5. Stochastic Control

5.1. **Some Preliminaries on Markov Processes.** We recall some basic notions related to Markov processes.

Definition: Given a homogeneous Markov process $\{X_t\}$, the **(infinitesimal) generator** $\mathcal{A}$ of $\{X_t\}$ is defined to be

$$\mathcal{A}f(x) = \lim_{t\downarrow 0} \frac{\mathbb{E}_x[f(X_t)] - f(x)}{t},$$

defined for f in the domain $D_{\mathcal{A}}$ of the generator, where

$$D_{\mathcal{A}}(x) \doteq \{f : \mathbb{R}^n \to \mathbb{R} \text{ such that the limit exists at } x\},$$

and

$$D_{\mathcal{A}} = \underset{x\in\mathbb{R}^n}{\cap} D_{\mathcal{A}}(x).$$

Fact: If X is an Itô diffusion, satisfying

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t$$

then, as can be verified by Itô's formula, $C_0^2(\mathbb{R}^n) \subset D_{\mathcal{A}}$ and, for $f \in C_0^2(\mathbb{R}^n)$,

$$\mathcal{A}f(x) = \sum_i b_i(x)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j}(\sigma\sigma^T)_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Dynkin's Formula

Let $f \in \mathcal{C}_0^2(\mathbb{R}^n)$, and τ be a stopping time with $\mathbb{E}[X(\tau)] < \infty$. Then

$$\mathbb{E}_x[f(X_\tau)] = f(x) + \mathbb{E}_x[\int_0^\tau \mathcal{A}f(X_s)ds]. \tag{5.53}$$

Remark: If τ is the first exit time of a bounded set D, then $\mathbb{E}x[\tau] < \infty$ and (5.53) holds for any $f \in C^2$.

Suppose you are given a strong Markov process $\{X_t\}$ that admits a canonical representation, i.e., it is is defined on $(\Omega, \mathcal{B}, \{\mathcal{B}_t\}, \mathbb{P})$, where Ω is the (appropriate) path space, $\{\mathcal{B}_t\}\}$ is the σ-algebra generated by the projections and

$$X_t(\omega) = \omega_t, \qquad t \in [0, \infty).$$

Then, for every $t \in [0, \infty)$, we can define the shift operators $\theta_t : \Omega \mapsto \Omega$ given by

$$(\theta_t X)_s(\omega) \doteq X(\theta_t\omega) = \omega(t + s).$$

Likewise for any stopping time τ,

$$(\theta_\tau X)_s(\omega) = (\theta_t X)_s(\omega), \qquad \text{on the set } \tau(\omega) = t.$$

Recall that if X is a strong Markov process, then given any measurable function f, τ_D an exit time from an open set D and stopping time $\tau \leq \tau_D$, we have

$$(5.54) \qquad \eta = \int_0^{\tau_D} g(Y_s)\,ds \quad \Rightarrow \quad \theta_\tau \eta = \int_\tau^{\tau_D} g(Y_s)\,ds.$$

5.2. **The homogeneous Markov case.** Suppose $\mathcal{U}$ is a control parameter space and $M(\mathbb{R}^n : \mathcal{U})$ is the space of measurable functions from $\mathbb{R}^n$ to $\mathcal{U}$. Let b and σ be drift and dispersion coefficients; $b : \mathbb{R}^n \times \mathcal{U} \mapsto \mathbb{R}^n$; $\sigma : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}^{n\times m}$, and consider the associated family of controlled processes: given $u \in \mathcal{M}(\mathbb{R}^n : \mathcal{U})$.

$$(5.55) \qquad dY_t^u = b\left(Y_t, u(Y_t)\right) dt + \sigma(Y_t, u(Y_t))dW_t$$

where W is an m-dimensional Brownian motion. For the moment, we will not specify any conditions on b, σ, but simply assume that they are regular enough that there exists a weak solution to the SDE, (5.55), which is unique in distribution.

Suppose you are given a running cost $F : \mathbb{R}^n \times \mathcal{U} \mapsto \mathbb{R}$ and a terminal cost $K : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}$, an open set $G \subset \mathbb{R}^n$ and let τ be the exit time from the domain G. Define, for a fixed $u \in \mathcal{M}(\mathbb{R}^n : \mathcal{U})$, the cost functional

$$V^u(y) = \mathbb{E}_y \left[\int_0^\tau F^u(Y_s^u)ds + K^u\left(Y_\tau^u\right) \right],$$

where we use the abbreviation $F^u(y) = F(y, u(y))$ and, likewise, $K^u(y) = K(y, u(y))$.

The *stochastic control problem* is to infimize the cost over all possible controls, and to find the corresponding "optimal control" u^* that achieves the infimum if it exists. In other words, we would like to find V, when

$$V(y) \doteq \inf_u J^u(y)$$

and identify u^* such that $V(y) = J^{u^*}(y)$, if it exists. V is called the *value function* of the stochastic control problem.

When V is sufficiently smooth, it is often possible to characterize V as a solution to a second-order PDE (as opposed to the deterministic case, where it is a solution to a first-order PDE), which is referred to as the Hamilton-Jacobi-Bellman (HJB) equation.

Theorem 5.1. *If $V \in \mathcal{C}^2(G)$ and an optimal Markov control u^* exists, then*

$$(5.56) \qquad \inf_{u\in\mathcal{U}} \{F^u(y) + (\mathcal{A}^u V)(y)\} = 0 \qquad \text{for all } y \in G,$$

and

$$(5.57) \qquad V(y) = K(y) \qquad \text{for all } y \in \partial_R G,$$

where $\partial_R G$ are the regular points of ∂G for $Y_t^{u^*}$. Moreover, the infimum in (5.56) is attained if $u = u^*(y)$.

Proof. Let u be any Markov control. Since

$$J^u(y) = \mathbb{E}_y\left[\int_0^\tau F^u(Y_s^u)ds + K^u(Y_\tau^u)\right],$$

by the strong Markov property (see (5.54)) we know that for any stopping time $\alpha \le \tau$,

$$\begin{aligned}
\mathbb{E}_y\left[J^u(Y_\alpha)\right] &= \mathbb{E}_y\left[\mathbb{E}_{Y_\alpha}\left[\int_0^\tau F^u(Y_r)\,dr + K^u(Y_\tau)\right]\right] \\
&= \mathbb{E}_y\left[\mathbb{E}_y\left[\theta_\alpha\left(\int_0^\tau F^u(Y_r)\,dr + K^u(Y_\tau)\right)|\mathcal{F}_\alpha\right]\right] \\
&= \mathbb{E}_y\left[\mathbb{E}_y\left[\int_\alpha^\tau \left(F^u(Y_r)\,dr + K^u(Y_\tau)\right)|\mathcal{F}_\alpha\right]\right] \\
&= \mathbb{E}_y\left[\mathbb{E}_y\left[\int_0^\tau F^u(Y_r)\,dr + K^u(Y_\tau) - \int_0^\alpha F^u(Y_r)\,dr\right]\right] \\
&= J^u(y) - \mathbb{E}_y\left[\int_0^\alpha F^u(Y_r)\,dr\right].
\end{aligned}$$

Therefore,

$$J^u(y) = \mathbb{E}_y\left[\int_0^\alpha F^u(Y_r)\,dr + \mathbb{E}_y\left[J^u(Y_\alpha)\right]\right].$$

Now, let $W \subset G$ be of the form

$$\mathcal{O} = \{z \in G : |z - x| < \varepsilon\}$$

where $\varepsilon > 0$ is small enough that $\overline{\mathcal{O}} \subset G$. Put $\alpha = \tau_{\mathcal{O}}$, to be the first exit time of Y from $\mathcal{O}$. Given an arbitrary $v \in U$, if an optimal Markov control $u^*(y)$ exists, choose

$$u(z) = \begin{cases} v & \text{if } z \in \mathcal{O} \\ u^*(z) & \text{otherwise} \end{cases}$$

Then

$$V(Y_\alpha) = J^{u_*}(Y_\alpha) = J^u(Y_\alpha),$$

and so

$$V(y) \le J^u(y) = \mathbb{E}_y\left[\int_0^\alpha F^v(Y_r)\,dr\right] + \mathbb{E}_y\left[V(Y_\alpha)\right].$$

Since $V \in \mathcal{C}^2(G)$ by assumption, Dynkin's formula shows that

$$\mathbb{E}_y\left[V(Y_\alpha)\right] = V(y) + \mathbb{E}_y\left[\int_0^\alpha \mathcal{A}^u V(Y_r)\,dr\right].$$

Therefore,

$$V(y) \le \mathbb{E}_y\left[\int_0^\alpha F^v(Y_r)\,dr\right] + V(y) + \mathbb{E}_y\left[\int_0^\alpha \mathcal{A}^v V(Y_r)\,dr\right].$$

In other words,

$$\mathbb{E}_y\left[\int_0^\alpha \left(F^v(Y_r)+\mathcal{A}^v V(Y_r)\right)\, dr\right]\geq 0.$$

Therefore,

$$\frac{\mathbb{E}_y\left[\int_0^\alpha \left(F^v(Y_r)+\mathcal{A}^v V(Y_r)\right)\, dr\right]}{\mathbb{E}^y[\alpha]}\geq 0 \qquad \text{for every } \mathcal{O}.$$

Since F^v and $\mathcal{A}^v V$ are continuous at y, seding $\varepsilon\downarrow 0$, we obtain (5.56). The boundary conditions follow immediately from the definition of a regular point and the terminal condition. □

We also state the converse, which can be proved using Dynkin's formula (the proof is left to the reader).

Theorem 5.2. *Let h be a bounded function in $\mathcal{C}^2(G)\cap\mathcal{C}\left(\overline{G}\right)$ such that for every $v\in U$,*

$$F^v(y)+\mathcal{A}^v h(y)\geq 0 \qquad \forall y\in G;$$

with boundary values

$$h(y)=K(y) \qquad y\in\partial G.$$

Then

$$h(y)\leq J^u(y)$$

for all Markov controls u. Moreover, if for each $y\in G$, there exists $u_0(y)$ such that

$$F^{u_0(y)}(y)+(\mathcal{A}^{u_0(y)}h)(y)=0$$

then $u=u_0(y)$ is a Markov control such that

$$h(y)=J^{u_0(y)}(y),$$

and so u_0 is an optimal control.

5.3. **The Inhomogeneous Markov Case.** Just as in the optimal stopping problem, we can reduce the inhomogeneous Markov case to the homogeneous case. Suppose you consider Markov controls of the form

$$u=u(t,X_t(\omega)),$$

and, for any given control u, let the associated controlled process satisfy

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

Then, by introducing $Y_s=(t+s,X_{t+s})$, the system equation becomes

$$dY_t=dY^w=b(Y_t^u,u(Y_t))dt+\sigma(Y_t,u(Y_t))dW_t,$$

where $u:\mathbb{R}^{n+1}\mapsto U$, with $u(Y_t):u(t,X_t)$.

Note that in this case, for $v\in U$,

$$A^v f(y)=\frac{\partial f}{\partial t}(y)+\sum b_i(y,v)\frac{\partial f}{\partial x_i}+\sum a_{ij}(y,v)\frac{\partial^2 f}{\partial_{x_i},\partial_{x_j}},$$

where now $y \in \mathbb{R}_+ \times \mathbb{R}^d$ and, as usual, $a_{ij}(y) = (\sigma^T \sigma)_{ij}$.

Note that in this case we can consider the stochastic control problem on a fixed horizon $[0, t_i]$ by choosing $G \subseteq \mathbb{R}^{n+1}$ appropriately so that $\tau = t_1$.

Example 1: (The linear regulation problem)

Suppose

$$dY_t = (L_t Y_t + M_t u_t)\, dt + \sigma_t dW_t$$

and the cost is of the form

$$J^u(y) = \mathbb{E}_y \left[\int_0^{t_1} \left(Y_s^T C_s Y_G + u_s^T D_s u_s \right) ds + X_{t_1}^T R X_{t_1} \right]$$

where $L_t, M_t, \sigma_t, C_t, D_t, R$ are continuous and non-random. Also, assume C_t, R are symmetric, non-negative definite and D_t is symmetric, positive definite $\forall t$.

The associated stochastic control problem admits the following interpretation: the aim is to find a control w that makes $|Y_t|$ small first and such that the "energy" used ($\sim u^T D u$) is small.

The HJB-equation in this case is easily seen to be

$$\begin{aligned} 0 &= \inf_v \{ F^u(t,x) + \mathcal{A}^v V(t,x) \} \\ &= \frac{\partial V}{\partial t} + \inf_v \left\{ x^T C_t x + v^T D_t v + \sum_i (L_t x + M_t v) \frac{\partial V}{\partial x_i} + \frac{1}{2} \sum a_{i_j} \frac{\partial^2 V}{\partial_{x_i} \partial_{x_j}} \right\}, \end{aligned}$$

with the boundary condition

$$V(t_1, x) = x^T R_x.$$

Let us try to find a solution V of the form

$$V(t,x) = x^T S_t x + a_t \tag{5.58}$$

where each X_t is a symmetric, non-negative definite matrix, a_t, s_t are continuously differentiable and non-random.

To have the boundary conditions coincide, let us set

$$S_{t_1} = R \quad a_{t_1} = 0.$$

Also, substituting (5.58) into the HJB equation, we have

$$F^v(t,x) + \mathcal{A}^u V(t,x) = x^T S_t' x + a_t' + x^T C_t x + u^T D_t v \tag{5.59}$$

$$+ (L_t x + M_t v)^T (S_t x + S_t^T x) + \sum_{i,j} a_{ij} (S_t)_{i,j}. \tag{5.60}$$

The minimum of this expression is obtained when

$$\frac{\partial}{\partial v_i} \left(F^v(t,x) + \mathcal{A}^v(t,x) \right) = 0,$$

i.e., when

$$2D_t v + 2M_t^T S_t x = 0$$

or when

$$v = -D_t^{-1} M_t^T S_t x. \tag{5.61}$$

Substituting this back into (5.60), algebraic manipulations show that the right hand side of (5.60) is zero if

$$\begin{aligned} S_t' &= -2L_t^T S_t + S_t M_t D_t^{-1} S_t - C_t; \quad t < t_1, \\ a_1' &= -tr(\sigma\sigma^T S)_t \qquad\qquad\qquad\qquad t < t_j. \end{aligned}$$

This is a Riccati type equation and determines S_t uniquely. Combining this with the boundary condition, we obtain

$$a_t = \int_t^{t_1} tr(\sigma\sigma^T S)_s \, ds.$$

Moreover, from (5.61) we conclude that the optimal control satisfies

$$u^*(b, x) = -D_t^{-1} M_t^T S_x x, \quad t < t_1$$

and the minimum cost is

$$V(t, x) = x^T S_t x + \int_t^{t_1} (tr(\sigma\sigma^T)S)_s \, ds, \quad t < t_1.$$

When compared with the solution to the corresponding deterministic problem, the formula shows that the extra cost due to the noise in the system is given by

$$a_t = \int_t^{t_1} tr(\sigma\sigma^T S)_s \, ds.$$