17 Continuous-time stochastic systems

The discussion in this section will not be rigorous. A stochastic controllable dynamical system of jump type is given by a function

$$q : \mathbb{R}^+ \times \{(x, y) \in S \times S : x \neq y\} \times A \to \mathbb{R}^+.$$

We assume that the state-space S is countable. We write $q_{xy}(t, a) = q(t, x, y, a)$. For x, y distinct, $q_{xy}(t, a)$ gives the rate of jumping from x to y when at time t we choose action a. It is convenient to write

$$q_{xx}(t, a) = -\sum_{y \neq x} q_{xy}(t, a).$$

We consider Markov controls $u : \mathbb{R}^+ \times S \to A$ and set

$$q^u_{xy}(t) = q_{xy}(t, u(t, x)).$$

Then the controlled process $(X_t)_{t \geqslant s}$ for control u, starting from (s, x_s) satisfies $X_s = x_s$ and, for all $t \geqslant s$ and $x \in S$, conditional on $X_t = x$, as $\delta \downarrow 0$,

$$X_{t+\delta} = \begin{cases} x, & \text{with probability } 1 + q^u_{xx}(t)\delta + o(\delta), \\ y, & \text{with probability } q^u_{xy}(t)\delta + o(\delta), \text{ for all } y \neq x. \end{cases}$$

We consider the same sorts of control problem as in Section 15, where now we take an expectation in defining the cost functions

$$V^u(s, x) = \mathbb{E}^u_{(s,x)} \left(\int_s^\tau c(X_t, U_t) dt + C(X_\tau) \right), \quad V(s, x) = \inf_u V^u(s, x).$$

We now give a derivation of the optimality equation for V. Suppose we start at (t, x) and choose action a until time $t + \delta$, then switch to an optimal control. On comparing the resulting expected total cost with that of an optimal control from the outset, we obtain

$$V(t, x) \leqslant c(x, a)\delta + \mathbb{E}(V(t + \delta, X_{t+\delta}) | X_t = x).$$

Now expand to first order in δ

$$\begin{aligned} \mathbb{E}(V(t + \delta, X_{t+\delta}) | X_t = x) &= V(t + \delta, x)(1 + q_{xx}(t, a)\delta) + \sum_{y \neq x} V(t + \delta, y) q_{xy}(t, a)\delta + o(\delta) \\ &= V(t, x) + \dot{V}(t, x)\delta + \sum_{y \in S} q_{xy}(t, a) V(t, y)\delta + o(\delta). \end{aligned}$$

So

$$0 \leqslant \{c(x, a) + \dot{V}(t, x) + QV(t, x, a)\}\delta + o(\delta),$$

with equality if a is chosen optimally, where

$$QV(t, x, a) = \sum_{y \in S} q_{xy}(t, a) V(t, y).$$

Thus we obtain the optimality equation

$$\inf_a \{c(x,a) + \dot{V}(t,x) + QV(t,x,a)\} = 0$$

and we expect to find the optimal control as the minimizing action a.

Now we shall give an analogous discussion in the case of a diffusive stochastic controllable dynamical system. We specify two functions

$$\sigma, b : \mathbb{R}^+ \times \mathbb{R} \times A \to \mathbb{R}.$$

The function σ^2 is the *diffusivity* and determines the size of the stochastic fluctuations or *noise* in the dynamics. The function b is the *drift* and determines the average velocity. Given a choice of Markov control $u : \mathbb{R}^+ \times \mathbb{R} \to A$, set $\sigma^u(t,x) = \sigma(t,x,u(t,x))$ and $b^u(t,x) = b(t,x,u(t,x))$. The corresponding dynamics can be described infinitesimally[29], conditional on $X_t = x$, by

$$X_{t+\delta} = x + \sigma^u(t,x)\Delta + b^u(t,x)\delta + o(\delta),$$

as $\delta \to 0$, where $\mathbb{E}(\Delta) = 0$ and $\mathbb{E}(\Delta^2) = \delta$. We define cost functions V^u and V exactly as in the jump case.

Let us now derive the optimality equation for V. Suppose we start at (t,x) and choose action a until time $t+\delta$, then switch to an optimal control. On comparing the resulting expected total cost with that of an optimal control from the outset, we obtain

$$V(t,x) \leqslant c(x,a)\delta + \mathbb{E}\left(V(t+\delta, x + \sigma^u(t,x)\Delta + b^u(t,x)\delta + o(\delta))\right).$$

We expand to fisrt order in δ

$$\begin{aligned}&\mathbb{E}\left(V(t+\delta, x + \sigma^u(t,x)\Delta + b^u(t,x)\delta + o(\delta))\right)\\ &\quad = V(t,x) + \dot{V}(t,x)\delta + V'(t,x)(b(t,x,a)\delta + \sigma(t,x,a)\Delta) + \frac{1}{2}V''(t,x)\sigma(t,x,a)^2\Delta^2 + o(\delta).\end{aligned}$$

So

$$0 \leqslant \{c(x,a) + \dot{V}(t,x) + LV(t,x,a)\}\delta + o(\delta)$$

with equality if a is chosen optimally, where

$$LV(t,x,a) = \frac{1}{2}\sigma(t,x,a)^2 V''(t,x) + b(t,x,a)V'(t,x).$$

Thus the optimality equation is

$$\inf_a \{c(x,a) + \dot{V}(t,x) + LV(t,x,a)\} = 0$$

and we expect to find the optimal control as the minimizing action a

[29]A rigorous formulation rests on the theory of stochastic integration. The infinitesimal formula given is replaced by the stochastic integral equation

$$X_t = x + \int_0^t \sigma^u(s,X_s)dB_s + \int_0^t b^u(s,X_s)ds,$$

where $(B_t)_{t\geqslant 0}$ is a Brownian motion.

Example (Escape to the boundary). Consider the diffusive controllable dynamical system in $[-1,1]$, with constant diffusivity $\sigma^2 = 1$ and with drift $b(t,x,u) = u$. Suppose we wish to minimize

$$V^u(x) = \frac{1}{2}\mathbb{E}_x^u\left(\tau + \int_0^\tau U_s^2 ds\right),$$

where $\tau = \inf\{t \geqslant 0 : |X_t| = 1\}$, $x \in [-1,1]$, and $U_s = u(X_s)$. The optimality equation is

$$\inf_u \left\{\frac{1+u^2}{2} + uV'(x) + \frac{1}{2}\sigma^2 V''(x)\right\} = 0.$$

The left hand side is minimized to

$$\frac{1}{2}(\sigma^2 V''(x) - V'(x)^2 + 1)$$

by taking $u = -V'(x)$. We can solve the differential equation with boundary conditions $V(-1) = V(1) = 0$ to obtain

$$V'(x) = -\tanh \lambda x,$$

where $\lambda = 1/\sigma^2$. It follows easily now that $V(x)$ is increasing in λ, and takes the limiting values $V(x) = 0$ as $\lambda \to 0$ and $V(x) = 1 - |x|$ as $\lambda \to \infty$. This fits with the intuitively reasonable idea that noise makes it easier to escape to the boundary.