STOCHASTIC CONTROL ANALYSIS

Oluwadamilola Oyekan

Abstract

In this paper, we discuss the Stochastic Control Problem, which deals with the best way to control a stochastic system and has applications in fields as diverse as robotics, finance, and industry. After quickly reviewing mathematical probability and stochastic integration, we prove the HamiltonJacobi-Bellman equations for stochastic control, which transform the Stochastic Control Problem into a partial differential equation.

CONTENTS

1. INTRODUCTION

The Stochastic Control Problem arises when we wish to control a continuoustime random system optimally. For example, we wish to fly an airplane through turbulence, or manage a stock portfolio, or precisely maintain the temperature and pressure of a system in the presence of random air currents and temperatures. To build up to a formal definition of this problem, we will first review mathematical probability and stochastic calculus. These will give us the tools to talk about Itô processes and diffusions, which will allow us to state the Stochastic Control Problem formally. Finally, we will spend the latter half of this paper proving necessary conditions and sufficient conditions for its solution.

2. Mathematical Probability

In this section we briskly define the mathematical model of probability required in the rest of this paper. We define probability spaces, random variables, and conditional expectation. For more detail, see Williams [4].

Definition 2.1 (Probability Space). A probability space is a triple $(\Omega, \mathcal{F}, P)$, where Ω is an arbitrary set, $\mathcal{F}$ is a sigma algebra over Ω, and P is a measure on sets in $\mathcal{F}$ with $P(\Omega) = 1$.

Intuitively, Ω is a set of scenarios which could occur in an experiment, $\mathcal{F}$ contains the subsets of Ω to which it is possible to assign a probability, called events, and for an event $E \in \mathcal{F}, P(S)$ is the probability that the outcome of the experiment is one of the scenarios in E, that is, that E occurs. We will use this probability space for the rest of this section.

Definition 2.2 (Random Variable). A random variable is a measurable function $X : \Omega \to \mathbb{R}^n$ where $\mathbb{R}^n$ is equipped with the Borel sigma-algebra. For a Borel subset $B \in \mathbb{R}^n$, we abuse notation by writing

$$P(X \in B) = P(\{\omega : X(\omega) \in B\}).$$

Intuitively, $X(\omega)$ is a quantity of interest associated with scenarios in the experiment, and $P(X \in B)$ is the probability that the X realized in the experiment will end up with a value in B. X will be a random variable on $(\Omega, \mathcal{F}, P)$ for the rest of this section.

Definition 2.3 (Independence). The random variables $X_1, X_2, \ldots, X_n$ are independent if for all Borel sets $B_1, B_2, \ldots, B_n$,

$$\begin{aligned} P\left((X_1 \in B_1) \cap \ldots \cap (X_n \in B_n)\right) &:= P\left(X_1^{-1}(B_1) \cap \ldots \cap X_n^{-1}(B_n)\right) \\ &= P\left(X_1^{-1}(B_1)\right) \times \ldots \times P\left(X_n^{-1}(B_n)\right). \end{aligned}$$

Intuitively, this means that knowing information about some of the X_i tells you nothing about the other X_i.

It is often more convenient to think about random variables than underlying scenarios and events. Given a random variable, we can "work backwards" to define the set of events required to support our random variable in a natural way:

Definition 2.4 (Sigma Algebra Generated By Random Variable). The sigmaalgebra generated by X is

$$\mathcal{F}_X = \left\{X^{-1}(B) : B \text{ Borel } \right\}.$$

It is easy to see that this is indeed a sigma-algebra. $\mathcal{F}_X$ is the set of events whose occurence or lack of occurence may be decided by the value of X. We can assign a measure to the events in $\mathcal{F}_X$ by using the probabilities already defined for X. For $E \in \mathcal{F}_X$

$$P(E) = P(X \in X(E)).$$

Definition 2.5 (Expectation). The expectation of X is

$$E[X] = \int_\Omega X dP$$

Intuitively, $E[X]$ is the average value of X observed in the experiment.

Definition 2.6 (Conditional Expectation). Let $\mathcal{A}$ be a sub-sigma-algebra of $\mathcal{F}$, and let $E[|X|] < \infty$. Then a conditional expectation of $X \in \mathbb{R}^n$ given $\mathcal{A}$ is an $\mathcal{A}$-measurable function $E[X \mid \mathcal{A}] : \Omega \to \mathbb{R}^n$ satisfying the equation

$$\int_A E[X \mid \mathcal{A}] dP = \int_A X dP \tag{2.7}$$

for all $A \in \mathcal{A}$.

That such a function in fact exists follows from the Radon-Nikodym theorem. Uniqueness is easier to see; this function is essentially unique in the sense that two $\mathcal{A}$-measurable functions satisfying (2.7) will be identical almost everywhere. If (as is almost always the case) $\mathcal{A}$ is the sigma algebra generated by some random variable Y, then intuitively the conditional expectation of X given $\mathcal{A}$, evaluated at the scenario ω, is the best guess of the value of $X(\omega)$ given the value $Y(\omega)$.

Definition 2.8 (Stochastic Process). A stochastic process is an indexed set of random variables $\{X_t\}_{t \in T}$ defined on the probability space $(\Omega, \mathcal{F}, P)$. The sample path of X_t under scenario ω is the function $t \to X_t(\omega)$.

Intuitively, X_t represents the state of a random system at time t, and the sample path under scenario ω is the way the system evolves with time under that scenario. In this paper, usually $T = [0, \infty)$.

Definition 2.9 (Adapted Process). Let $\{\mathcal{F}_t\}_{t \geq 0}$ be an increasing sequence of sigma algebras, that is, $\mathcal{F}_{t_1} \subset \ldots \subset \mathcal{F}_{t_n}$ for $t_1 < \ldots < t_n$. Then a stochastic process $X_t(\omega)$ is called $\mathcal{F}_t$-adapted if for every $t \geq 0, X_t(\omega)$ is $\mathcal{F}_t$ measurable.

Intuitively, this means that X_t does not incorporate any information from the future. All reasonable stochastic processes are adapted to a relevant sequence of sigma algebras.

We end this chapter by introducing a notation we will use in the rest of the paper:

Definition 2.10 (Indicator Function). The indicator function $\mathbf{1}_{\{Q\}}$ has value 1 if Q is true, and 0 otherwise.

3. Stochastic Calculus

In this section we rapidly review Brownian Motion, Stochastic Integration, and Itô's Lemma. For more detail, see the excellent Lalley notes [2]. We begin by defining the most important stochastic process in this paper, and arguably the most important stochastic process yet known.

Definition 3.1 (Brownian Motion). A one-dimensional Brownian Motion is a stochastic process $W_t \in \mathbb{R}$ with the following three properties:

(1) For $0 \leq t_0 < t_1 < \ldots < t_n < \infty$, the increments

$$W_{t_1} - W_{t_0}, \ldots, W_{t_n} - W_{t_{n-1}}$$

are independent and $W_{t_j} - W_{t_{j-1}}$ has the normal distribution with mean zero and variance $t_j - t_{j-1}$, that is,

$$P\left(W_{t_j} - W_{t_{j-1}} < x\right) = \int_{y=-\infty}^{x} \frac{1}{\sqrt{2\pi\left(t_j - t_{j-1}\right)}} \exp\left(\frac{-y^2}{2\left(t_j - t_{j-1}\right)}\right) dy$$

(2) The sample paths of W_t are almost surely continuous.

(3) $W_0 = 0$

An m-dimensional Brownian Motion is a stochastic process $B_t \in \mathbb{R}^m$ with independent components, each of which is a one-dimensional Brownian Motion. In the rest of this paper, we let $\mathcal{F}_t$ be the sigma-algebra generated by the sample path of B_t up to time t.

It should not be obvious that such a process exists, and indeed the first construction came many years after the process was first described. For details on its construction, see Lalley's Notes [2]. We now state an important lemma about the variation of Brownian motion which will be useful later.

Lemma 3.2 (Quadratic Variation of Brownian Motion). Let W_t be a one-dimensional Brownian Motion, and let $\{P_n\}_{n \in \mathbb{N}}$ be the nth dyadic rational partition of the interval $[t_1, t_2]$. That is,

$$P_n^j = t_1 + \frac{(t_2 - t_1)\, j}{2^n}.$$

for $j = 0, 1, \ldots, 2^n$. Then

$$\lim_{n\to\infty}\sum_{j=1}^{2^n}\left(W_{P_n^j}-W_{P_n^{j-1}}\right)^2=(t_2-t_1)\ \text{a.s.}$$

Proof. By the independent increments property for Brownian Motion, we know that $\left(W_{P_n^j}-W_{P_n^{j-1}}\right)^2$ are independent and identically distributed random variables. These random variable have a distribution well known in statistics, (chi-square), with expectation $\frac{(t_2-t_1)}{2^n}$. The law of large numbers then implies that the sum of all of these, which is the average of 2^n independent random variables, converges to (t_2-t_1) as $n\to\infty$.

Although Brownian Motion is fascinating in its own right, its primary interest to us will be in allowing us to define many other stochastic processes through the Itô integral.

Definition 3.3 (Itô Integral). Let $f:\Omega\times\mathbb{R}_0^+\to\mathbb{R}$ be an $\mathcal{F}_t$-adapted step stochastic process, that is, there exist $0\le t_0<t_1<\ldots<t_n<\infty$ such that

$$f(\omega,t)=\begin{cases} f_0(\omega) & \text{if } 0\le t<t_1\\ f_1(\omega) & \text{if } t_1\le t<t_2\\ \vdots & \vdots\end{cases}$$

where $f_i(\omega)$ is measurable at time t_i. Moreover, assume that $f\in L^2(\omega,t)$, that is,

$$\int_0^t E\left[f(s)^2\right]ds<\infty$$

Then the Itô integral of f with respect to the one-dimensional Brownian motion W_t is

$$\int_0^t f(\omega,s)dW_s=\left[\sum_{i=0}^{I(t)-1}f_i(\omega)\left(W_{t_i+1}-W_{t_i}\right)\right]+f_{I(t)}(\omega)\left(W_t-W_{t_{I(t)}}\right)$$

where $I(t)=\max\{j:t_j\le t\}$.

Let $g:\Omega\times\mathbb{R}\to\mathbb{R}\in L^2(\omega,t)$ be a measurable, $\mathcal{F}_t$-adapted stochastic process, (not necessarily a step-process), and let $\{f^n\}_{n\in\mathbb{N}}$ be a series of $\mathcal{F}_t$-adapted stepstochastic functions converging to g in L^2. That is,

$$\lim_{n\to\infty}\int_0^t E\left[|g(s)-f^n(s)|^2\right]ds=0$$

Then the Itô integral of g with respect to the one-dimensional Brownian motion W_t is

$$\int_0^t g(\omega, s) dW_s = \lim_{n\to\infty} \int_0^t f^n(\omega, s) dW_s$$

Finally, let $h : \Omega\times\mathbb{R} \to \mathbb{R}$ be a measurable, $\mathcal{F}_t$-adapted stochastic process in $L^2(t)$, i.e., with

$$\int_0^t |h(s)|^2 ds < \infty \text{ a.s.}$$

(but not necessary with finite expected square), and let $\{g^n\}_{n\in\mathbb{N}}$ be a series of $\mathcal{F}_{t^-}$ adapted functions in $L^2(\omega, t)$ converging to h in $L^2(t)$. Then the Itô integral of h with respect to the one-dimensional Brownian motion W_t is

$$\int_0^t h(\omega, s) dW_s = \lim_{n\to\infty} \int_0^t g^n(\omega, s) dW_s$$

The multidimensional Itô integral is defined component-wise: If $\sigma : \Omega \times \mathbb{R}_0^+ \to \mathbb{R}^{n\times m}$ is measurable, $\mathcal{F}_t$-adapted, and with integrated square almost surely finite, and $B_t \in \mathbb{R}^m$ is an m-dimensional Brownian motion, then

$$\int_0^t \sigma(\omega, s) dB_s = \begin{pmatrix} \sum_{j=1}^m \int_0^t \sigma_{1j}(\omega, s) dB_s^j \\ \vdots \\ \sum_{j=1}^m \int_0^t \sigma_{nj}(\omega, s) dB_s^j \end{pmatrix}$$

There are many other details involved in defining the Itô integral, such as showing that approximating functions always exist and that the limits do not depend on which sequence of functions is used. We refer the interested reader to Kuo's intuitive book [1].

Theorem 3.4 (Itô's Isometry). Let $g : \Omega \times \mathbb{R} \to \mathbb{R} \in L^2$ be a measurable, $\mathcal{F}_{t^-}$ adapted stochastic process. Then

$$E\left[\left(\int_0^t g(\omega, s) dB_s\right)^2\right] = E\left[\int_0^t g(\omega, s)^2 ds\right]$$

Proof. Let $\{f^n\}_{n\in\mathbb{N}}$ be a series of $\mathcal{F}_t$-adapted step-stochastic functions converging to g in L^2. By the quadratic variation formula for Brownian Motion, it follows that

$$E\left[\left(\int_0^t f^n(\omega,s)dB_s\right)^2\right] = E\left[\int_0^t f^n(\omega,s)^2 ds\right]$$

So

$$\begin{aligned}
E\left[\left(\int_0^t g(\omega,s)dB_s\right)^2\right] &= E\left[\left(\lim_{n\to\infty}\int_0^t f^n(\omega,s)dB_s\right)^2\right] \\
&= E\left[\lim_{n\to\infty}\int_0^t f^n(\omega,s)^2 ds\right] \\
&= E\left[\int_0^t g(\omega,s)^2 ds\right].
\end{aligned}$$

Definition 3.5 (Itô Process). An Itô process is a stochastic process $Z_t(\omega)$ satisfying an equation of the form

$$Z_t(\omega) - Z_0 = \int_{s=0}^t b(s,\omega)ds + \int_{s=0}^t \sigma(s,\omega)dB_s \tag{3.6}$$

where $Z_t \in \mathbb{R}^n, \omega \in \Omega, b : \mathbb{R}\times\Omega \to \mathbb{R}^n, \sigma : \mathbb{R}\times\Omega \to \mathbb{R}^{n\times m}, B_t$ is an m-dimensional Brownian motion, and $Z_0 = z.b$ is known as the drift coefficient and σ is known as the diffusion coefficient.

We often write (3.6) in differential form as

$$dZ_t(\omega) = b(t,\omega)dt + \sigma(t,\omega)dB_t$$

Theorem 3.7 (Itô's Lemma). Let $u(x,t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ be twice continuously differentiable with respect to x and once continuously differentiable with respect to t. Let $Z_t = \left(Z_t^1 \dots Z_t^n\right)$ be an n-dimensional Itô process. Then

$$du\left(Z_t,t\right) = \frac{\partial u\left(t,Z_t\right)}{\partial t}dt + \sum_{i=1}^n \frac{\partial u\left(t,Z_t\right)}{\partial x_i}dZ_t^i + \sum_{i,j=1}^n \frac{\partial^2 u\left(t,Z_t\right)}{\partial x_i \partial x_j}dZ_t^i dZ_t^j$$

where we compute $dZ_t^i dZ_t^j$ with

$$\begin{aligned}
dt \times dt &= 0 \\
dt \times dB_t^i &= 0 \\
dB_t^i \times dB_t^j &= \mathbf{1}_{\{i=j\}}.
\end{aligned}$$

Itô's Lemma is the equivalent of the chain rule in ordinary calculus. The importance of the 2nd order derivatives shows that Stochastic Calculus is fundamentally different from ordinary calculus. We will prove the special case where $Z_t = W_t$, the one-dimensional Brownian motion, and hope that this will at least convince the reader that the general result is plausible.

Proof.

$$\begin{aligned} u\left(W_t,t\right)-u\left(W_0,0\right) &= \sum_{j=1}^{2^n}\left[u\left(W_{jt/2^n},jt/2^n\right)-u\left(W_{(j-1)t/2^n},(j-1)t/2^n\right)\right] \\ &= \sum_{j=1}^{2^n}\left[u\left(W_{jt/2^n},jt/2^n\right)-u\left(W_{jt/2^n},(j-1)t/2^n\right)\right. \\ &\qquad \left.+u\left(W_{jt/2^n},(j-1)t/2^n\right)-u\left(W_{(j-1)t/2^n},(j-1)t/2^n\right)\right] \end{aligned}$$

We will examine the two terms separately, beginning with the first:

$$\begin{aligned} &\sum_{j=1}^{2^n}\left[u\left(W^{\partial t}_{jt/2^n,jt/2^n)-u\left(W_{jt/2^n},(j-1)t/2^n\right)\right]}\right.\right.\sum_{j=1}^{2^n}\left[\frac{\partial u\left(W_{jt/2^n},(j-1)t/2^n\right)}{\partial t}2^{-n}t+o\left(2^{-n}\right)\right] \\ &\xrightarrow{n\to\infty}\int_{s=0}^{t}\frac{\partial u\left(W_s,s\right)}{\partial t}ds \end{aligned}$$

As for the second term,

$$\begin{aligned} &\sum_{j=1}^{2^n}\left[u\left(W_{jt/2^n},(j-1)t/2^n\right)-u\left(W_{(j-1)t/2^n},(j-1)t/2^n\right)\right] \\ &\quad =\sum_{j=1}^{2^n}\left[\frac{\partial u\left(W_{(j-1)t/2^n},(j-1)t/2^n\right)}{\partial x}\left(W_{jt/2^n}-W_{(j-1)t/2^n}\right)\right. \\ &\qquad +\frac{1}{2}\frac{\partial^2 u\left(W_{(j-1)t/2^n},(j-1)t/2^n\right)}{\partial x^2}\left(W_{jt/2^n}-W_{(j-1)t/2^n}\right)^2 \\ &\qquad \left.+o\left(\left(W_{jt/2^n}-W_{(j-1)t/2^n}\right)^2\right)\right] \\ &\qquad \xrightarrow{n\to\infty}\int_0^t\frac{\partial u\left(W_s,s\right)}{\partial x}dW_s+\frac{1}{2}\int_0^t\frac{\partial^2 u\left(W_s,s\right)}{\partial x^2}ds \end{aligned}$$

by the definition of the Itô integral and by the Itô isometry.

4. Diffusions

In this chapter we slow down our pace considerably as we examine the probability models used in the Stochastic Control Problem.

Definition 4.1. A diffusion is an Itô process in which the drift and diffusion coefficients depend only upon time and the position of system, i.e., a process X_t satisfying

$$dY_t = b\left(Y_t, t\right) dt + \sigma\left(Y_t, t\right) dB_t$$

where $Y_t \in \mathbb{R}^n, b : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n, \sigma : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n \times m}, B_t$ is an m-dimensional Brownian motion, and $Y_0 = y$.

There is a simple intuition behind diffusions which gives them their name. A diffusion is a reasonable probabilistic model for the position of a particle in a fluid. X_t is the position of the particle at time $t, b(y,t)$ is the velocity of the fluid at the point y at time t, and $\sigma(y,t)$ describes the way in which the particle is randomly jostled by the fluid at the point y at time t. In particular, $\Sigma(y,t) := \sigma(y,t)\sigma(y,t)^T \in \mathbb{R}^{n \times n}$ is the instantaneous covariance matrix of the particle at the point y and time t, in units of $\frac{\text{distance}^{\;2}}{\text{time}}$.

Definition 4.2. A time-homogeneous diffusion is an Itô process in which the drift and diffusion coefficients depend only upon the position of system, i.e., a process X_t satisfying

$$dX_t = b\left(X_t\right) dt + \sigma\left(X_t\right) dB_t$$

where $X_t \in \mathbb{R}^n, b : \mathbb{R}^n \to \mathbb{R}^n, \sigma : \mathbb{R}^n \to \mathbb{R}^{n \times m}, B_t$ is an m-dimensional Brownian motion, and $X_0 = x$.

Intuitively, a time-homogeneous diffusion describes the position of a particle in a steady flow.

At first glance, time-homogeneous diffusions look like a special case of diffusions, but in fact, the two classes of stochastic processes are actually equivalent. In particular, any time-heterogeneous diffusion can be converted into a time-homogeneous
diffusion through the simple expedient of setting $X_t = (Y_t, t)$. The class of timehomogeneous diffusions is therefore rather broader than it might first appear. Consequently, we will focus mainly on time-homogeneous diffusions in the rest of this paper.

Because the drift and diffusion coefficients of a time-homogeneous diffusion do not depend upon its past positions, we can easily imagine that time-homogeneous diffusions are "memoryless" in the sense that the future of such a process depends only on its present state and not upon the past. We make this idea formal in the following definition and Lemma.

Definition 4.3 (Stopping Time). A stopping time is a function $\tau : \Omega \rightarrow [0, \infty]$ such that $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$ for all t.

Intuitively, a stopping time is a rule to choose a time based upon the progress of the Brownian Motion, (or the diffusion it determines). A common stopping time is the first exit time from a set $G, \tau_G = \inf\{t \geq 0 : X_t \notin G\}$. Stopping times allow us to formalize our idea of memorylessness in the following Lemma:

Lemma 4.4 (Strong Markov Property). Let τ be a stopping time, and X_t a timehomogeneous diffusion. Then

$$E^x\left[f\left(X_{\tau+t}\right) \mid \mathcal{F}_\tau\right] = E^{X_\tau}\left[f\left(X_t\right)\right] \tag{4.5}$$

where $0 \leq t < \infty$, and $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is bounded and measurable.

The result becomes more intuitive when $f(x) = \mathbf{1}_{\{x \in B\}}$ for a Borel set B. In this case, (4.5) reduces to:

$$P^x\left(X_{\tau+t} \in B \mid \mathcal{F}_\tau\right) = P^{X_\tau}\left(X_t \in B\right).$$

For the highly technical proof, we refer the reader to Øksendal [3]. This lemma should be intuitive if only for the following reason: Anyone wanting to simulate a diffusion would likely approximate continuous time by using small but discrete time increments and applying the definition of the Itô integral. Such an approximation is simply a discrete time Markov Chain on an uncountably infinite state space, so it should not be suprising that the limit of such processes is Markovian as well.

It is possible to prove even stronger results which say that the diffusion after a stopping time τ has the same distribution as if it had begun at X_τ instead of x.

We next prove some technical results that will be useful in later sections.

Theorem 4.6 (Dynkin's Formula). Let $Z_t(\omega)$ be an Itô process satisfying

$$dZ_t(\omega) = b(t, \omega)dt + \sigma(t, \omega)dB_t,$$

where $Z_t \in \mathbb{R}^n, \omega \in \Omega, b : \mathbb{R} \times \Omega \rightarrow \mathbb{R}^n, \sigma : \mathbb{R} \times \Omega \rightarrow \mathbb{R}^{n \times m}, B_t$ is an m-dimensional Brownian motion, and $Z_0 = z$.

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$, with $f \in C_0^2\left(\mathbb{R}^n\right)$, i.e., f is twice continuously differentiable with compact support. Let τ be a stopping time with finite expectation. Assume $\sigma(t, \omega)$ is bounded, at least on the set of (t, ω) where $f\left(Z_t(\omega)\right) \neq 0$. Then

$$E^z\left[f\left(Z_\tau\right)\right] =$$
$$f(z)+E^z\left[\int_0^\tau\left[b(t,\omega)^T\nabla f\left(Z_t\right)+\frac{1}{2}\sum_{i,j}H\left(Z_t\right)_{ij}\Sigma(t,\omega)_{ij}\right]dt\right]$$

where $\nabla f(z)$ is the gradient of f at $z, H(z)$ is the Hessian matrix of f at z, i.e., $H(z)_{ij} = \left.\frac{\partial^2 f}{\partial z_i \partial z_j}\right|_z$, and as before $\Sigma(t,\omega) = \sigma(t,\omega)\sigma(t,\omega)^T$ is the instantaneous covariance matrix of Z_t at time t.

Proof. The proof has two logical parts: first we apply Itô's Lemma to $df\left(Z_t\right)$, and second we adapt the result to $f\left(Z_\tau\right)$.

By Itô's Lemma,

$$\begin{aligned}
df\left(Z_t\right) =&\nabla f\left(Z_t\right)^T dZ_t+\frac{1}{2}\sum_{i,j}H\left(Z_t\right)_{ij}\left(dZ_t\right)_i\left(dZ_t\right)_j\\
=&\nabla f\left(Z_t\right)^T\left(b(t,\omega)dt+\sigma(t,\omega)dB_t\right)\\
&+\frac{1}{2}\sum_{i,j}H\left(Z_t\right)_{ij}\left(b(t,\omega)dt+\sigma(t,\omega)dB_t\right)_i\left(b(t,\omega)dt+\sigma(t,\omega)dB_t\right)_j\\
=&\nabla f\left(Z_t\right)^T b(t,\omega)dt+\nabla f\left(Z_t\right)^T\sigma(t,\omega)dB_t\\
&+\frac{1}{2}\sum_{i,j}H\left(Z_t\right)_{ij}\left(\sigma(t,\omega)dB_t\right)_i\left(\sigma(t,\omega)dB_t\right)_j\\
=&\nabla f\left(Z_t\right)^T b(t,\omega)dt+\nabla f\left(Z_t\right)^T\sigma(t,\omega)dB_t\\
&+\frac{1}{2}\sum_{i,j}H\left(Z_t\right)_{ij}\Sigma(t,\omega)dt.
\end{aligned}$$

So

$$\begin{aligned}
&E^z\left[f\left(Z_\tau\right)\right] =\\
&f(z)+E^z\left[\int_0^\tau b(t,\omega)^T\nabla f\left(Z_t\right)dt+\int_0^\tau\nabla f\left(Z_t\right)^T\Sigma(t,\omega)dB_t\right.\\
&\quad\left.+\frac{1}{2}\int_0^\tau\sum_{i,j}H\left(Z_t\right)_{ij}\Sigma(t,\omega)_{ij}dt\right].
\end{aligned}\tag{4.7}$$

Now let $\tau\wedge T=\min(\tau,T)$. By Doob's Optional Stopping Theorem,

$$E^z\left[\int_0^{\tau\wedge T}\nabla f\left(Z_t\right)^T\sigma(t,\omega)dB_t\right]=0$$

which makes the dB_t term in (4.7) vanish, proves the theorem for bounded stopping times τ, and concludes the first part of the proof. Since τ is in general not bounded, we must prove that

$$E^z\left[\int_0^{\tau}\nabla f\left(Z_t\right)^T\sigma(t,\omega)dB_t\right]=\lim_{T\to\infty}E^z\left[\int_0^{\tau\wedge T}\nabla f\left(Z_t\right)^T\sigma(t,\omega)dB_t\right]=0.$$

The key observation is that $\nabla f\left(Z_t\right)^T\sigma(t,\omega)$ is bounded since by assumption ∇f is continuous and nonzero only on a bounded set of $\mathbb{R}^n$, and by assumption $\sigma(t,\omega)$ is bounded wherever f is nonzero. So we may pick a bound D such that

$$\left|\nabla f\left(Z_t\right)^T\sigma(t,\omega)\right|\le D<\infty.$$

Then, using Itô's isometry, we get:

$$\begin{aligned}&\lim_{T\to\infty}E^z\left[\left(\int_0^{\tau}\nabla f\left(Z_t\right)^T\sigma(t,\omega)dt-\int_0^{\tau\wedge T}\nabla f\left(Z_t\right)^T\sigma(t,\omega)dt\right)^2\right]\\&=\lim_{T\to\infty}E^z\left[\int_{t=\tau\wedge T}^{\tau}\left(\nabla f\left(Z_t\right)^T\sigma(t,\omega)\right)^2dt\right]\\&\le\lim_{T\to\infty}(\tau-\tau\wedge T)D^2 &&=0.\end{aligned}$$

Combining this with (4.7) gives the result in general.

An immediate use for this theorem is in finding an explicit formula for a sort of stochastic derivative of the time-homogeneous diffusion X_t, called the generator of X_t :

Definition 4.8. The (infinitesimal) generator A of the time-homogeneous Itô diffusion X_t is defined as

$$Af(x)=\lim_{t\downarrow 0}\frac{E^x\left[f\left(X_t\right)\right]-f(x)}{t}\tag{4.9}$$

for $x\in\mathbb{R}^n$ and $f:\mathbb{R}^n\to\mathbb{R}$.

Theorem 4.10 (Formula for Generator). If $f\in C_0^2\left(\mathbb{R}^n\right)$, i.e., f is twice continuously differentiable with compact support, then the limit Af exists for all $x\in\mathbb{R}^n$ and

$$Af(x) = b(x)^T \nabla f(x) + \frac{1}{2} \sum_{i,j} \Sigma(x)_{ij} H(x)_{ij}$$

where $\nabla f(x)$ is the gradient of f at $x, H(x)$ is the Hessian matrix of f at x, i.e., $H(x)_{ij} = \left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_x$, and as before $\Sigma(x) = \sigma(x)\sigma(x)^T$ is the instantaneous covariance matrix of X_t at x.

Proof. This is a straightforward application of Dynkin's Formula, (Theorem 4.6), with $\tau = t$.

The first-order term in the generator formula represents change in f due to the drift in X_t. The second-order term, however, is perhaps not as intuitive. To get more insight into this term, consider the simplest diffusion, $X_t = B_t \in \mathbb{R}$, a Brownian motion in one dimension, and a function f such that $f(x) = x^2$ on an interval surrounding 0 , and elsewhere gracefully returning to 0 in such a way that $f \in C_0^2$. Begin at $X_t = 0$, and wait a small interval of time t. Then $\frac{f(X_t)}{t}$ is approximately distributed as a chi-square random variable with one degree of freedom, and so $f(X_t)$ has expectation approximately equal to t. Roughly speaking, the generator is:

$$\begin{aligned} Af(0) &= \lim_{t \downarrow 0} \frac{E\left[f\left(X_t\right)\right] - f(0)}{t} \\ &= \lim_{t \downarrow 0} \frac{t - 0}{t} \\ &= 1 \end{aligned}$$

which is the same as would have resulted from evaluating the generator formula. This movement is due not to a systemic drift of the Brownian Motion, but rather to its random fluctuations combined with the curvature of the function f.

Dynkin's Formula also proves a nice corollary:

Corollary 4.11. Suppose $f \in C_0^2(\mathbb{R}^n)$, and τ is a stopping time with $E^x[\tau] < \infty$. Then

$$E^x\left[f\left(X_\tau\right)\right] = f(x) + E^x\left[\int_0^\tau Af\left(X_t\right)dt\right]$$

Generators are useful because they are a gateway between time-homogeneous diffusions and differential equations. If we know the generator of a time-homogeneous diffusion at all points x and for all C_0^2 functions f, then we can recover b and Σ at all points x by choosing the function f expediently, (for example, make all its second partial derivates zero and all its first partial derivates zero except for in component j to recover $b_j(x)$). The ability to recover b and Σ means that we can recover the probability law of X_t. Thus we have a differential operator which encodes all the important information about our diffusion. We will make good use out of generators in our final section.

We are now ready to begin defining the Stochastic Control Problem. First, we need a model for a controlable random system X_t. We let X_t be like a diffusion, except that the the drift and diffusion coefficients also depend upon a vector of control parameters u_t. Thus

$$dX_t = b\left(X_t, u_t\right)dt + \sigma\left(X_t, u_t\right)dB_t \tag{5.1}$$

where $u_t \in \mathbb{R}^p$ and is $\mathcal{F}_t$-adapted, and the drift and diffusion functions, $b(y, v)$ and $\sigma(y, v)$, are known a priori.

For example, X_t might represent the position, velocity, and orientation of an airplane at time t, and u_t the values of the airplane controls at time t, (like the throttle, ailerons, and rudder positions). $b(y, v)$ would then represent the mean change over time in the state of the airplane at state y and with controls v, and $\sigma(y, v)$ would describe the random changes in the system at state y and controls v, with $\Sigma(y, v)$ the instantaneous covariance matrix at state y and controls v, as before. Any control u_t we choose at time t must be $\mathcal{F}_t$-adapted, because the only information we can use to make the control decision at time t is the information that is available at time t. Upon fixing an adapted control u, X_t is an Itô process.

We observe first that in many problems if X_t strays from a particular set of values then the problem ends, in the sense that no changes of the controls will affect the performance. For example, if X_t represents the airplane, then after the plane lands successfully (or crashes!) the problem is over. Or if $X_t \in \mathbb{R}$ represents the cash value of a portfolio, then if $X_t < 0$ the problem is over because the owner is bankrupt. We therefore introduce a subset $G \subset \mathbb{R}^n$ of the space of states of X_t in which the problem continues, and a stopping time $\tau_G = \inf \{t \geq 0 : X_t \notin G\}$ at which point the problem ends. To avoid trivialities, we assume that $X_0 \in G$.

In order to say that one control is superior to another, we need a utility function W describing how preferable the path $\{X_t, u_t\}_{0 \leq t \leq \tau_G}$ is. For full generality, nothing would be assumed about W, but as often occurs in mathematics, a few reasonable

simplifications can make a difficult problem tractable with little cost in usefulness. We therefore assume unapologetically that the utility function W is of the form

$$W\left(\{X_t, u_t\}_{0 \leq t \leq \tau_G}\right) = \int_0^{\tau_G} f\left(X_t, u_t\right) dt + g\left(X_{\tau_G}\right) \mathbf{1}_{\{\tau_G < \infty\}}$$

for continuous, known functions $f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$, and $g : \mathbb{R}^n \to \mathbb{R}$.

The function $f(x, v)$ represents the utility rate of the system at x with controls v, and the function $g(x), x \in \partial G$ represents the utility of being at state x when the problem ends. If $\tau_G = \infty$, then this term is not included in the utility function. In the airplane example, $f(x, v)$ might represent the fuel costs associated with the velocity and altitude of the plane in state x and with controls v, and $g(x)$ would describe how bad it is to crash and how good it is to land safely.

We want controls u that generally produce high utility paths $\{X_t, u_t\}_{0 \leq t \leq \tau_G}$. We thus define the performance of the control u with $X_0 = x$ as

$$\begin{aligned} J^u(x) &= E^x\left[W\left(\{X_t, u_t\}_{0 \leq t \leq \tau_G}\right)\right] \\ &= E^x\left[\int_0^{\tau_G} f\left(X_t, u_t\right) dt + g\left(X_{\tau_G}\right) \mathbf{1}_{\{\tau_G < \infty\}}\right]. \end{aligned}$$

We assume that $J^u(x)$ is finite for all choices of $u \in \mathbb{R}^p$ and $x \in G$ so that performance can be compared.

We can now state the stochastic control problem formally.

Definition 5.2 (Stochastic Control Problem). Given a controlable process $X_t \in \mathbb{R}^n$ satisfying

$$dX_t = b\left(X_t, u_t\right) dt + \sigma\left(X_t, u_t\right) dB_t$$

a set $G \subset \mathbb{R}^n$ with $X_0 \in G$, a utility rate $f(x, u)$ for points $x \in G$ and controls $u \in \mathbb{R}^p$, and a conclusion utility $g(x)$ for boundary points $x \in \partial G$, find the best possible performance

$$\Phi(x) := \sup_{u_t(\omega)} J^u(x) = \sup_{u_t(\omega)} E^x\left[\int_0^{\tau_G} f\left(X_t, u_t\right) dt + g\left(X_{\tau_G}\right) \mathbf{1}_{\{\tau_G < \infty\}}\right]$$

and an optimal control u^* such that

$$J^{u^*}(x) = \Phi(x)$$

at all points x.

6. The Hamilton-Jacobi-Bellman Equation

The controlable system X_t does not directly depend on past states, and the utility function W assigns utility to a state x with controls v at the rate $f(x,v)$ at time t, regardless of the past states and controls at times $s < t$. This Markovian theme suggests that the optimal control $u_t^*(\omega)$ might depend only upon the state of the system X_t, i.e., $u_t^*(\omega) = u_t^*\left(X_t(\omega)\right)$, a conjecture we will later prove. With this motivation, we examine only Markov controls in this section, those for which

$$u_t(\omega) = u_t\left(X_t(\omega)\right)$$

Let $\Phi_M(x) = \sup\left\{J^u(x) : u \text{ Markov control }\right\}$, and for an adapted control v and a function $h : \mathbb{R}^n \to \mathbb{R}, h \in C_0^2\left(\mathbb{R}^n\right)$, let

$$\left(L^v h\right)(x) = b(x,v)^T \nabla h(x) + \frac{1}{2} \sum_{i,j} \Sigma(x,v)_{ij} H(x)_{ij}$$

the generator introduced in 4.10.

Theorem 6.1 (Hamilton-Jacobi-Bellman Necessary Conditions). If

(1) $\Phi_M \in C^2(G) \cap C(\bar{G})$

(2) For all bounded stopping times $\alpha \leq \tau_G$, for all points $x \in G$, and for all controls $v \in \mathbb{R}^p$,

$$E^x\left[\Phi_M\left(X_\alpha\right) + \int_0^\alpha \left|L^v \Phi_M\left(X_t\right)\right| dt\right] < \infty$$

(3) An optimal Markov control u^* exists

(4) ∂G is regular for $X_t^{u^*}$, that is, if $X_0 \in \partial G$, then $\tau_G = 0$ a.s. then

(1) $\Phi_M(x) = g(x)$ for all boundary points $x \in \partial G$

(2) For all points $x \in G$,

$$\sup_{v \in \mathbb{R}^p} \left\{f(x,v) + \left(L^v \Phi_M\right)(x)\right\} = 0$$

(3) For all points $x \in G$, the supremum above is attained at the optimal Markov control u^* :

$$f(x,u^*) + \left(L^{u^*}\Phi_M\right)(x) = 0$$

Before proving the theorem, we note that conclusions 2 and 3 have a strong intuition behind them: $(L^u\Phi_M)(x)$ is the rate of expected change in the optimal performance at point x under control u. If $u = u^*$, then the rate of expected increase in the optimal performance should be equal to the negative utility rate, $-f(x,u^*)$, but if u is not optimal, the rate of expected increase in the optimal performance is less than the negative utility rate, because we are not being compensated enough for moving.

Proof. The conclusion that $\Phi_M(x) = g(x)$ on the boundary ∂G, (conclusion 1), is a direct consequence of the regularity of ∂G, (hypothesis 4). Since $\tau_G = 0$ a.s. under the optimal control u^* for all $x \in \partial G$,

$$\begin{aligned}
\Phi_M(x) &= E^x\left[\int_0^{\tau_G} f(X_t,u_t^*)\,dt + g(X_{\tau_G})\,1_{\{\tau_G<\infty\}}\right] \\
&= E^x\left[\int_0^{0} f(X_t,u_t^*)\,dt + g(X_0)\right] \\
&= g(X_0)\,.
\end{aligned}$$

To prove conclusions 2 and 3 , we will define a bounded stopping time $\alpha \leq \tau_G$, before which time we will use an arbitrary control v, and after which time we will use the optimal control u^*. By exploiting the strong Markov property and taking $\alpha \to 0$, we will be able to use a continuity argument to derive our result.

Formally, let $\tau_G \wedge T = \min(\tau_G, T)$, for some constant time $0 \leq T < \infty$, and let

$$u_t(x) = \begin{cases} v_t & \text{if } t \leq \tau_G \wedge T \\ u_t^* & \text{if } t > \tau_G \wedge T \end{cases} \tag{6.2}$$

Then by the Strong Markov Property,

$$\begin{aligned}
J^u(x) &= E^x\left[\int_0^{\tau_G} f(X_t,u_t)\,dt + g(X_{\tau_G})\,1_{\{\tau_G<\infty\}}\right] \\
&= E^x\left[\int_0^{\tau_G\wedge T} f(X_t,v_t)\,dt\right] + E^x\left[J^{u^*}(X_{\tau_G\wedge T})\right] \\
&= E^x\left[\int_0^{\tau_G\wedge T} f(X_t,v_t)\,dt\right] + E^x\left[\Phi_M(X_{\tau_G\wedge T})\right].
\end{aligned} \tag{6.3}$$

We then apply Dynkin's Formula to the second term in (6.3) above:

$$E^x\left[\Phi_M(X_{\tau_G\wedge T})\right] = \Phi_M(x) + E^x\left[\int_0^{\tau_G\wedge T} (L\Phi_M)(X_t)\,dt\right]$$

and substitute to get

$$J^u(x) = E^x\left[\int_0^{\tau_G} f\left(X_t, v_t\right) dt\right] + \Phi_M(x) + E^x\left[\int_0^{\tau_G \wedge T} \left(L\Phi_M\right)\left(X_t\right) dt\right].$$

By definition,

$$\Phi_M(x) \geq J^u(x) \tag{6.4}$$

SO

$$\Phi_M(x) \geq E^x\left[\int_0^{\tau_G \wedge T} f\left(X_t, v_t\right) dt\right] + \Phi_M(x) + E^x\left[\int_0^{\tau_G \wedge T} \left(L\Phi_M\right)\left(X_t\right) dt\right]$$

or
or

$$0 \geq E^x\left[\int_0^{\tau_G \wedge T} f\left(X_t, v_t\right) dt\right] + E^x\left[\int_0^{\tau_G \wedge T} \left(L\Phi_M\right)\left(X_t\right) dt\right]$$

$$0 \geq \frac{E^x\left[\int_0^{\tau_G \wedge T}\left[f\left(X_t, v_t\right) + \left(L\Phi_M\right)\left(X_t\right)\right] dt\right]}{E^x\left[\tau_G \wedge T\right]}$$

Now take time $T \to 0$, so that $\tau_G \wedge T = \min\left(\tau_G, T\right) \to 0$. By the continuity of f and $(L\Phi_M\ )$, (the latter since $\Phi_M \in C^2(G)$, (hypothesis 1) and L is a 2^{nd} -order differential operator), it must be that

$$0 \geq f(x, v(x)) + \left(L\Phi_M\right)(x) \tag{6.5}$$

for all Markov controls v. However, if $v = u^*$, then u (defined in (6.2)) is identically u^*. As a result, (6.4) holds with equality, i.e.,

$$\Phi_M(x) = J^{u^*}(x).$$

Following the same reasoning as previously, we find that (6.5) also holds with equality, i.e.,

$$0 = f\left(x, u^*(x)\right) + \left(L\Phi_M\right)(x).$$

This is conclusion 3. Combining this result with the inequality (6.5), we obtain conclusion 2.

We should emphasize that this theorem only provides necessary conditions for a Markov control to be optimal, but not sufficient conditions. It is analogous to the theorem that at a local maximum, a C^1 function $f : \mathbb{R}^n \to \mathbb{R}$ has a zero gradient, even though local minima and saddle points also have zero gradients. We therefore turn our attention to sufficient conditions.

Theorem 6.6 (Hamilton-Jacobi-Bellman Sufficient Conditions). Suppose ϕ : $\mathbb{R}^n \to \mathbb{R}$ is such that

(1) $\phi \in C^2(G) \cap C(\bar{G})$
(2) For every point $x \in G$ and control $v \in \mathbb{R}^p$,

$$f(x, v) + (L^v \phi)(x) \leq 0$$

(3) For every starting point $X_0 = x$,

$$\lim_{t \to \tau_G} \phi(X_t) = g(X_{\tau_G}) \mathbf{1}_{\{\tau_G < \infty\}} \text{ a.s.}$$

(4) For every Markov control v and starting point $X_0 = x$, the set of random variables

$$\{\max(-\phi(X_\tau), 0) : \tau \text{ stopping time, } \tau \leq \tau_G\} \text{ is uniformly integrable} \tag{6.7}$$

then $\phi(x) \geq J^v(x)$ for all points $x \in G$ and controls $v \in \mathbb{R}^p$. If we have also found a candidate optimal Markov control $u(x)$ such that
(1) For every point $x \in G$

$$f(x, u) + (L^u \phi)(x) = 0$$

(2) For the candidate u, at each starting point $X_0 = x$, the set of random variables

$$\{\phi(X_\tau) : \tau \text{ stopping time, } \tau \leq \tau_G\} \text{ is uniformly integrable} \tag{6.8}$$

then

$$\phi = \Phi_M \text{ and } u \text{ is an optimal Markov control, i.e. } J^u(x) = \Phi_M(x).$$

Proof. Let $\tau_G^T = \min(\tau_G, T, \inf\{t > 0 : |X_t| > T\})$. By Dynkin's Formula, at any control v,

$$E^x\left[\phi\left(X_{\tau_G^T}\right)\right] = \phi(x) + E^x\left[\int_0^{\tau_G^T} (L^v \phi)(X_t)\, dt\right]$$

and since by assumption

$$(L^v \phi)(x) \leq -f(x, v), \tag{6.9}$$

a simple rearrangement gives

$$\begin{aligned}
\phi(x) &\geq E^x\left[\int_0^{\tau_G^T} f(X_t, v_t)\, dt\right] + E^x\left[\phi\left(X_{\tau_G^T}\right)\right] \\
&= E^x\left[\int_0^{\tau_G^T} f(X_t, v_t)\, dt + \phi\left(X_{\tau_G^T}\right)\right].
\end{aligned}$$

This holds for all times T, so it also holds in the limit infimum:

$$
\begin{aligned}
\phi(x) &\geq \liminf_{T\to\infty} E^x\left[\int_0^{\tau_G^T} f\left(X_t, v_t\right) dt + \phi\left(X_{\tau_G^T}\right)\right] \\
&\geq E^x\left[\liminf_{T\to\infty}\left[\int_0^{\tau_G^T} f\left(X_t, v_t\right) dt + \phi\left(X_{\tau_G^T}\right)\right]\right] \\
&= E^x\left[\int_0^{\tau_G} f\left(X_t, v_t\right) dt + g\left(X_{\tau_G}\right)\mathbf{1}_{\{\tau_G<\infty\}}\right] \\
&= J^v(x).
\end{aligned} \tag{6.10}
$$

Thus

$$
\phi(x) \geq \sup_{v\in\mathbb{R}^p} J^v(x) = \Phi_M(x). \tag{6.11}
$$

The movement of the limit infimum through the expectation in (6.10) is justified by Fatou's Lemma with the uniform integrability conditions in (6.7).

If, however, $v = u$, the candidate for an optimal Markov control, then (6.9) holds with equality, and instead of taking the limit infimum we take the limit, moving it through the expectation in (6.10) by the Vitali Convergence Theorem and the uniform integrability conditions (6.8). We then find that

$$
\phi(x) = J^u(x)
$$

which combined with (6.11) gives the result.

Finally, we turn our attention to the theorem alluded to at the beginning of this section. This theorem justifies our examination of Markov controls in this section, confirming our intuition that a Markovian system with a "Markovian" utility function should have an optimal $\mathcal{F}_t$-adapted control that is also Markovian.

Theorem 6.12 (Optimality of Markov Controls). If

(1) $\Phi_M \in C^2(G) \cap C(\bar{G})$

(2) For all bounded stopping times $\alpha \leq \tau_G$, for all points $x \in G$, and for all adapted controls v, (not only Markov controls),

$$
E^x\left[\Phi_M\left(X_\alpha\right) + \int_0^\alpha \left|L^v \Phi_M\left(X_t\right)\right| dt\right] < \infty
$$

(3) Φ_M is bounded over all x,

$$
\left|\Phi_M(x)\right| \leq D \tag{6.13}
$$

then $\Phi_M = \Phi$.

Proof. The proof is similar to that of the sufficient conditions: As before, let $\tau_G^T = \min\left(\tau_G, T, \inf\left\{t > 0 : |X_t| > T\right\}\right)$. Then by Dynkin's Formula, for all adapted controls v and points $x \in G$,

$$E^x\left[\Phi_M\left(X_{\tau_G^T}\right)\right]=\Phi_M(x)+E^x\left[\int_0^{\tau_G^T}\left(L^v\Phi_M\right)(X_t)\,dt\right]$$

But by the necessary conditions proved earlier, at each point x,

$$\left(L^v\Phi_M\right)(x)\le -f(x,v)$$

for all $v\in\mathbb{R}^p$, regardless of how v is chosen. So

$$\Phi_M(x)\ge E^x\left[\Phi_M\left(X_{\tau_G^T}\right)+\int_0^{\tau_G^T}f\left(X_t,v_t\right)dt\right]$$

This is also true in the limit, which we will take through the expectation by the Bounded Convergence Theorem together with boundedness of Φ_M, (hypothesis 3):

$$\begin{aligned}\Phi_M(x)&\ge\lim_{T\to\infty}E^x\left[\Phi_M\left(X_{\tau_G^T}\right)+\int_0^{\tau_G^T}f\left(X_t,v_t\right)dt\right]\\&\ge E^x\left[\lim_{T\to\infty}\left[\Phi_M\left(X_{\tau_G^T}\right)+\int_0^{\tau_G^T}f\left(X_t,v_t\right)dt\right]\right]\\&=E^x\left[\Phi_M\left(X_{\tau_G}\right)1_{\{\tau_G<\infty\}}+\int_0^{\tau_G}f\left(X_t,v_t\right)dt\right]\\&=J^v(x).\end{aligned}$$

Therefore

$$\Phi_M(x)\ge\Phi(x).$$

Having proved these three theorems about optimal stochastic control, it is rather gratifying to note that we have also solved the deterministic control problem as a special case! The deterministic control problem is the same as that developed in the previous section, except that $\sigma=0$ in (5.1), i.e., the system X_t satisfies the deterministic differential equation

$$dX_t=b\left(X_t,u_t\right)dt$$

Much simplification ensues. Firstly, the generator A defined in (4.9) simplifies to become a simple first-order differential operator. In particular,

$$Af(x)=\nabla f(x)^Tb(x)$$

The added uniform integrability and boundedness conditions in the sufficient conditions and the Markov Control theorems, (equations (6.7), (6.8), and (6.13)), are designed to pull limits through expectation operators, which is no longer an issue under deterministic control because performance is also deterministic. So if we can find a solution to the equation in the necessary conditions, that solution is automatically optimal!

Theorem 6.14 (Deterministic Hamilton-Jacobi-Bellman Equations). If

(1) $\Phi \in C^2(G) \cap C(\bar{G})$

(2) For all times $T \leq \tau_G$, for all points $x \in G$, and for all controls $v \in \mathbb{R}^p$,

$$\Phi(X_T) + \int_0^T \left| \nabla \Phi(X_t)^T b(x, v) \right| dt < \infty$$

(3) An optimal Markov control u^* exists

(4) ∂G is regular for $X_t^{u^*}$, that is, if $X_0 \in \partial G$, then $\tau_G = 0$

then a solution $(\Phi(x), u^*(x))$ to the equations

$$\Phi(x) = g(x) \quad \forall x \in \partial G$$

and

$$f(x, v) = -\nabla \Phi(x)^T b(x) \quad \forall x \in G$$

exists and that solution is the optimal performance and an optimal control. Moreover, for all points $x \in G$,

$$\sup_{v \in \mathbb{R}^p} \left\{ f(x, v) + \nabla \Phi(x)^T b(x) \right\} = 0$$

7. Acknowledgements

I'd like to thank my mentors, Dr. Ramakhrishna and Dr. Alain, for their enthusiasm and generosity throughout this research.

REFERENCES

[1] H.H. Kuo. Introduction to stochastic integration. Springer Verlag, 2006.

[2] Steven Lalley. Notes for stat 385 and stat 390.

[3] B.K. Øksendal. Stochastic differential equations: an introduction with applications. Springer Verlag, 2003.

[4] D. Williams. Probability with martingales. Cambridge Univ Pr, 1991.