12 Itô Calculus and Controlled Diffusion Process

Let $\mathcal{P}_t(x, A) = \mathbf{P}[X_t \in A | X_0 = x]$ be the transition probability kernel of a diffusion process $\{X_t\}$ in $\mathbb{R}^n$. Recall that a diffusion process is defined by the following two properties:

1. For small t, X_t is concentrated around a ball of radius r centered at X_0:

$$\mathcal{P}_t(x, \mathbb{B}^n(x, r)^c) = o(t),$$

where $\mathbb{B}^n(x, r) := \{x' \in \mathbb{R}^n | \|x' - x\| \leq r\}$.

2. Bounded local mean and covariance:

$$\begin{aligned}
&\exists b(x) \in \mathbb{R}^n, A(x) = A^\top(x) \succeq 0 \quad \text{s.t.} \\
&\lim_{h \to 0} \frac{1}{h} \mathbf{E}[(X_h - X_0) \mathbb{I}_{\{\|X_h - X_0\| \leq r\}} | X_0 = x] = b(x) \\
&\lim_{h \to 0} \frac{1}{h} \mathbf{E}[(X_h - X_0)(X_h - X_0)^\top \mathbb{I}_{\{\|X_h - X_0\| \leq r\}} | X_0 = x] = A(x),
\end{aligned}$$

where $b(x)$ is called the local *drift coefficient* and $A(x)$ is called the *diffusion matrix*.

Formally, the two properties imply the following recursive relation:

$$X_{t+dt} = X_t + b(X_t)dt + \sigma(X_t)(W_{t+dt} - W_t),$$

where $A(x) = \sigma(x)\sigma(x)^\top$ and $W_{t+dt} - W_t$ has the same law as $\sqrt{dt}Z, Z \sim \mathcal{N}(0, I_n)$. The expression is useful for numerical simulation by choosing a proper step size dt.

To facilitate a more precise definition of diffusion process in term of stochastic differential equation (SDE), we need the definition of filtration and adapted process, which intuitively constrains the current event of a random process to depend only on the previous events.

Definition 12.1 (Filtration and adapted process) *Let the probability space be defined by the Kolmogorov tuple $(\Omega, \mathcal{F}, \mathbb{P})$ and $\{\mathcal{F}_t\}_{t \geq 0}$ be a sequence of sub-σ algebra of $\mathcal{F}$. $\{\mathcal{F}_t\}_{t \geq 0}$ is called a filtration if, for any $t < s$, $\mathcal{F}_t \subseteq \mathcal{F}_s \subseteq \mathcal{F}$. (There are also additional regularity conditions, like right-continuity, that we will not worry about here.) Further, a random process $\{X_t\}_{t \geq 0}$ is called an adapted process to $\{\mathcal{F}_t\}_{t \geq 0}$ if X_t is $\mathcal{F}_t$-measurable for each t.*

Recall that an ordinary differential equation (ODE) can be written equivalently as an integral equation:

$$\begin{aligned}
dX_t &= X_{t+dt} - X_t \\
&= b(X_t)dt \\
X_t &= X_0 + \int_0^t b(X_s)ds.
\end{aligned}$$

Similarly, we can write the diffusion SDE in the following form, introduced by Kiyoshi Itô:

$$\begin{aligned} dX_t &= X_{t+dt} - X_t \\ &= b(X_t)dt + \sigma(X_t)(W_{t+dt} - W_t) = b(X_t)dt + \sigma(X_t)dW_t \\ X_t &= X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s, \end{aligned}$$

where the last term is a *stochastic integral* or *Itô integral*. The reason we write SDE in this form is because the noise W_t is in general not differentiable with respect to time. Also notice that the integral is well-defined for an adapted process under some regularity condition, as will be shown in the next section. Further, the *infinitesimal generator* of the diffusion SDE for a suitable function $f : \mathbb{R}^n \to \mathbb{R}$ can be expressed in the following form, called the *second-order diffusion operator*:

$$\mathcal{A}f(x) = b(x)^\top \nabla f(x) + \frac{1}{2}\mathbf{tr}(A(x)\nabla^2 f(x)),$$

where $\mathcal{A}$ can be viewed as the probability intensity operator for the diffusion process $\{X_t\}_{t\geq 0}$ if f is the probability density function at any time. As a result,

$$\mathbf{E}[f(X_T)|X_0 = X] = f(x) + \mathbf{E}[\int_0^t \mathcal{A}f(X_s)ds].$$

12.1 Stochastic integration

Let $\{X_t\}_{t\geq 0}$ be an adapted process and $\{W_t\}_{t\geq 0}$ be a standard Brownian motion ($b(x) = 0, A(x) = I$). If X_t is constant on intervals $[t_{i-1}, t_i)$ with $0 = t_0 < t_1 < \ldots < t_n = t$, then the Itô integral can be defined as follows: $0 \leq t_0 < t_1 \ldots < t_n = t$:

$$\begin{aligned} I_t &= \int_0^t X_s dW_s \\ &:= \sum_{i=0}^n X(t_i)(W_{t_i+1} - W_{t_i}). \end{aligned}$$

It can be shown that the Itô integral has zero mean, using the fact that $W_{t_{i+1}} - W_{t_i}$ is independent of X_s and W_s, $\forall s \leq t_i$:

$$\begin{aligned} \mathbf{E}[I] &= \sum_{i=0}^n \mathbf{E}[X(t_i)(W_{t_{i+1}} - W_{t_i})] \\ &= \sum_{i=0}^n \mathbf{E}[X(t_i)]\mathbf{E}[W_{t_{i+1}} - W_{t_i}] \\ &= 0, \end{aligned}$$

Further, the second moment of the integral can be shown to be:

$$\mathbf{E}[I^2] = \mathbf{E}[X_{t_i}^2]\mathbf{E}[(W_{t_{i+1}} - W_{t_i})^2]$$

$$
\begin{aligned}
&= \sum_{i=0}^{n} \mathbf{E}[X_{t_i}^2 (W_{t_{i+1}} - W_{t_i})] + \sum_{i=0}^{n} \sum_{j \neq i} \mathbf{E}[X_{t_i} X_{s_j} (W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j})] \\
&= \int \mathbf{E}[X_s^2] ds + 0 \\
&= \mathbf{E}[\int X_s^2 ds],
\end{aligned}
$$

where the second-to-last equality uses the fact that $W_{t_{i+1}} - W_{t_i}$ is independent of $W_{s_{i+1}} - W_{s_i}, X_{t_i}, X_{s_j}$ if $s_{i+1} < t_i$; if $t_{i+1} < s_i$, $W_{s_{i+1}} - W_{s_i}$ will be independent of $W_{t_{i+1}} - W_{t_i}, X_{t_i}, X_{s_j}$. Then the integral can be extended to arbitrary adapted processes satisfying $\mathbf{E}\left[\int_0^t W_s^2 ds\right] < \infty$ using a limiting procedure. A related property of Itô integral is that $\{I_t\}_{t \geq 0}$ forms a *Martingale* defined as follows.

Definition 12.2 (Martingale) *For the probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$*,* $\{M_t\}$ *is a Martingale with respect to* $\{\mathcal{F}_t\}$ *if* $\forall s < t$*,* $\mathbf{E}[M_t | \mathcal{F}_s] = M_s$*.*

The Martingale property of Itô process follows from the fact that $\mathbf{E}[I_t(\mathcal{F}_s)] = I_s(\mathcal{F}_s) + \mathbf{E}[I_t(\mathcal{F}_s) - I_s(\mathcal{F}_s)] = I_s(\mathcal{F}_s) + [I_t(\mathcal{F}_s) - I_s(\mathcal{F}_s)] = I_s(\mathcal{F}_s)$ using the independence property of Brownian motion. Therefore, in the diffusion process, $M_t := X_t - X_0 - \int_0^t b(X_s) ds = \int_0^t \sigma(X_s) dW_s$ is a Martingale.

12.2 Preview on controlled diffusion process

So far we have not applied any control on the diffusion process. Suppose now we can control the drift coefficient and diffusion coefficient of the process and let $b : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}^n, A : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}^{n \times m}, A(x, u) = \sigma(x, u)\sigma(x, u)^\top$, the SDE for *controlled diffusion process* is then defined as:

$$dX_t = b(X_t, u)dt + \sigma(X_t, u)dW_t. \tag{1}$$

If the control actions $\{U_s\}_{s=0}^T$ is an adapted process, the Itô integral will be well-defined. Therefore for adapted control, Eq. (1) is equivalent to:

$$X_t = X_0 + \int_0^t b(X_s, U_s)ds + \int_0^t \sigma(X_s, U_s)dW_s.$$

Fix the control $U = \{u_s\}_{s=0}^T$, the SDE reduces to a diffusion process:

$$dX_t^U = b(X_t^U, u_t) + \sigma(X_t^U, u_t)dW_s.$$

The finite-horizon expected cost, cost-to-go and value function for the controlled diffusion process are then defined as follows:

$$
\begin{aligned}
J_T(g) &:= \mathbf{E}^g[\int_0^T c_t(X_t, U_t)dt + c_T(X_T)] \\
J_t(x; g) &:= \mathbf{E}^g[\int_t^T c_t(X_t, U_t)dt + c_T(X_T) | X_t = x] \\
V_t(x) &:= \min_g J_t(x; g).
\end{aligned}
$$

As will be shown later, the value function for the controlled Markov process turns out to satisfy the Hamilton-Jacobi-Bellman equation:

$$\begin{aligned}\frac{\partial V_t(x)}{\partial x} &= \min_{u\in\mathcal{U}}\{c_t(x,u) + b(x,u)^\top \nabla_x V_t(x) + \frac{1}{2}\mathbf{tr}(\sigma(x,u)\sigma(x,u)^\top \nabla_x^2 V_t(x))\} \\ &=: \mathcal{A}^u V_t.\end{aligned}$$

Further, we have:

$$\begin{aligned}X_{t+dt} &= X_t + b(X_t,u)dt + \sigma(X_t,U_t)dt \\ X_{t+dt} &\sim \mathcal{N}(X_t + b(X_t,u)dt, \sigma(X_t,u)\sigma(X_t,u)^\top),\end{aligned}$$

which can be discretized as:

$$X_{k+1} = X_k + hb(X_k) + \sqrt{h}\sigma(X_k,u_k)Z_k.$$

12.3 Controlled diffusion processes

Before we formally introduce the controlled diffusion processes, we first quickly go over the definition of a general (uncontrolled) diffusion process, which takes the following form

$$dX_t = b(X_t,t)dt + \sigma(X_t,t)dW_t \tag{2}$$

with initial condition $X_0 = x$ and time $0 \le t \le T$. In some scenarios, (2) can also be written as

$$X_{t+dt} = X_t + b(X_t,t)dt + \sigma(X_t,t)dW_t, \tag{3}$$

where $b(X_t,t)$ and $\sigma(X_t,t)$ are respective the deterministic drift term and diffusion matrix, which can be time-varying and $A(x) = \sigma(x)\sigma(x)^\top$ for the drift, and $dW_t = W_{t+dt} - W_t$. In order to simulate the uncontrolled diffusion processes (2) and (3) on computer, we usually fix a small time step $h > 0$ and consider the following discrete/numerical equation based on Euler scheme

$$X_{(k+1)h} = X_{kh} + b(X_{kh},kh)h + \sigma(X_{kh},kh)\cdot\sqrt{h}\xi_k, \tag{4}$$

where $\xi_k \overset{\text{i.i.d.}}{\sim} N(0,I_m)$ and diffusion matrix $\sigma(X_{kh},kh) \in \mathbb{R}^{n\times m}$. In the limit of infinitesimal h in (4), we can restore the analytical solution to (2) and (3)

$$X_t = X_0 + \int_0^t b(X_s,s)\ ds + \int_0^t \sigma(X_s,s)dW_s. \tag{5}$$

The following picture gives the idea of (4) for simple 1-dimensional case

We now formally introduce the controlled diffusion processes. Consider the state space $\mathcal{X} \in \mathbb{R}^n$ and control space $\mathcal{U} \in \mathbb{R}^m$, and the deterministic drift term $b(x,t)$ and diffusion matrix $\sigma(x,t)$ become $b(x,u,t)$ and $\sigma(x,u,t)$. Two types of control actions are considered

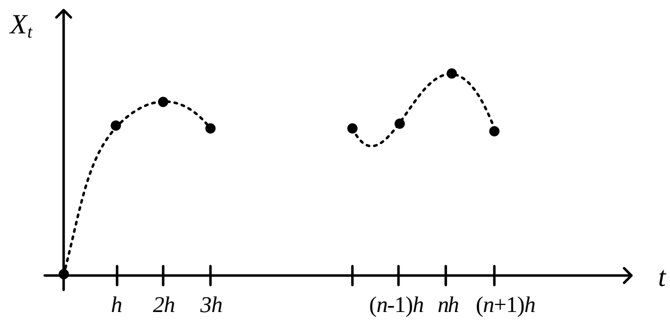

Figure 1: Discrete scheme for 1-D numerical simulation.

- Constant control. For constant control, we consider a fixed control action $u \in \mathcal{U}$. Hence, the controlled diffusion process becomes

$$dX_t^u = b(X_t^u, u, t)dt + \sigma(X_t^u, u, t)dW_t,$$

or in the discrete-time form

$$X_{(k+1)h}^u = X_{kh}^u + b(X_{kh}^u, u, kh)h + \sigma(X_{kh}^u, u, kh)\sqrt{h}\xi_k.$$

- State feedback control. Consider the control policy $g : \mathbb{R}^n \times [0, T] \to u$, deterministic drift term $b^g(x, t) := b(x, g(x, t), t)$ and diffusion matrix $\sigma^g(x, t) := \sigma(x, g(x, t), t)$, and the controlled diffusion process becomes

$$dX_t = b^g(X_t, t)dt + \sigma^g(X_t, t)dW_t \tag{6}$$

with control action $u_t = g(X_t, t)$. A general solution to (6) takes the following form

$$X_t^g = X_0^g + \int_0^t b^g(X_s, s)ds + \int_0^t \sigma^g(X_s, s)dW_s. \tag{7}$$

Fix a control action, for example state feedback policy g. Then the cost-to-go function

$$J^g(x) := \mathbf{E}^g \left[c_T(X_T) + \int_0^T c_t(X_t, u_t)dt \right] \tag{8}$$

with running cost $c_t(X_t, u_t)$ and the expectation $\mathbf{E}^g$ is taken with respect to the probability measure under which X_t satisfies (6). Our goal is to find the optimal control policy $g^* : \mathbb{R}^n \times [0, 1] \mapsto u$ that minimize the cost-to-go function (8) and satisfies the optimality equation

$$V(x, t) = \min_g \mathbf{E}^g \left[c_T(X_T) + \int_0^T c_t(X_t, u_t)dt \right], \tag{9}$$

where X_t satisfies controlled diffusion process (6). Since policy g is of state-feedback form, $\{x_t^g\}$ is a Markov process. Hence, for any $0 \le r < t \le T$,

$$J_r(x; g) = \mathbb{E}^g \left\{ \int_r^t c_s(x_s, u_s) ds + J_t(X_t; g) \big| X_r = x \right\}. \tag{10}$$

Assume the optimal policy g^* exists, then the value function $V_t(x) = J_t(x; g^*)$ also exists. In the following subsection, we first introduce Itô formula (lemma), a cornerstone result in stochastic calculus, and use this lemma to cast (9) into the well-known stochastic Hamilton–Jacobi–Bellman (HJB) equation.

12.4 Stochastic Hamilton–Jacobi–Bellman equation

We first state the Itô's lemma for a function $f(X_t, t)$ that is twice continuously differentiable in X and differentiable in t.

Lemma 12.1 *Assume X_t is an Itô diffusion process that satisfies $dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t$. Let $f(X_t, t)$ be a function that is twice continuously differentiable in X_t and differentiable in t. Then for every $t \ge 0$,*

$$f(X_t, t) = f(X_0, 0) + \int_0^t \frac{\partial f}{\partial s}(X_s, s) ds + \int_0^t \nabla f(X_s, s)^\top dX_s + \frac{1}{2} \int_0^t \nabla^2 f(X_s, s) d\langle X, X \rangle_s,$$

or equivalently in the differential form

$$df(X_t, t) = \frac{\partial f}{\partial t}(X_t, t)dt + \nabla f(X_t, t)^\top dX_t + \frac{1}{2} dX_t^\top \nabla^2 f(X_t, t) dX_t.$$

Lemma 12.1 (Itô's lemma) can also be applied to general semi-martingales, and interested readers may refer to the supplementary discussion in [1]. Next, we show a brief scheme for proving Lemma 12.1.

The proof of Itô's lemma is based on the application of second order Taylor expansion to stochastic function $f(X_t, t)$. Assume X_t is an Itô diffusion process that satisfies $dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t$, and function $f(X_t, t) : \mathbb{R}^n \times [0, \infty) \mapsto \mathbb{R}$ is twice continuously differentiable in X and differentiable in t. Let $y_t := f(X_t, t)$. First, we apply the first order Taylor expansion to y_{t+dt} and show why it does not work. Expanding y_{t+dt} to the first order, we have

$$\begin{aligned} y_{t+dt} &= f(X_{t+dt}, t + dt) = f(X_t + b(X_t, t)dt + \sigma(X_t, t)dW_t, t + dt) \\ &= f(X_t, t) + \frac{\partial f}{\partial t}(X_t, t)dt + \nabla f(X_t, t)^\top dX_t + O(dt^2) \end{aligned}$$

where

$$\nabla f(X_t, t)^\top dX_t = \nabla f(X_t, t)^\top b(X_t, t)dt + \nabla f(X_t, t)^\top \sigma(X_t, t)dW_t.$$

Since $dW_t = W_{t+dt} - W_t \overset{\text{def}}{=} \sqrt{dt} \cdot Z$ where $Z \sim N(0, I_m)$ and $\sqrt{dt} \gg dt$ when $dt \to 0$, we need to preserve the second order term to control the truncation error and avoid the dominance of term consisting $\sqrt{dt}$. Before we give the second order Taylor expansion, by convention, we have $dt^2 = 0$, $dt \cdot dW_t = 0$, and $dW_t \cdot dW_t = 0$, which is summarized in the following table

	dt	dW_t
dt	0	0
dW_t	0	dt

Applying the second order Taylor expansion to y_{t+dt}, we have

$$\begin{aligned} y_{t+dt} &= f(X_t,t) + \frac{\partial f}{\partial t}(X_t,t)dt + \nabla f(X_t,t)^\top dX_t + \frac{1}{2}dX_t^\top \nabla^2 f(X_t,t)dX_t + O(dt) \\ &= f(X_t,t) + \frac{\partial f}{\partial t}(X_t,t)dt + \nabla f(X_t,t)^\top \left[b(X_t,t)dt + \sigma(X_t,t)dW_t\right] \\ &\quad + \frac{1}{2}\left[b(X_t,t)dt + \sigma(X_t,t)dW_t\right]^\top \nabla^2 f(X_t,t)\left[b(X_t,t)dt + \sigma(X_t,t)dW_t\right] \\ &= f(X_t,t) + \left\{\frac{\partial f}{\partial t}(X_t,t) + b(X_t,t)^\top \nabla f(X_t,t) + \frac{1}{2}\mathrm{tr}[\sigma(X_t,t)\sigma(X_t,t)^\top \nabla^2 f(X_t,t)]\right\} dt \\ &\quad + \nabla f(X_t,t)^\top \sigma(X_t,t)dW_t \\ &= y_t + \left[\frac{\partial f}{\partial t}(X_t,t) + \mathcal{A}_t f(X_t,t)\right] dt + \nabla f(X_t,t)^\top \sigma(X_t,t)dW_t. \end{aligned} \tag{11}$$

where generator $\mathcal{A}_t f(X_t,t) = b(X_t,t)^\top \nabla f(X_t,t) + 1/2 \cdot \mathrm{tr}[\sigma(X_t,t)\sigma(X_t,t)^\top \nabla^2 f(X_t,t)]$. Equality (11) can be rewritten as

$$dy_t = \left[\frac{\partial f}{\partial t}(X_t,t) + \mathcal{A}_t f(X_t,t)\right] dt + \nabla f(X_t,t)^\top \sigma(X_t,t)dW_t,$$

or equivalently, in integral form

$$y_t = y_0 + \int_0^t \left[\frac{\partial f}{\partial t}(X_t,t) + \mathcal{A}_t f(X_t,t)\right] dt + \int_0^t \nabla f(X_t,t)^\top \sigma(X_t,t)dW_t. \tag{12}$$

The preceding derivations verify the Itô's lemma in Lemma 12.1. Especially, when $\sigma \equiv 0$, we have $dy_t = [\partial f(X_t,t)/\partial t + b(X_t,t)^\top \nabla f(X_t,t)]dt$. With Itô's lemma, we will next derive the stochastic HJB equation for controlled diffusion process.

The following theorem gives the formulation of stochastic HJB equation

Theorem 12.1 *For a twice continuously differentiable value function $V(x,t)$ that satisfies the optimality equation*

$$V(x,t) = \min_g \mathbf{E}^g \left[c_T(X_T) + \int_t^T c_t(X_\tau, u_\tau)d\tau \;\middle|\; X_t = x \right], \tag{13}$$

we can cast it into the stochastic HJB equation

$$\frac{\partial}{\partial t}V_t(x) = -\min_{u\in\mathcal{U}} \left\{c_t(x,u) + \mathcal{A}_t^u V_t(x)\right\} \tag{14}$$

with generator $\mathcal{A}_t f(X_t,t) = b(X_t,t)^\top \nabla f(X_t,t) + 1/2 \cdot tr[\sigma(X_t,t)\sigma(X_t,t)^\top \nabla^2 f(X_t,t)]$, terminal condition $V_T(x) = c_T(x)$, and optimal control policy

$$g^*(x,t) = \arg\min_{u\in\mathcal{U}} \left[c(x,u) + \mathcal{A}_t^u V_t(x)\right]. \tag{15}$$

Proof: Assume value function $V_t(x)$ is C^2 differentiable in x. Fix $0 \le r < s \le T$. Value function $V_s(X_s^g)$ satisfies the Itô's formula given in (12)

$$V_s(X_s^g) = V_r(X_r^g) + \int_r^s \left[\frac{\partial}{\partial t}V_t(X_t^g) + \mathcal{A}_t^g V_t(X_t^g)\right] dt + \int_r^s \nabla V_t(X_t^g)^\top \sigma(X_t, t) dW_s \tag{16}$$

Taking expectation on both sides of (16) subject initial condition $X_s = x$ gives

$$\mathbf{E}^g \left[V_s(X_s)|X_r = x\right] = V_r(x) + \mathbf{E}^g \left\{ \int_r^s \left[\frac{\partial}{\partial t}V_t(X_t^g) + \mathcal{A}_t^g V_t(X_t^g)\right] dt \;\middle|\; X_r = x \right\} + 0. \tag{17}$$

With some manipulations, (17) can be rewritten as

$$V_r(X) = -\mathbf{E}^g \left\{ \int_r^s \left[\frac{\partial}{\partial t}V_t(X_t^g) + \mathcal{A}_t^g V_t(X_t^g)\right] dt \;\middle|\; X_r = x \right\} + \mathbf{E}^g \left[V_s(X_s)|X_r = x\right] \tag{18}$$

Given a policy g and the optimal policy g^*, we define a policy $\bar{g} : \mathbb{R}^n \times [0, T] \to \mathcal{U}$ such that

$$\bar{g}(t) = \begin{cases} g(t) & t \le s \\ g^*(t) & t > s \end{cases}$$

which implies that policy $g(t)$ is selected on time interval $[0, s]$ and $g^*(t)$ is selected on time interval $(s, T]$. Value function $V_r(X)$ and cost-to-go $J_r(x; \bar{g})$ satisfy

$$\begin{aligned} V_r(X) \le J_r(X; \bar{g}) &= \mathbf{E}^{\bar{g}} \left[\int_r^s c_t(X_t, u_t) dt + J_s(X_s; \bar{g}) \;\middle|\; X_r = x \right] \\ &= \mathbf{E}^{\bar{g}} \left[\int_r^s c_t(X_t, u_t) dt \;\middle|\; X_r = x \right] + \mathbf{E}^{\bar{g}} \left[J_s(X_s; \bar{g}) \;\middle|\; X_r = x \right] \\ &= \mathbf{E}^{\bar{g}} \left[\int_r^s c_t(X_t, u_t) dt \;\middle|\; X_r = x \right] + \mathbf{E}^{\bar{g}} \left[V_s(X_s) \;\middle|\; X_r = x \right] \end{aligned} \tag{19}$$

Subtracting (18) from (19), we have

$$\mathbb{E}^g \left\{ \int_r^s \left[c_t(X_t, u_t) + \frac{\partial}{\partial t}V_t(X_t) + \mathcal{A}_t^g V_t(X_t)\right] dt \;\middle|\; X_r = x \right\} \ge 0, \qquad \forall g. \tag{20}$$

Equality "=" in (20) holds when $g = g^*$. Taking $s \to r$ in (20), we have

$$\frac{\partial}{\partial t}V_t(x) + \min_{u \in \mathcal{U}} \left[c_t(x, u) + \mathcal{A}_t^u V_t(x)\right] = 0 \tag{21}$$

with boundary $V_T(x) = c_T(x)$, which implies the stochastic HJB equation given in (14). This completes the proof. ∎

From the preceding derivations, if $\exists V_t(x)$ that is twice differentiable in x, once differentiable in t and solves stochastic HJB equation (14), then $V_t(x)$ is the value function and optimal control policy $g^*(x,t) = \arg\min_{u\in\mathcal{U}} \left[c(x,u) + \mathcal{A}_t^u V_t(x)\right]$. By the end of this lecture, we apply Theorem 12.1 to solve the LQR problem in continuous-time. Detailed derivations will be given in the lecture of week 13.

Example 12.1 (LQR in continuous-time) *Consider the state space $\mathcal{X} = \mathbb{R}^n$ and control space $\mathcal{U} = \mathbb{R}^m$. The linear drift term*

$$b(X,u) = AX + Bu$$

with $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$. Diffusion coefficient $\sigma(x,u) = \Gamma \in \mathbb{R}^{n\times r}$ and $A(x) = \Gamma\Gamma^\top$. When consider constant control, the controlled diffusion process becomes

$$dX_t = (AX_t + Bu)dt + \Gamma dW_t.$$

Running cost function $c_t(X,u) = X^\top QX + u^\top Ru$ with Q being semi-positive definite and R being positive definite, and the terminal cost $c_T(X) = X^\top Q_T X$.The value function or solution to HJB equation takes the following form

$$V_t(X) = X^\top K_t X + f(t)$$

where $K_t = K_t^\top \geq 0$, $K_T = Q_T$, $f(T) = 0$

$$\begin{aligned}\frac{dK_t}{dt} &= -Q - K_tA - A^\top K_t + K_tBR^{-1}B^\top K_t,\\ \frac{df(t)}{dt} &= -tr\left(\Gamma\Gamma^\top K_t\right).\end{aligned}$$

The optimal policy takes the form

$$g^*(X,t) = -R^{-1}B^\top K_t X,$$

and the optimal controlled diffusion process becomes

$$dX_t = (A - BR^{-1}B^\top K_t)X_t dt + \Gamma dW_t.$$

References

[1] J.-F. Le Gall *Brownian Motion, Martingales, and Stochastic Calculus.* Springer, 2016.