17. Stochastic control in continuous time

Two types

- Jump type : a continuous-time Markov chain whose transition rates we control.
- Diffusive type : a diffusion process whose diffusivity and drift we control

A continuous-time stochastic controllable dynamical system of "jump type" is given by a map

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

The state-space S is assumed here to be countable.

- $q(t,x,y,a)$ is the "rate of 'jumping'" from x to y at time t for action a
- We assume that the total 'jump rate'

$$\sum_{y \neq x} q(t,x,y,a) \quad \text{is finite for all } t, x, a.$$

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

$$q^u_{xy}(t) = q(t, x, y, u(t,x)), \quad x \neq y, \quad q^u_{xx}(t) = -\sum_{y \neq x} q^u_{xy}(t).$$

Then

$$Q^u(t) = (q^u_{xy}(t) : x, y \in S)$$

is a Q-matrix.

The controlled process will be the associated time-inhomogeneous continuous-time Markov chain.

The controlled process $(X_t)_{t \geq s}$, starting from (s, x) with control u, satisfies:

- $X_s = x$,
- for all $t \geq s$, all $y, z \in S$ with $y \neq z$, all events $B \subseteq \{X_t = y\}$ depending only on $(X_r : s \leq r \leq t)$, all $\delta \downarrow 0$

$$X_{t+\delta} = \begin{cases} z & \text{with probability } q^u_{yz}(t)\,\delta + O(\delta^2), \\ y & \text{with probability } 1 + q^u_{yy}(t)\,\delta + O(\delta^2) \end{cases}$$

As in §15

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

Often V can be found by solving the <u>optimality equation</u>

$$\inf_a \left\{ c(x,a) + \dot{V}(t,x) + QV(t,x,a) \right\} = 0 \qquad (t,x) \in \tilde{S}$$

and we determine $u(t,x)$ as the minimizing action.

Here

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

Non-rigorous derivation of optimality equation

Start at (s, x) and choose action a for time δ, switching then to an optimal control. Then

$$V(s,x) \leq c(x,a)\delta + \mathbb{E}^{a}_{(s,x)} V(s+\delta, X_{s+\delta}) + o(\delta),$$

with equality if a is optimal. Now

$$\mathbb{E}^{a}_{(s,x)} V(s+\delta, X_{s+\delta}) = V(s,x) + \dot{V}(s,x)\delta + \sum_{y \in S} q(s,x,y,a) V(s,y)\delta + o(\delta)$$

so we have

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

with equality if a is optimal.

A continuous-time stochastic controllable dynamical system of <u>diffusive type</u> is given by two functions

$$b : \mathbb{R}^+ \times \mathbb{R} \times A \longrightarrow \mathbb{R}, \qquad \sigma : \mathbb{R}^+ \times \mathbb{R} \times A \longrightarrow \mathbb{R}^+$$

Here the state-space is (for simplicity) $\mathbb{R}$.

- $b(t, x, a)$ is the <u>drift</u>, the mean velocity at time t and position x under action a,
- $\sigma^2(t, x, a)$ is the <u>diffusivity</u> and measures the infinitesimal variance of the state's motion.

Given a choice of Markov control $u : \mathbb{R}^+ \times \mathbb{R} \to A$, define

$$b^u(t,x) = b(t,x,u(t,x)), \qquad \sigma^u(t,x) = \sigma(t,x,u(t,x)).$$

The behaviour of the controlled process $(X_t)_{t \geq s}$ over a small time interval is given by

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

where Δ is independent of $(X_r : s \leq r \leq t)$ and $\mathbb{E}(\Delta) = 0$, $\mathbb{E}(\Delta^2) = \delta$.

To make a rigorous theory we would use stochastic differential equations

$$X_t = x + \int_s^t \sigma^u(r,X_r)\,dB_r + \int_s^t b^u(r,X_r)\,dr.$$

Non-rigorous derivation of the optimality equation

Start at (s,x) and choose action a for time δ, switching then to an optimal control. Then

$$V(s,x) \leqslant \left\{ c(x,a)\delta + \mathbb{E}^{a}_{(s,x)} V(s+\delta, X_{s+\delta}) + o(\delta) \right\}$$

with equality if a is optimal. Write $\sigma = \sigma(s,x,a)$, $b = b(s,x,a)$. Under $\mathbb{P}^{a}_{(s,x)}$ we have

$$V(s+\delta, X_{s+\delta}) = V(s,x) + \dot{V}(s,x)\delta + V'(s,x)(\sigma\Delta + b\delta) + \tfrac{1}{2}V''(s,x)\sigma^2\Delta^2 + o(\delta)$$

so

$$\mathbb{E}^{a}_{(s,x)} V(s+\delta, X_{s+\delta}) = V(s,x) + \{\dot{V}(s,x) + \underbrace{V'(s,x)b + \tfrac{1}{2}V''(s,x)\sigma^2}_{LV(s,x,a)}\}\delta + o(\delta)$$

and so

$$c(x,a) + \dot{V}(s,x) + LV(s,x,a) \geqslant 0 \quad \text{with equality if } a \text{ is optimal.}$$

Example - escape to the boundary

Diffusive controllable dynamical system in $[-1, 1]$ with $\sigma^2 = \text{const}$, $\quad b(t,x,a) = a$.

We wish to minimize, over stationary Markov controls u,

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

where $\tau = \inf\{t \geq 0 : |X_t| = 1\}$ and $U_s = u(X_s)$.