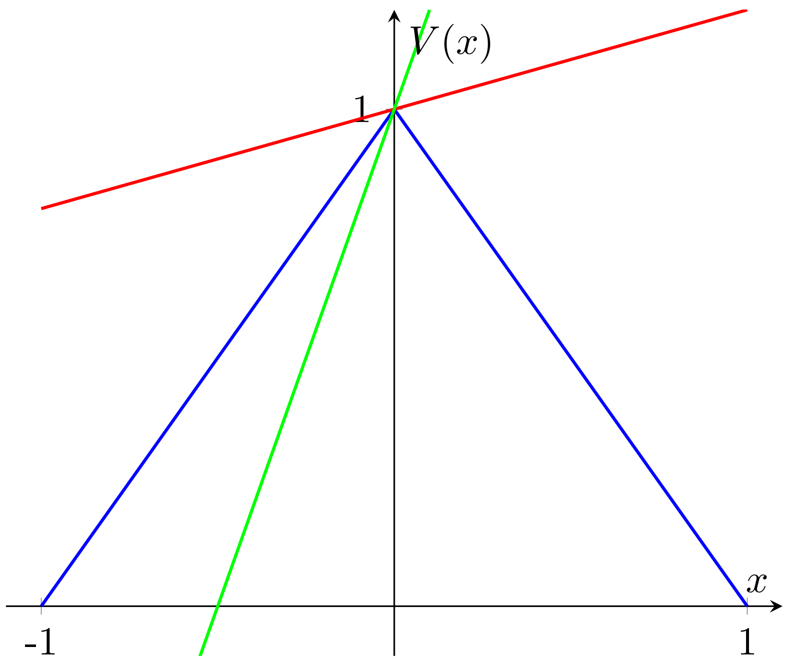

Figure 2.8.1: Red line is an acceptable tangent to the graph, while the green line is not. The green line does not keep the graph on one side locally.

$\varphi(t,x) = a \cdot (x - x_0) + b(t - t_0) + V(t_0, x_0)$ such that $\varphi(t,x) - f(t,x)$ has a local extrema at (t_0, x_0).

For the second derivative tangent planes are not sufficient, because the second derivatives of the tangent planes are always zero. Therefore, we need to appeal to the quadratic functions of the form $\varphi(t,x) = (x - x_0) \cdot A(x - x_0) + a \cdot (x - x_0) + b(t - t_0) + V(t_0, x_0)$. Note that since we have first derivative on t, we only use first order term on t. If a function $\varphi(t,x)$ touches $V(t,x)$ at point (t_0, x_0) from above (resp. below), i.e., (t_0, x_0) is a local minimum (maximum) point for $\varphi - V$, then we call $(b, a, A) = (\partial_t \varphi(t_0, x_0), \nabla \varphi(t_0, x_0), D^2 \varphi(t_0, x_0))$ a superderivative (resp. sub) of V at point (t_0, x_0). Such functions $\varphi(t,x)$ are called test functions. See Figure 2.8.2.

A function $\underline{V}$ (resp. $\overline{V}$) is called a viscosity subsolution (resp. super) of (2.8.3) if for any superderivative (resp. sub) (b, a, A) at point (t_0, x_0), we have

$$\begin{cases} 0 \geq b + F\big(t_0, x_0, \varphi(t_0, x_0), a, A\big), & \text{(resp. } 0 \leq) \\ V(T, x) \leq g(x) \quad \text{(resp. } 0 \geq) \end{cases} \tag{2.8.5}$$

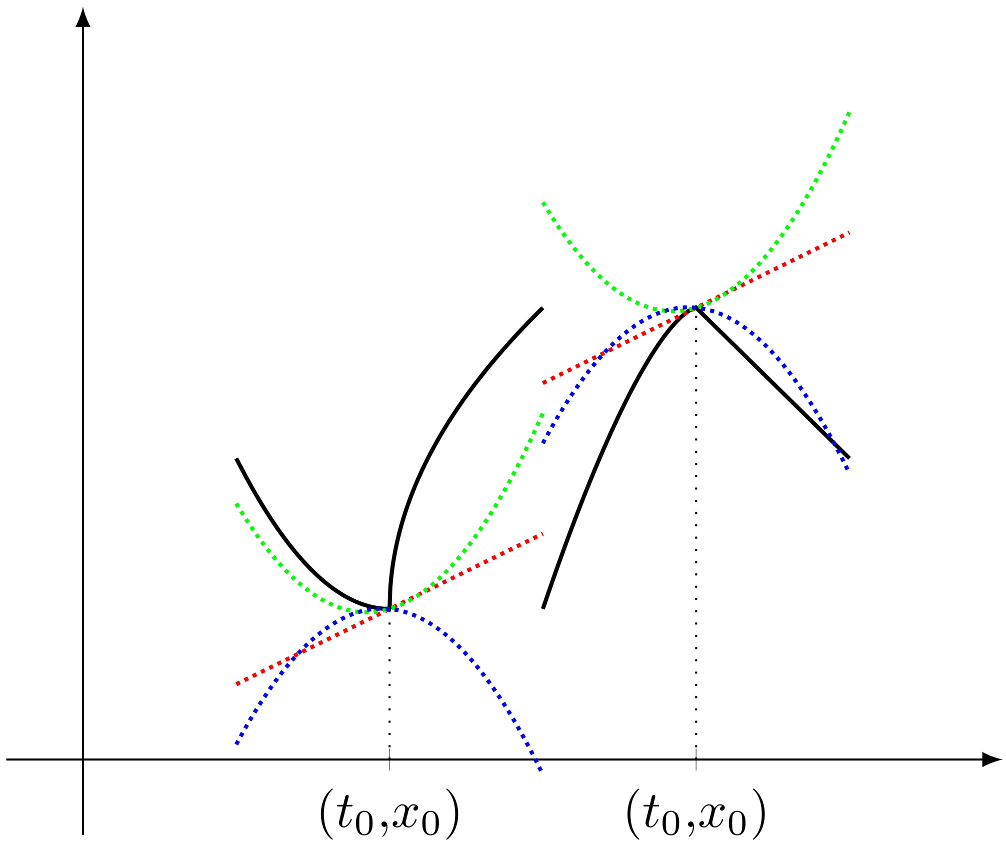

Figure 2.8.2: The test functions that represent the sub and superderivatives of a function at the points of nondifferentiability as well as other points that the function is differentiable.

A function is called a viscosity solution if it is a sub and a super solution.

Remark 2.8.1. *The subsolution (resp. super) is named in accordance with submartingale (resp. super). If $U(t,x) \in \mathrm{C}^{1,2}$ is a subsolution (resp. super) to the equation*

$$0 = \partial_t V(t,x) + C(t,x) + \mu(t,x)\partial_x V(t,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx} V(t,x), \tag{2.8.6}$$

then $Y_t := U(t, X_t)$ is a submartingale (resp. super) martingale, where

$$\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t. \tag{2.8.7}$$

For the first order HJB equation, finding the direction of the inequality for the subsolution or super solutions is not particularly obvious. It can be determined while we add a singular perturbation term εB_t to the deterministic state process and send $\varepsilon \to 0$. For instance, in Exercise 5, if we set the state process $\mathrm{d}X_t^u = u_t \mathrm{d}t + \varepsilon \mathrm{d}B_t$, we obtain the HJB

$$\begin{cases} 0 = 1 - |V'(x)| + \frac{\varepsilon^2}{2} V''(x) \\ V(\pm 1) = 0 \end{cases} \tag{2.8.8}$$

In this case, the subsolution (resp. super) can satisfies

$$\begin{cases} 0 \geq 1 - |V'(x)| + \frac{\varepsilon^2}{2} V''(x) \quad (\text{resp. } 0 \leq) \\ V(\pm 1) \leq 0 \quad (\text{resp. } 0 \geq) \end{cases} \tag{2.8.9}$$

By sending $\varepsilon \to 0$, we obtain the following inequalities for the Eikonal equation, (2.8.8).

$$\begin{cases} 0 \geq 1 - |V'(x)| \quad (\text{resp. } 0 \leq) \\ V(\pm 1) \leq 0 \quad (\text{resp. } 0 \geq) \end{cases} \tag{2.8.10}$$

We can easily check that the function $V(t,x) = 1 - |x|$ is a viscosity solution of the Eikonal equation, (2.8.8). For example, at all points $x \neq 0$, the sub and superderivatives are equal to $-\text{sgn}(x)$, which obviously satisfy the equation with equality. At $x = 0$, we only have super derivatives and the set of subderivatives is empty set. This makes the supersolution property to hold obviously by the false antecedent. The set of superderivatives contains all $m \in [-1, 1]$. Since $0 \leq 1 - |m|$, then $V(t,x) = 1 - |x|$ is also a subsolution. All the above arguments are valid since the boundary condition $V(t, \pm 1) = 0$ validates the sub and supersolution properties.

On the other hand, the function $\tilde{V} = \min\{|x|, 1 - |x|\}$ only satisfies the subsolution property.

Exercise 17. *Show that $\tilde{V}$ is a viscosity subsolution to the Eikonal equation (2.8.10), but not a super solution. Explore the points $x = \pm\frac{1}{2}$.*

The existence and uniqueness of the solution to the nonlinear parabolic equations in the class of functions with linear growth is studies in several papers. For instance, see [2]. For the HJB equations derived from the optimal control problems, the value function of the control problem is usually the viscosity solution of the HJB. The uniqueness is due to a technical lemma, Ishii's lemma. However, in most cases, if we manage to show that any viscosity solution is indeed a classical solution, $\mathrm{C}^{1,2}$, then the verification theorem, Theorem 2.9.1 shows that any classical solution is equal to the value function, and therefore, uniqueness is obtained.

2.9 Verification

Instead of going through viscosity solutions, we consider the cases where the HJB equation has a $C^{1,2}$ solution, which is a candidate for the value function. Then, we introduce a verification theorem, which approves that the solution is indeed the value function.

Theorem 2.9.1 (Verification). *Let HJB equation, (2.7.4) has a $C^{1,2}$ solution, $v(t,x)$. Then, $v(t,x) \geq V(t,x)$. In addition, assume that there exists $u^*(t,x)$ such that*

$$u^*(t,x) \in \mathit{argmin}_u \Big\{ C(t,x,u) + \frac{1}{2} D^2 v(t,x) \cdot a(t,x,u) + \nabla v(t,x) \cdot \mu(t,x,u) \Big\} \tag{2.9.1}$$

and $u^*(t,x)$ is an admissible Markovian (feedback) control, i.e., (2.4.4) has a (strong) solution, X^*_t, and $u^*_t = u^*(t, X^*_t)$ is admissible, $u^* \in \mathcal{A}$. Then, $v(t,x) = V(t,x)$.

Proof. Since v satisfies (2.7.4) in classical sense, we conclude that for any value of $u \in U$, we have

$$0 \geq \partial_t v(t,x) + C(t,x,u) + \mu(t,x,u) \nabla v(t,x) + \frac{1}{2} a(t,x,u) \cdot D^2 v(t,x) \tag{2.9.2}$$

Since $v(t,x) \in \mathrm{C}^{1,2}$, it follows from the Itô formula that for any $u \in \mathcal{A}_{t,T}$,

$$\begin{aligned} v(T, X_T^{t,x,u}) &= v(t,x) \\ &+ \int_0^T \Big(\partial_t v(s, X_s^{t,x,u}) + \mu(s, X_s^{t,x,u}, u_s)\cdot \nabla v(s, X_s^{t,x,u}) + \frac{1}{2} a(s, X_s^{t,x,u}, u_s)\cdot D^2 v(s, X_s^{t,x,u})\Big)\mathrm{d}s \\ &+ \int_t^T \sigma(s, X_s^{t,x,u}, u_s)\mathrm{d}B_s \end{aligned} \tag{2.9.3}$$

By (2.9.2) and terminal condition $v(T,x) = g(x)$, we obtain

$$g(X_T^u) \geq v(t,x) - \int_0^T C(s, X_s^{t,x,u}, u_s)\mathrm{d}s + \int_t^T \sigma(s, X_s^{t,x,u}, u_s)\mathrm{d}B_s \tag{2.9.4}$$

If we take conditional expectation, the martingale property of the stochastic integral implies that

$$\mathbb{E}_{t,x}\left[g(X_T^u) + \int_0^T C(s, X_s^{t,x,u}, u_s)\mathrm{d}s\right] \geq v(t,x) \tag{2.9.5}$$

Now, by taking supremum over $u \in \mathcal{A}_{t,T}$, we obtain

$$V(t,x) = \sup_{u\in\mathcal{A}_{t,T}} \mathbb{E}_{t,x}\left[g(X_T^u) + \int_0^T C(s, X_s^{t,x,u}, u_s)\mathrm{d}s\right] \geq v(t,x). \tag{2.9.6}$$

Now, assume that for some $u^*(t, X_t^*) \in \mathcal{A}$, $Y_t^{u^*} = v(t, X_t^{u^*}) + \int_0^t C(s, X_s^{u^*}, u_s^*)\mathrm{d}s$ is a martingale. By applying the Itô formula, we have

$$\begin{aligned} Y_t^{u^*} =& v(t,x) + \int_t^T \sigma(s, X_s^{t,x,u^*}, u_s)\mathrm{d}B_s + \int_0^T \Big(\partial_t v(s, X_s^{t,x,u^*}) + \mu(s, X_s^{t,x,u^*}, u_s^*)\cdot\nabla v(s, X_s^{t,x,u^*}) \\ &\qquad + \frac{1}{2} a(s, X_s^{t,x,u^*}, u_s^*)\cdot D^2 v(s, X_s^{t,x,u^*}) + C(s, X_s^{u^*}, u_s^*)\Big)\mathrm{d}s \\ =& v(t,x) + \int_t^T \sigma(s, X_s^{t,x,u^*}, u_s)\mathrm{d}B_s \end{aligned} \tag{2.9.7}$$

Therefore, the martingale property of Y^{u^*}, implies that

$$\begin{aligned} C(s, X_s^{u^*}, u_s^*) + &\partial_t v(s, X_s^{t,x,u^*}) + \mu(s, X_s^{t,x,u^*}, u_s^*)\cdot\nabla v(s, X_s^{t,x,u^*}) \\ &+ \frac{1}{2} a(s, X_s^{t,x,u^*}, u_s^*)\cdot D^2 v(s, X_s^{t,x,u^*}) = 0, \quad \text{a.s.} \end{aligned} \tag{2.9.8}$$

Therefore, for u^*, all the inequalities in (2.9.4)-(2.9) holds as equality and we obtain $v = V$. $\square$

We apply Theorem (2.9.1) to the following examples.

Exercise 18 (Merton optimal investment problem)**.** *Remember a self-financing portfolio with a Black-Scholes risky asset,* $\mathrm{d}S_t = S_t(\mu\mathrm{d}t + \sigma\mathrm{d}B_t)$*, under zero interest rate satisfies*

$$\mathrm{d}X_t = \theta_t(\mu\mathrm{d}t + \sigma\mathrm{d}B_t) \tag{2.9.9}$$

where θ is the amount of money invested in the risky asset. Merton problem is

$$\sup_{\theta} \mathbb{E}[U(X_T^{\theta})] \tag{2.9.10}$$

The HJB is given by

$$\begin{cases} 0 = \partial_t v + \sup_\theta \left\{ \frac{\theta^2\sigma^2}{2} \partial_x^2 v + \theta\mu\partial_x v \right\} \\ v(T,x) = U(x) \end{cases} \tag{2.9.11}$$

This simplifies to

$$\begin{cases} 0 = \partial_t v - \frac{(\mu\partial_x v)^2}{\sigma^2\partial_x^2 v} \\ v(T,x) = U(x) \end{cases} \tag{2.9.12}$$

For $U(x) = -e^{-\lambda x}$, with $\lambda > 0$, use separation of variables $u(t,x) = -f(t)e^{-\lambda x}$ to find a closed form solution for the HJB.

Exercise 19. *For $U(x) = x^\lambda$, use separation of variables $u(t,x) = f(t)x^\lambda$ to find a closed form solution for the HJB. Is there any restriction on the value of λ?*

Exercise 20. *For $U(x) = \ln x$, can you suggest a separation of variables?*

Exercise 21 (Merton optimal consumption problem)**.** *It is similar to the Merton optimal investment problem except, the investor is consuming from the account and what matters is the utility of consumption.*

$$\mathrm{d}X_t = \theta_t(\mu \mathrm{d}t + \sigma \mathrm{d}B_t) - c_t \mathrm{d}t \tag{2.9.13}$$

where θ is the amount of money invested in the risky asset and c is the consumption with $c_t \geq 0$. Merton problem is

$$\sup_{\theta,c} \mathbb{E}\Big[\int_0^\infty e^{-\gamma t} U(c_t)\mathrm{d}t\Big],\ \gamma > 0 \tag{2.9.14}$$

Show that the HJB is given by

$$0 = \sup_\theta \left\{ \frac{\theta^2\sigma^2}{2} v'' + \theta\mu v' \right\} + \sup_{c\geq 0} \left\{ U(c) - cv' \right\} - \gamma v \tag{2.9.15}$$

or

$$0 = \frac{(\mu v')^2}{\sigma^2 v''} + \sup_{c\geq 0} \left\{ U(c) - cv' \right\} - \gamma v \tag{2.9.16}$$

For $U(c) = \frac{x^{1-\lambda}}{1-\lambda}$, with $\lambda \in (0,1)$, we have

$$\sup_{c\geq 0} \left\{ U(c) - cv' \right\} = U(c^*) - c^* v' \tag{2.9.17}$$

with $U'(c^*) = v'$ or c^*

2.9.1 Martingale approach

The martingale principle for optimal control ([3]) is another verification result which does not require the differentiability of the value function.

Theorem 2.9.2. *Assume that there exists a function $v(t,x)$ such that for all $u \in \mathcal{A}$ the process $\{Y_t^u\}_{t\geq 0}$*

$$Y_t^u := v(t, X_t^u) + \int_0^t C(s, X_s^u, u_s)\mathrm{d}s \tag{2.9.18}$$

is a super martingale and that for some $u^* \in \mathcal{A}$, $\{Y_t^{u^*}\}_{t\geq 0}$ is a martingale. Then, u^* is an optimal control and the value function is equal to $v(t,x)$.

Proof. By the supermartingale property, we have

$$Y_t^u \leq \mathbb{E}[Y_T^u|\mathcal{F}_t] = \mathbb{E}\Big[g(X_T^u) + \int_t^T C(s, X_s^u, u_s)\mathrm{d}s\Big|\mathcal{F}_t]\Big] + \int_0^t C(s, X_s^u, u_s)\mathrm{d}s \tag{2.9.19}$$

Thus,

$$v(t,x) = Y_t^u - \int_0^t C(s, X_s^u, u_s)\mathrm{d}s \leq \mathbb{E}[Y_T^u|\mathcal{F}_t] = \mathbb{E}\Big[g(X_T^u) + \int_t^T C(s, X_s^u, u_s)\mathrm{d}s\Big|\mathcal{F}_t]\Big], \tag{2.9.20}$$

for all $u \in \mathcal{A}_{t,T}$. Given $X_t^u = x$, we obtain

$$v(t,x) \leq \sup_{u\in\mathcal{A}_{t,T}} \mathbb{E}\Big[g(X_T^u) + \int_t^T C(s, X_s^u, u_s)\mathrm{d}s\Big|X_t^u = x]\Big] = V(t,x). \tag{2.9.21}$$

If for some u^*, Y^{u^*} is a martingale, then in all the above, the inequality turns into equality and we have $v(t,x) = V(t,x)$. □

The above theorem can also be regarded as a verification that a possible candidate $V(t,x)$ is a value function of the optimal control problem. The following example shows the use of this theorem.

Example 12. *In Example 10, let $U(c) = \frac{c^{1-\alpha}}{1-\alpha}$ when $c \geq 0$ and $-\infty$ otherwise. Consider the function $V(t,x) = e^{-\gamma(T-t)}(T - t + e^{A/\alpha t})^\alpha e^{-AT}U(x)$ with $A = r(1-\alpha) - \gamma + \frac{(\mu-r)(1-\alpha)}{2\sigma^2\alpha}$.*

$$Y_t^u = V(t, X_t^u) + \int_0^t e^{-\gamma s}U(c_s)\mathrm{d}s. \tag{2.9.22}$$

By Itô formula, we have

$$\begin{aligned} dY_t^u =&\Big(\partial_t V + \big(\theta_t(\mu - r) + rX_t^u - c_t\big)\partial_x V + \frac{1}{2}\sigma^2\theta_t^2\partial_{xx}V + e^{-\gamma t}U(c_t)\Big)\mathrm{d}t \\ &+ \sigma\theta_t\partial_x V\mathrm{d}B_t. \end{aligned} \tag{2.9.23}$$

By direct calculation, one can see that

$$\partial_t V + \big(\theta_t(\mu - r) + rX_t^u - c_t\big)\partial_x V + \frac{1}{2}\sigma^2\theta_t^2\partial_{xx}V + e^{-\gamma t}U(c_t) \leq 0, \ \ \mathbb{P}\text{-a.s.} \tag{2.9.24}$$

for all values of θ and c. In addition, for $c_t^*(x) = (\partial_x V(t,x))^{-1/\alpha}$ and $\theta_t^*(x) = -\frac{(\mu-r)\partial_x V(t,x)}{\sigma^2\partial_{xx}V(t,x)}$,

$$\partial_t V + \big(\theta_t(\mu - r) + rX_t^u - c_t\big)\partial_x V + \frac{1}{2}\sigma^2\theta_t^2\partial_{xx}V + e^{-\gamma t}U(c_t) = 0, \ \ \mathbb{P}\text{-a.s.} \tag{2.9.25}$$

It remains to show that c^* and θ^* are admissible Markov controls, i.e., to show that

$$dX_t^* = \Big(\theta_t(X_t^*)\big((\mu - r)\mathrm{d}t + \sigma \mathrm{d}B_t\big) - rX_t^*\mathrm{d}t\Big) - c_t^*(X_t^*)\mathrm{d}t, \tag{2.9.26}$$

has a strong solution. We leave the details as an exercise.

Remark 2.9.1. *If the control problem is with infimum instead of supremum,*

$$V(t,x) = \inf_{u \in \mathcal{A}_{t,T}} \mathbb{E}\left[\int_t^T C(s, X_s^{t,x,u}, u_s)\mathrm{d}t + g(X_T^{t,x,u})\right], \tag{2.9.27}$$

Theorem 2.9.2 is should be modified. More precisely, Y_t^u is a submartingale.
For deterministic cases, supermartingale (resp. submartingale) means nonincreasing (resp. nondecreasing).

Finding a candidate for a value function and an optimal control is the subject of the future sections.

2.10 Some nonstandard HJB equations

In this section, we provide a review of some HJB equations that come from stochastic singular control, optimal stopping time, stochastic impulse control, and switching problems. The treatment of such problems via HJB equations is similar, however, we first need to derive HJB equations.

2.10.1 Stochastic singular control problems

Consider the stochastic control problem below:

Example 13.

$$\inf_{u_t} \mathbb{E}\Big[\int_0^T \Big(\frac{1}{2}X_t^2 - u_t\Big)\mathrm{d}t + X_T^2 - X_T\Big] \tag{2.10.1}$$

with $u_t \geq 0$ (takes nonnegative values) and

$$\mathrm{d}X_t = (X_t + u_t)\mathrm{d}t + \sigma \mathrm{d}B_t \tag{2.10.2}$$

The HJB is formally written as

$$\begin{cases} \partial_t V + x\partial_x V + \frac{\sigma^2}{2}\partial_x^2 V + \frac{1}{2}x^2 + \inf_{u \geq 0} u(\partial_x V - 1) = 0 \\ V(T,x) = x^2 - x \end{cases} \tag{2.10.3}$$

Note that

$$\inf_{u \geq 0} u(\partial_x V - 1) = \begin{cases} -\infty & \partial_x V - 1 < 0 \\ 0 & \partial_x V - 1 \geq 0 \end{cases} \tag{2.10.4}$$

Therefore, it is natural to assume that the value function satisfies $\partial_x V - 1 \geq 0$. In fact, we expect so find the value functions such that if $\partial_x V > 1$, then $\partial_t V + x\partial_x V + \frac{\sigma^2}{2}\partial_x^2 V + \frac{1}{2}x^2 = 0$ and when $\partial_x V = 1$, the we only have $\partial_t V + x\partial_x V + \frac{\sigma^2}{2}\partial_x^2 V + \frac{1}{2}x^2 \leq 0$. More precisely, inequality $\partial_x V - 1 \geq 0$ divides $[0,T] \times \mathbb{R}$ into two regions:

- $\mathbf{N} = \{(t,x) : \partial_x V(t,x) > 1\}$
- $\mathbf{R} = \{(t,x) : \partial_x V(t,x) = 1\}$

Inside $\mathbf{N}$*, we expect that* $\partial_t V + x\partial_x V + \frac{\sigma^2}{2}\partial_x^2 V + \frac{1}{2}x^2 = 0$ *holds. This equations has a solution of the form* $v(t,x) = a(t)x^2 + b(t)x + c(t)$.

We denote the boundary of $\mathbf{N}$ *by* $\partial\mathbf{N}$ *and we define* $R(t)$ *such that* $(t, R(t)) \in \partial\mathbf{N}$*, assuming that* $R(t)$ *is uniquely determined. In addition, we expect to see that the value function and its first derivative are the same in the both sides of* $\partial\mathbf{N}$*, particularly,* $\partial_x V(t, R(t)) = 1$*. Therefore, if* v *is the solution to*

$$\begin{cases} \partial_t v + x\partial_x v + \frac{\sigma^2}{2}\partial_x^2 v + \frac{1}{2}x^2 = 0 \\ v(T,x) = x^2 - x \end{cases} \tag{2.10.5}$$

then, we anticipate to write

$$V(t,x) = \begin{cases} v(t,x) & x \geq R(t) \\ x - R(t) + v(t, R(t)) & x < R(t) \end{cases} \tag{2.10.6}$$

Solving (2.10.5)*, we obtain that* $v(t,x) = \frac{5e^{2(T-t)}-1}{4}x^2 - e^{T-t}x - \frac{\sigma^2}{4}\Big(T - t - \frac{5e^{2(T-t)}-1}{2}\Big)$ *and, therefore,* $R(t) = 2\frac{e^{T-t}+1}{5e^{2(T-t)}-1}$.

Exercise 22.

$$\inf_{u_t \geq 0} \mathbb{E}\Big[\int_0^\tau e^{-rt}\Big(\mu - u_t\Big)\mathrm{d}t\Big] \tag{2.10.7}$$

where $\tau = \inf\{s \geq 0 : X_t = 0\}$ *with*

$$\mathrm{d}X_t = (\gamma X_t - u_t)\mathrm{d}t + \sigma \mathrm{d}B_t, \quad X_0 = x \geq 0 \tag{2.10.8}$$

Write the HJB and obtain the nonlinear term by evaluating the infimum.

Singular control problems are easy to detect; the running cost and drift of the SDE are both linear in the control. To see this, consider the control problem

$$\inf_u \mathbb{E}\Big[\int_0^T \big(C(s, X_s^u) + au_s\big)\mathrm{d}s + g(X_T^u)\Big] \tag{2.10.9}$$

where u_t takes values in a closed convex cone $\mathcal{C}$ and

$$\begin{cases} \mathrm{d}X_t^u = (\mu(t, X_t^u) + Au_t)\mathrm{d}t + \sigma(t, X_t^u)\mathrm{d}B_t \\ X_0^u = x \in \mathbb{R}^d \end{cases} \tag{2.10.10}$$

The heuristic derivation of the HJB yields

$$\begin{cases} 0 = \partial_t V(t,x) + \frac{1}{2}a(t,x,u) \cdot D^2 V(t,x) + \nabla V(t,x) \cdot \mu(t,x) + C(t,x) + \inf_u \Big\{au + Au \cdot \nabla V(t,x)\Big\} \\ V(T,x) = g(x) \end{cases} \tag{2.10.11}$$

$a = \sigma^\intercal \sigma$. The infimum above is given by

$$\inf_{\hat{u} \in \mathcal{C},\ |\hat{u}|=1, \lambda \geq 0} \lambda \hat{u} \cdot \big(a + A\nabla V(t,x)\big) = \inf_{\lambda \geq 0} \lambda \inf_{|\hat{u}|=1} \hat{u} \cdot \big(a + A\nabla V(t,x)\big) = \inf_{\lambda \geq 0} \lambda \mathcal{H}(\nabla V(t,x)) \tag{2.10.12}$$

where $\mathcal{H}(\nabla V(t,x)) := \inf_{|\hat{u}|=1} \hat{u} \cdot \big(a + A\nabla V(t,x)\big)$. If $\mathcal{H}(\nabla V(t,x)) < 0$, then the infimum is $-\infty$ and the problem becomes degenerate. To avoid degeneracy, we must have $\mathcal{H}(\nabla V(t,x)) \geq 0$, in which case infimum is attained in $\lambda = 0$ and we have

$$\begin{cases} 0 = \partial_t V(t,x) + \frac{1}{2} a(t,x,u) \cdot D^2 V(t,x) + \nabla V(t,x) \cdot \mu(t,x) + C(t,x) \\ V(T,x) = g(x) \end{cases} \tag{2.10.13}$$

For simplicity of notation, we set

$$\mathcal{L}V(t,x) := \partial_t V(t,x) + \frac{1}{2} a(t,x,u) \cdot D^2 V(t,x) + \nabla V(t,x) \cdot \mu(t,x) + C(t,x) \tag{2.10.14}$$

The interpretation of the HJB requires some *variational inequalities*:[6]

1. for all (t,x), $-\mathcal{L}V(t,x) \geq 0$ and $\mathcal{H}\nabla V(t,x) \geq 0$.
2. If $\mathcal{H}(\nabla V(t,x)) > 0$, then $-\mathcal{L}V(t,x) = 0$.

The description of the optimal control in singular control problems need the notion of local time of SDEs. For instance, in Exercise 22, the optimal control is the local time of the process, $\mathrm{d}X_t = \gamma X_t \mathrm{d}t + \sigma \mathrm{d}B_t$ at the point $\hat{x}$, existence of which is obtained through solving the HJB and going through verification step.

2.10.2 Optimal stopping problem

Let X be given by the SDE

$$\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t \tag{2.10.15}$$

in a filtered probability space and consider the following problem

$$\sup \mathbb{E}\left[\int_0^{\tau\wedge T} e^{-rt} C(t, x_t)\mathrm{d}t + g(X_{\tau\wedge T})\right] \tag{2.10.16}$$

where the supremum is over all stopping time τ adapted to the filtration. These class of problems, which are not stochastic control problems, called optimal stopping problems. To solve optimal stopping problems, we can write the following DPP:

$$V(t,x) = \mathbb{E}_{t,x}\left[\int_0^{\eta} e^{-r(s-t)} C(t, X_s)\mathrm{d}s + e^{-r(\eta - t)} g(X_\eta)\right] \tag{2.10.17}$$

for any stopping time η with values in $[t,T]$. The DPP above leads to the following variational HJB equation:

1. $V(t,x) \geq g(x)$ and $-\mathcal{L}V(t,x) \geq 0$
2. if $V(t,x) > g(x)$, then $-\mathcal{L}V(t,x) = 0$

where $\mathcal{L}$ is given in (2.10.14).

2.11 Backward stochastic differential equations

The backward stochastic differential equations (BSDE) is an alternative way to tackle the optimal control problems. One advantage of the BSDEs is that they can cover the nonMarkovian optimal control problems.

[6] For more details on why we write the variational inequality this way, see [5, Chapter 11] and the notion of sub and super solutions.

The derivation of the BSDEs from an optimal control problem follows from the Itô martingale representation theorem.

Theorem 2.11.1. *In a probability space which hosts a Brownian motion B, let $\mathbb{F}^B := \{\mathcal{F}_t^B\}_t$ be the right-continuous augmented filtration generated by a Brownian motion B. Let X be a $\mathcal{F}_T^B$-measurable random variable with finite expectation. Then, there exists a $\mathbb{F}^B$-progressively measurable process $Z_t(\omega)$ such that*

$$X = \mathbb{E}[X] + \int_0^T Z_t \mathrm{d}B_t \tag{2.11.1}$$

In addition, if M is a continuous martingale with respect to $\mathbb{F}$, then, there exists a $\mathbb{F}^B$-progressively measurable process $\phi_(\omega)$ such that

$$M_t = M_0 + \int_0^t Z_s \mathrm{d}B_s, \quad \textit{for} \quad t \geq 0. \tag{2.11.2}$$

In the above theorem, Z can interpreted as the sensitivity of the martingale M with respect to the Browanian noise B.

Consider for a given stochastic process $C_t(\omega)$

$$Y_t := \mathbb{E}\left[\int_t^T L_s(\omega)\mathrm{d}s + \xi \Big| \mathcal{F}_t^B\right]. \tag{2.11.3}$$

We simply assume that $r_s(\omega)$ and $L_s(\omega)$ are $\mathbb{F}^B$-progressively measurable processes and ξ is a $\mathcal{F}_T^X$-measurable and integrable random variable. For the martingale define by

$$M_t := \mathbb{E}\left[\int_0^T L_s(\omega)\mathrm{d}s + \xi \Big| \mathcal{F}_t^B\right], \tag{2.11.4}$$

and by the Itô martingale representation theorem, we have

$$M_T = \int_0^T L_s(\omega)\mathrm{d}s + \xi = M_t + \int_t^T Z_s \mathrm{d}B_s, \tag{2.11.5}$$

for some $\mathbb{F}^B$-progressively measurable process $Z_t(\omega)$. Since $Y_T = \xi$, we have

$$M_T = \int_0^T L_s(\omega)\mathrm{d}s + Y_T = M_t + \int_t^T Z_s \mathrm{d}B_s, \tag{2.11.6}$$

On the other hand, since $\int_0^t L_s(\omega)\mathrm{d}s$ is $\mathcal{F}_t^B$-measurable, we can write

$$M_t = \int_0^t L_s(\omega)\mathrm{d}s + \mathbb{E}\left[\int_t^T L_s(\omega)\mathrm{d}s + \xi \Big| \mathcal{F}_t^B\right] = \int_0^t L_s(\omega)\mathrm{d}s + Y_t. \tag{2.11.7}$$

Therefore,

$$\int_0^T L_s(\omega)\mathrm{d}s + Y_T = \int_0^t L_s(\omega)\mathrm{d}s + Y_t + \int_t^T Z_s \mathrm{d}B_s, \tag{2.11.8}$$

or,

$$Y_t = Y_T + \int_t^T L_s(\omega)\mathrm{d}s - \int_t^T Z_s \mathrm{d}B_s. \tag{2.11.9}$$

The BSDE (2.11.9) can be written formally by

$$\begin{cases} dY_t = -L(t,\omega)\mathrm{d}t + Z_t \mathrm{d}B_t \\ Y_T = g(X_T) \end{cases} \tag{2.11.10}$$

Note that we could have written the above forwardly, i.e., $Y_t = Y_0 + \int_0^t L_s(\omega)\mathrm{d}s - \int_0^t Z_s \mathrm{d}B_s$. However, this is not very useful, because Y_0 is not known. See example below.

Example 14. *Recall that the solution to the linear equation*

$$\begin{cases} 0 = \partial_t v(t,x) + C(t,x) + [\mu \cdot \nabla v](t,x) + \frac{1}{2}[\sigma^\intercal \sigma \cdot D^2 v](t,x) \\ v(T,x) = g(x) \end{cases} \tag{2.11.11}$$

is given by the Feynmann-Kac formula:

$$V(t,x) = \mathbb{E}\left[\int_t^T C(s,X_s)\mathrm{d}s + g(X_T)\Big| X_t = x\right], \tag{2.11.12}$$

where

$$\mathrm{d}X_t = \mu(t,X_t)\mathrm{d}t + \sigma(t,X_t)\mathrm{d}B_t. \tag{2.11.13}$$

If we define $Y_t = V(t,X_t)$, then (2.11.9) is given by

$$Y_t = g(X_T) + \int_t^T C(s,X_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}B_s. \tag{2.11.14}$$

In the above, g and L are known. Given that we can find Z, Y_t is known. However, if we write the equation forward

$$Y_t = Y_0 + \int_0^t C(s,X_s)\mathrm{d}s - \int_0^t Z_s \mathrm{d}B_s. \tag{2.11.15}$$

Here, $Y_0 = V(0,X_0)$ is not known.

The following theorem shows that a BSDE can generalize the Feynmann-Kac formula to nonlinear equations.

Theorem 2.11.2. *Assume that the following semilinear[7] PDE has a solution $V(t,x) \in \mathrm{C}^{1,2}$.*

$$\begin{cases} 0 = \partial_t v(t,x) + [\mu \cdot \nabla v](t,x) + \frac{1}{2}[\sigma^\intercal \sigma \cdot D^2 V](t,x) + C(t,x,v(t,x),\nabla v(t,x)) \\ V(T,x) = g(x). \end{cases} \tag{2.11.16}$$

Then, $Y_t = V(t,X_t)$ and $Z_t = \sigma(t,X_t)\nabla V(t,X_t)$ satisfy the BSDE

$$\begin{cases} dY_t = -C(t,X_t,Y_t,\sigma^{-1}(t,X_t)Z_t)\mathrm{d}t + Z_t \mathrm{d}B_t \\ Y_T = g(X_T) \end{cases} \tag{2.11.17}$$

where

$$\mathrm{d}X_t = \mu(t,X_t)\mathrm{d}t + \sigma(t,X_t)\mathrm{d}B_t. \tag{2.11.18}$$

[7] The an equation with linear second order term and possibly nonlinear first order term is called semilinear.

Proof. By applying Itô lemma on $Y_t = V(t, X_t)$, we obtain

$$\begin{aligned} dY_t =& \big(\partial_t V(t,X_t) + \mu(t,X_t)\nabla V(t,X_t) + \frac{1}{2}\sigma^2(t,X_t)\partial_{xx}V(t,X_t)\big)\mathrm{d}t + \sigma(t,X_t)\nabla V(t,X_t)\mathrm{d}B_t \\ =& -C(t,X_t,V(t,X_t),\nabla V(t,X_t))\mathrm{d}t + \sigma(t,X_t)\nabla V(t,X_t)\mathrm{d}B_t \\ =& -C(t,X_t,Y_t,\sigma^{-1}(t,X_t)Z_t)\mathrm{d}t + Z_t\mathrm{d}B_t \end{aligned} \tag{2.11.19}$$

□

In the above, the second equality is obtained from the PDE and the third equality is from the definition of Y and Z. In addition, by the terminal condition we have, $Y_T = V(T, X_T) = g(X_T)$. In general, the use of BSDEs over the HJB is preferable when the regularity of the function $V(t,x)$ is not established. Because by the existence theorem for the BSDEs, we already know that Z_t exists as stochastic process even if $\sigma(t,X_t)\nabla V(t,X_t)$ does not make sense in cases when ∇V does not exist as specific points.

Motivated by Theorem 2.11.2, we can define a general BSDE as

$$\begin{cases} dY_t = -C(t,Y_t,Z_t,\omega)\mathrm{d}t + Z_s\mathrm{d}B_s \\ Y_T = \xi \end{cases}, \tag{2.11.20}$$

The argument ω inside L represents a general dependence on the randomness. In particular, it can represent solution X_t of a possible path-dependent SDE

$$\mathrm{d}X_t = \mu(t, X.)\mathrm{d}t + \sigma(t, X.)\mathrm{d}B_t, \tag{2.11.21}$$

a possible control process u_t, or both X_t and u_t. We discuss such dependencies in further details in Section 2.11.3.

Remark 2.11.1. *In (2.11.20), we blended σ^{-1} inside the Lagrangian $L(s,X_s,Y_s,\sigma^{-1}Z_s,\omega)$ into it and simply write $L(s,X_s,Y_s,Z_s,\omega)$.*

Theorem 2.11.3 (Existence and uniqueness theorem for BSDEs)**.** *Assume that*

i) *$\mu(t,x)$ and $\sigma(t,x)$ are Lipschitz in t and x*

ii) *for all (t,x) with $0 < \lambda|x|^2 \le x^\top\sigma^\top\sigma x$ for all x and for some λ (in particular σ is invertible)*

iii) *$L(t,x,\varrho,\Pi,\omega)$ is progressively measurable in ω and Lipschitz in other variables (t,x,ϱ,Π) and is decreasing in ϱ a.s.*

Then, for any $\mathcal{F}_T$-measurable square-integrable random variable ξ, there exists a couple (Y_t, Z_t) such that (2.11.20) holds, i.e.,

$$Y_t = \xi + \int_t^T L(s,X_s,Y_s,Z_s,\omega)\mathrm{d}s - Z_s\mathrm{d}B_s, \tag{2.11.22}$$

and there exists a constant C that only depends on the Lipschitz constant of μ, σ, and L and the constant λ that

$$\mathbb{E}\left[\sup_{t\in[0,T]}|Y_t|^2 + \int_0^T Z_s^2\mathrm{d}s\right] \le C\left(\mathbb{E}[\xi^2]\right) \tag{2.11.23}$$

2.11.1 Linear BSDEs

Linear BSDEs take the form

$$dY_t = -(\alpha_t Y_t + \beta_t Z_t + L_t)\mathrm{d}t + Z_t \mathrm{d}B_t, \tag{2.11.24}$$

where α, β, and L are arbitrary progressively measurable processes. One can write a closed-form solution for a linear BSDE. To do so, note that

$$dY_t = -(\alpha_t Y_t + L_t)\mathrm{d}t + Z_t dB_t^\beta, \tag{2.11.25}$$

where by the Girsanov theorem, $dB_t^\beta := \beta_t \mathrm{d}t + \mathrm{d}B_t$ is a Brownian motion under the probability $\mathbb{P}^\beta$ given in

$$\frac{d\mathbb{P}^\beta}{d\mathbb{P}}\Big|_{\mathcal{F}_t} := \exp\left(\int_0^t \beta_s \mathrm{d}B_s - \frac{1}{2}\int_0^t \beta_s^2 \mathrm{d}s\right) \tag{2.11.26}$$

If we define $\tilde{Y}_t := e^{\int_0^t \alpha_s \mathrm{d}s} Y_t$, then

$$d\tilde{Y}_t = e^{\int_0^t \alpha_s \mathrm{d}s}\left(L_t \mathrm{d}t + Z_t dB_t^\beta\right). \tag{2.11.27}$$

In other words,

$$\tilde{Y}_t = \tilde{Y}_T - \int_t^T e^{\int_0^s \alpha_z dz}\left(L_s \mathrm{d}t + Z_s dB_s^\beta\right), \tag{2.11.28}$$

or,

$$Y_t = e^{\int_t^T \alpha_s \mathrm{d}s}\tilde{Y}_T - \int_t^T e^{\int_t^s \alpha_z dz}\left(L_s \mathrm{d}t + Z_s dB_s^\beta\right) = e^{\int_t^T \alpha_s \mathrm{d}s}\xi - \int_t^T e^{\int_t^s \alpha_z dz}\left(L_s \mathrm{d}t + Z_s dB_s^\beta\right). \tag{2.11.29}$$

After taking conditional expectation with respect to $\mathbb{P}^\beta$, we obtain

$$Y_t = \mathbb{E}^\beta\left[e^{\int_t^T \alpha_s \mathrm{d}s}\xi - \int_t^T e^{\int_t^s \alpha_z dz} L_s \mathrm{d}t \Big| \mathcal{F}_t\right]. \tag{2.11.30}$$

If we define

$$\Gamma_t = e^{\int_t^T \alpha_s \mathrm{d}s}\frac{d\mathbb{P}^\beta}{d\mathbb{P}}\Big|_{\mathcal{F}_t} = \exp\left(\int_0^t \beta_s \mathrm{d}B_s + \int_0^t (\alpha_s - \frac{1}{2}\beta_s^2)\mathrm{d}s\right), \tag{2.11.31}$$

by changing the measure back to $\mathbb{P}$

$$Y_t = \Gamma_t^{-1}\mathbb{E}\left[\Gamma_T \xi - \int_t^T \Gamma_s L_s \mathrm{d}t \Big| \mathcal{F}_t\right]. \tag{2.11.32}$$

2.11.2 Comparison principle

Consider

$$\begin{cases} dY_t^{(i)} = -L^{(i)}(t, Y_t^{(i)}, Z_t^{(i)}, \omega)\mathrm{d}t + Z_s^{(i)}\mathrm{d}B_s \\ Y_T^{(i)} = \xi^{(i)} \end{cases}, \tag{2.11.33}$$

for $i = 1,\ 2$. The following theorem provides a sufficient condition for comparing $Y^{(1)}$ and $Y^{(2)}$, and hence, it is called comparison principle.

Theorem 2.11.4. *Assume that for $i = 1,\ 2$, $(Y^{(i)}, Z^{(i)})$ is the solution for (2.11.33) and $L(t, y, z, \omega)$ is*

Lipschitz in (y, z) uniformly in ω and t. Further assume that $\xi^{(1)} \geq \xi^{(2)}$ and, for each value (t, y, z), $L^{(1)}(t, y, z, \omega) \leq L^{(2)}(t, y, z, \omega)$, a.s. Then, $Y^{(1)} \geq Y^{(2)}$, a.s.

Proof. □

2.11.3 Maximum principle

For the purpose of control theory, we assume that L takes the form $L(t, x, \varrho, \Pi, u_t)$, where u_t is a progressively measurable process that represents a control at time t. In this case, we denote the solution to the BSDE by (Y^u, Z^u), where

$$\begin{cases} Y_t^u = \xi + \int_t^T F(s, X_s, Y_s^u, Z_s^u, u_s)\mathrm{d}s - \int_t^T Z_s^u \mathrm{d}B_s \\ Y_T = \xi \in \mathcal{F}_t^X \\ \mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t \end{cases} \tag{2.11.34}$$

and $F(s, X_s^u, Y_s^u, Z_s^u, u_s) = L(s, X_s, Y_s, \sigma^{-1}(s, X_s^u, u_s)Z_s, u_s)$. Then, a control problem can be written as

$$Y_t = \text{esssup}_{u \in \mathcal{A}_{t,T}} Y_t^u. \tag{2.11.35}$$

Note that the set of admissible controls are defined specific to a particular problem. In addition, if the terminal condition $\xi = g(X_t^u)$, the problem is a Markovian control problem that was studied in previous sections. Otherwise, it is not Markovian, and therefore, the value cannot be written as a function $V(t, x)$. In fact, value function takes a more general form of

$$Y_t = \text{esssup}_{u \in \mathcal{A}_{t,T}} \mathbb{E}\left[\int_t^T F(s, X_s^u, Y_s^u, Z_s^u, u_s)\mathrm{d}s + \xi \middle| \mathcal{F}_t^X\right]. \tag{2.11.36}$$

In the above, the Lagrangian F not only depends on the state of the system X_s^u and control u_s, but also depends on the history of the value Y_s^u and the sensitivity of the value with respect to the Brownian noise Z_t^u.

To solve such optimal control problems, we require the following comparison principle.

Theorem 2.11.5 (Maximum principle for BSDEs)**.** *The process*

$$Y_t = \text{esssup}_{u \in \mathcal{A}_{t,T}} \mathbb{E}\left[\int_t^T F(s, X_s^u, Y_s^u, Z_s^u, u_s)\mathrm{d}s + \xi \middle| \mathcal{F}_t^X\right] = \text{esssup}_{u \in \mathcal{A}_{t,T}} Y_t^u \tag{2.11.37}$$

there exists a Z such that (Y, Z) satisfies the BSDE

$$\begin{cases} Y_t = \xi + \int_t^T F^*(s, X_s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}B_s \\ Y_T = \xi \in \mathcal{F}_T^X \\ \mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t \end{cases} \tag{2.11.38}$$

where

$$F^*(s, x, y, z) := \sup_{u \in U} F(t, x, y, z, u), \quad \textit{and} \quad u^*(t, x, y, z) := \text{argmax}_{u \in U} F(t, x, y, z, u) \tag{2.11.39}$$

Chapter 3

Numerical evaluation of stochastic control problems

3.1 DPP based approximation

Recall from (2.5.13), the value function and the optimal control in discrete stochastic control problem (2.5.4) can be evaluated through.

$$\begin{cases} V(t,x) = \inf_u C(t,x,u) + \mathbb{E}[V(t+1, X^u_{t+1})|X_t = x] \\ V(T,x) = g(x) \\ u^*_t(x) \in A(x) := \operatorname{argmax}_u C(t,x,u) + \mathbb{E}[V(t+1, X^u_{t+1})|X_t = x] \end{cases} \tag{3.1.1}$$

where

$$X^u_{t+1} = x + \mu(t,x,u) + \sigma(t,x,u)\xi_{t+1} \tag{3.1.2}$$

and $\{\xi_t\}_{t=1}^T$ is a sequence of i.i.d. random variables with mean 0 and variance 1. In the above, we need to evaluate (1) conditional expectation $\mathbb{E}[\cdot|X_t = x]$ and (2) the infimum over control u. Given $V(t+1,\cdot)$, these calculations can be done separately. However, the separate evaluation creates inefficiency in the calculations. Therefore, we propose the one-shot approximation of $\inf_u C(t,x,u)+\mathbb{E}[V(t+1, X^u_{t+1})|X_t = x]$ through the following methods.

3.1.1 Evaluation of infimum

First note that

$$\begin{aligned} C(t,x,u(x)) + &\mathbb{E}[V(t+1, X^u_{t+1})|X_t = x] \\ &= \mathbb{E}[C(t,X,u(X)) + V(t+1, X + \mu(t,X,u(X)) + \sigma(t,X,u(X))\xi_{t+1})|X = x] \end{aligned} \tag{3.1.3}$$

Denoting the right-hand side above by $\Phi(x,u)$, then we seek a function $\hat{u}^*(x)$ such that

$$\hat{u}^*(\cdot) \in A(X) := \operatorname{argmin}_{u(\cdot)} \mathbb{E}[\Phi(X, u(X))] \tag{3.1.4}$$

The following Lemma guarantees that $\hat{u}^*$ solves (3.1.1).

Lemma 3.1.1. *Assume that $\hat{u}^* \in A(X)$ defined above. Then, $\hat{u}^*(x) \in A(x)$ for all x in the set of values of X.*

Proof. For a complete proof, we need a measurable selection theorem and some other conditions. For simplicity, we only provide the sketch of the proof. Assume that $u^*(x) \notin A(x)$ with a positive probability on the set of values of X. We denote the set by B. Then, $x \in B$, we define $\tilde{u}^*(x)$ such that

$$\tilde{u}^*(x) \in \operatorname{argmax}_u C(t, x, u) + \mathbb{E}[V(t+1, X^u_{t+1}) | X_t = x]$$

and $\tilde{u}^*(x) = \hat{u}^*(x)$. Therefore,

$$\mathbb{E}[\Phi(X, \tilde{u}^*(X))] < \mathbb{E}[\Phi(X, \hat{u}^*(X))]$$

which contradicts the definition of $\hat{u}^*$. □

From the nonparametric point of view, we can approximate $\hat{u}^*(\cdot)$ in (3.1.4) by a nonparametric model $u(\cdot;\theta)$ via the following optimization

$$\theta^* \in \operatorname{argmin}_\theta \mathbb{E}[\Phi(X, u(X;\theta))] \tag{3.1.5}$$

For instance, $u(\cdot;\theta)$ can be a neural network with parameter θ.

3.2 Multi-step evaluation

While (3.1.1) is a classic way to evaluate optimal control problems, it is not very efficient. There are two reasons for the lack of efficiency. First, we have to run a loop over the number of time steps. Second, after evaluation of each nonparametric, we need to evaluate the value function. The total number of parameters is the number of parameters in each step times the number of steps, which for nonparametric models is massive and potentially needs a lot of memory. Therefore, we propose the following optimization in place of (3.1.1).

$$\inf_{u(\cdot,\cdot)} \mathbb{E}\left[\sum_{t=0}^{T-1} C(t, X^u_t, u(t, X^u_t)) + g(X^u_T)\right] \tag{3.2.1}$$

where

$$\begin{cases} X^u_{t+1} = X^u_t + \mu(t, X^u_t, u(t, X^u_t)) + \sigma(t, X^u_t, u(t, X^u_t))\xi_{t+1} \\ X_0 \text{ is a random variable.} \end{cases} \tag{3.2.2}$$

The justification for the above problem is the same as Lemma 3.1.1 for one-step DPP method. Note that we can use a nonparametric model, $u(t, x; \theta)$, to approximate the minimization problem by

$$\theta^* \in \operatorname{argmin}_\theta \mathbb{E}\left[\sum_{t=0}^{T-1} C(t, X^u_t, u(t, X^u_t; \theta)) + g(X^u_T)\right] \tag{3.2.3}$$

with

$$\begin{cases} X^u_{t+1} = X^u_t + \mu(t, X^u_t, u(t, X^u_t; \theta)) + \sigma(t, X^u_t, u(t, X^u_t; \theta))\xi_{t+1} \\ X_0 \text{ is a random variable.} \end{cases} \tag{3.2.4}$$

When an approximate optimal strategy, $u(t, x; \theta^*)$, is found, then one can find the value function through evaluating

$$\mathbb{E}_{t,x}\left[\sum_{s=t}^{T-1} C(s, X^u_s, u(t, X^u_s; \theta^*)) + g(X^u_T)\right] \tag{3.2.5}$$

with

$$\begin{cases} X^u_{s+1} = X^u_s + \mu(s, X^u_s, u(s, X^u_s; \theta^*)) + \sigma(s, X^u_s, u(s, X^u_s; \theta^*))\xi_{s+1} \\ X_t = x \end{cases} \tag{3.2.6}$$

Of course, the evaluation of the conditional expectation above is a different problem.

3.3 Numerical methods based on BSDEs

Recall from Section 2.11 that the solution $V(t, x)$ to the semilinear equation

$$\begin{cases} 0 = \partial_t v(t, x) + [\mu \cdot \nabla v](t, x) + \frac{1}{2}[\sigma^\intercal \sigma \cdot D^2 V](t, x) + C(t, x, v(t, x), \nabla v(t, x)) \\ V(T, x) = g(x). \end{cases} \tag{3.3.1}$$

is related to the BSDE

$$\begin{cases} dY_t = -C(t, X_t, Y_t, \sigma^{-1}(t, X_t)Z_t)\mathrm{d}t + Z_t \mathrm{d}B_t \\ Y_T = g(X_T) \end{cases} \tag{3.3.2}$$

by

$$Y_t = V(t, X_t) \text{ and } Z_t = [\sigma \nabla V](t, X_t) \tag{3.3.3}$$

where

$$\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t. \tag{3.3.4}$$

One way to interpret the solution to the PDE is to find functions $\mathbf{Y}(t, x)$ and $\mathbf{Z}(t, x)$ such that

$$\mathbf{Y}(t, X_t) = g(X_T) - \int_t^T \Big(C(s, X_s, \mathbf{Y}(s, X_s), \mathbf{Z}(s, X_s))\mathrm{d}s + \mathbf{Z}(s, X_s)\mathrm{d}B_s\Big) \tag{3.3.5}$$

Functions $\mathbf{Y}$ and $\mathbf{Z}$ can be approximated by neural networks $\hat{\mathbf{Y}}(t, x, \alpha)$ and $\hat{\mathbf{Z}}(t, x; \beta)$ and the equations (3.3.4) and (3.3.5) can be approximated discretely by

$$\begin{cases} \hat{\mathbf{Y}}(t_n, \hat{X}_{t_n}) = \hat{\mathbf{Y}}(t_{n+1}, \hat{X}_{t_{n+1}}) - C(t_n, \hat{X}_{t_n}, \hat{\mathbf{Y}}(t_n, \hat{X}_{t_n}), \hat{\mathbf{Z}}(t_n, \hat{X}_{t_n}))\Delta t + \hat{\mathbf{Z}}(t_n, \hat{X}_{t_n})\Delta B_{t_{n+1}} \\ \hat{X}_{t_{n+1}} = \hat{X}_{t_n} + \mu(t_n, \hat{X}_{t_n})\Delta t + \sigma(t_n, \hat{X}_{t_n})\Delta B_{t_{n+1}}. \end{cases} \tag{3.3.6}$$

Specifically, we shall determine $\hat{\mathbf{Y}}$ and $\hat{\mathbf{Z}}$ such that $g(\hat{X}_T)$ is as close to $\hat{\mathbf{Y}}(T, \hat{X}_T)$ as possible. In other words, we can find the approximate solution $\hat{\mathbf{Y}}(t, x, \alpha^*)$ and $\hat{\mathbf{Z}}(t, x; \beta^*)$ with

$$(\alpha^*, \beta^*) \in \operatorname*{argmin}_{\alpha, \beta} \mathbb{E}\Big[\Big(\hat{\mathbf{Y}}(T, \hat{X}_T) - g(\hat{X}_T)\Big)^2\Big] \tag{3.3.7}$$

Bibliography

[1] G. Barles and P. E. Souganidis, Convergence of approximation schemes for fully nonlinear second order equations, in Asymptotic analysis, no. 4, 1991, pp. 2347–2349.

[2] M. G. Crandall, H. Ishii, and P.-L. Lions, User's guide to viscosity solutions of second order partial differential equations, Bulletin of the American Mathematical Society, 27 (1992), pp. 1–67.

[3] M. Davis and P. Varaiya, Dynamic programming conditions for partially observable stochastic systems, SIAM Journal on Control, 11 (1973), pp. 226–261.

[4] N. El Karoui, N. Du Huu, and M. Jeanblanc-Picqué, Compactification methods in the control of degenerate diffusions: existence of an optimal control, Stochastics, 20 (1987), pp. 169–219.

[5] W. H. Fleming and H. M. Soner, Controlled Markov processes and viscosity solutions, vol. 25, Springer, 2006.

[6] U. Haussmann, Existence of optimal markovian controls for degenerate diffusions, in Stochastic differential systems, Springer, 1986, pp. 171–186.

[7] N. V. Krylov, On the selection of a markov process from a system of processes and the construction of quasi-diffusion processes, Mathematics of the USSR-Izvestiya, 7 (1973), p. 691.

[8] L. C. G. Rogers and D. Williams, Diffusions, Markov processes and martingales: Volume 2, Itô calculus, vol. 2, Cambridge university press, 2000.

[9] N. Touzi, Optimal stochastic control, stochastic target problems, and backward SDE, vol. 29, Springer, 2012.

[10] B. Tsirel'son, An example of a stochastic differential equation having no strong solution, Theory of Probability & Its Applications, 20 (1976), pp. 416–418.