Lecture 9. Stochastic optimal control.

Optimal control: • We have a system $y = y(t)$ and we can influence the system via control $\alpha(t)$. y solves

$$\dot{y}(t) = f(y(t), \alpha(t)).$$

• **Goal**: Choose α to maximise/minimise some function of y, α.

Our goal. Study the stochastic version, avoid technicalities by focusing on the "verification thm".

Def 9.1. (Controlled SDE).

Consider $(\Omega, \mathcal{F}, \mathbb{P})$ with a d-dim BM B_t. The basic object of stochastic control theory is an SDE with a control input:

$$(*)\begin{cases} dX_t^\alpha = \mu(t, X_t^\alpha, \alpha_t)\, dt + \sigma(t, X_t^\alpha, \alpha_t)\, dB_t \\ X_0^\alpha = x. \end{cases}$$

where $\mu : \mathbb{R}_+ \times \mathbb{R}^n \times A \to \mathbb{R}^n$, $\sigma : \mathbb{R}_+ \times \mathbb{R}^n \times A \to \mathbb{R}^n$, where A is the control set, and α_t is the control process / control law

Def 9.2. The control law α is called an admissible strategy if

i) α_t is $\mathcal{F}_t$-adapted

ii) $\alpha_t(\omega) \in A$ for every t, ω.

iii) $(*)$ has a unique strong solution.

Def 9.3. An admissible strategy α is called a Markov strategy if

$\alpha_t = \alpha(t, X_t^\alpha)$ for some fcn $\alpha : \mathbb{R}_+ \times \mathbb{R}^n \longrightarrow A$.

Remark. With α Markovian, X_t^α is Markovian.

- easy to implement,
- method in our course automatically gives rise to Markov strategies

Problem Formulation

- What to optimise? $\rightarrow$ A payoff functional.
  - Associated to α.
  - In the form of expectation.

Def 9.4 (Expected payoff function)

Recall the cylinder set $D_T = [0,T) \times D$ with boundary Γ_T, let

$$\Psi^\alpha : D_T \times A \to \mathbb{R} \quad , \quad \text{(running payoff)}$$

$$\Phi : \Gamma_T \longrightarrow \mathbb{R}, \quad \text{(final payoff).}$$

We define the (finite time horizon) expected payoff function J^α:

$$J^\alpha(t,x) = \mathbb{E}_{t,x}\left[\Phi(X_T^\alpha) + \int_t^T \Psi^\alpha(s, X_s^\alpha)\, ds \right].$$

and the value function V:

$$V(t,x) = \sup_{\alpha \in A} J^{\alpha}(t,x) = J^{\alpha^*}(t,x).$$

such a α^*, if exists, is called an optimal control

Our goal: Find α^* and V.

Remark. Two other common formulations are

i) Indefinite time horizon:

Consider $\tau := \min\left\{ \inf\{s \geq t : X_s^{\alpha} \notin D\}, T \right\}$,

$$J^{\alpha}(t,x) = \mathbb{E}_{t,x}\left[\Phi(\tau, X_{\tau}^{\alpha}) + \int_t^{\tau} \Psi^{\alpha}(s, X_s^{\alpha})\,ds \right]$$

ii) Infinite time horizon $t \in [0,\infty)$.

$$J_{\lambda}^{\alpha}(t,x) = \mathbb{E}_{t,x}\left[\int_t^{\infty} e^{-\lambda s} \Psi^{\alpha}(s, X_s^{\alpha})\,ds \right]$$

where $\lambda > 0$ (discounting).

Two questions to ask:

1. Does α^* exist? $J^{\alpha^*}(t,x) \leq J^{\alpha}(t,x)$ for all $\alpha \in A$.

2. If so, how to find it?

We assume α^* exists now, and answer 2.

Idea: Dynamic programming

Lemma 9.5 (Bellman's principle of Optimality).

If α^* is optimal on $[t, T]$, then it is also optimal on every subinterval $[s, T]$, where $s \in [t, T]$.

Prf. Iterated expectation.

Now we shall derive a PDE, and show that solving the control problem is equivalent to solving this PDE.

$\boxed{\text{Strategy.}}$

- Fix $(t, x) \in D_T$.
- Choose $h > 0$ small, $t+h < T$.
- choose an arbitrary $\alpha \in A$.

Comparing 2 scenarios. i) Let $\alpha_1 = \alpha^*$ for all $s \in [t, T]$.

ii). Let $\alpha_2 = \begin{cases} \alpha, & s \le t+h \\ \alpha^*, & s > t+h. \end{cases}$

- Expected payoff for i).

$$J^{\alpha_1}(t, x) = J^{\alpha^*}(t, x) = V(t, x).$$

- Expected payoff for ii) :

$$J^{\alpha_2}(t,x) = \mathbb{E}_{t,x}\left[\Phi(X_T^{\alpha_2}) + \int_t^T \Psi(s, X_s^{\alpha_2})\,ds\right].$$

$$= \mathbb{E}_{t,x}\left[\Phi(X_T^{\alpha^*}) + \int_{t+h}^T \Psi(s, X_s^{\alpha^*})\,ds + \int_t^{t+h} \Psi(s, X_s^{\alpha})\,ds\right]$$

$$= \mathbb{E}_{t,x}\left[\int_t^{t+h} \Psi(s, X_s^{\alpha})\,ds + \mathbb{E}_{t+h, X_{t+h}^{\alpha}}\left[\Phi(X_T^{\alpha^*}) + \int_{t+h}^T \Psi(s, X_s^{\alpha^*})\,ds\right]\right]$$

Markov

$$= \mathbb{E}_{t,x}\left[V(t+h, X_{t+h}^{\alpha}) + \int_t^{t+h} \Psi(s, X_s^{\alpha})\,ds\right].$$

- Trivially, $V(t,x) \overset{①}{\geq} J^{\alpha_2}(t,x)$.

- Assume V smooth, Apply Itô on V :

$$V(t+h, X_{t+h}^{\alpha}) = V(t,x) + \int_t^{t+h} \left(\frac{\partial V}{\partial t} + \mathcal{L}^{\alpha} V\right) ds + \int_t^{t+h} \cdots\, dB_s.$$

$$\mathbb{E}_{t,x}\left[V(t+h, X_{t+h}^{\alpha})\right] \overset{②}{=} V(t,x) + \mathbb{E}_{t,x}\left[\int_t^{t+h} \left(\frac{\partial V}{\partial t} + \mathcal{L}^{\alpha} V\right) ds\right].$$

Combining ① and ② :

$$\mathbb{E}_{t,x}\left[\int_t^{t+h} \Psi(s, X_s^{\alpha})\,ds\right] \leq -\mathbb{E}_{t,x}\left[\int_t^{t+h} \left(\frac{\partial V}{\partial t} + \mathcal{L}^{\alpha} V\right) ds\right]$$

Take it to the limit now :

Divide by h, move it in the expectation, let $h \to 0$:

$$\boxed{\frac{\partial V}{\partial t}(t,x) + \mathcal{L}^{\alpha} V(t,x) + \psi^{\alpha}(t,x) \le 0}$$

Obs. i) The equality holds iff $\alpha = \alpha^*$

ii) (t,x) is arbitrary $\Rightarrow$ holds for all $(t,x) \in D_T$.

iii) $V(T,x) = \phi(x)$.

Taking sup on both sides, we arrive at.

Thm 9.6 (Hamilton-Jacobi-Bellman eqn).

If the value fcn $V \in C^{1,2}$, and α^* exists, then V solves the HJB-eqn:

$$\boxed{\begin{aligned} &\frac{\partial V}{\partial t}(t,x) + \sup_{\alpha \in A} \left\{ \psi^{\alpha}(t,x) + \mathcal{L}^{\alpha} V(t,x) \right\} = 0, \quad (t,x) \in (0,T) \times \mathbb{R}^n \\ &V(T,x) = \phi(x), \quad x \in \mathbb{R}^n \end{aligned}}$$

and for each (t,x) the supremum is attained by $\alpha = \alpha^*$.

Prf. (we sketched).

Remark. HJB says, assume V regular, then.

"V optimal & α^* exists $\Longrightarrow$ V solves the HJB."

This is the necessary condition!

Question. Suppose we solved HJB, have we found V and α^*?

Yes! " $\Leftarrow$ " also holds, HJB is also sufficient.

Thm 9.7 (The verification thm).

If $H \in C^{1,2}$ solves the HJB-eqn, g is admissible and for each (t,x), $\sup_\alpha \{ \Psi(t,x,\alpha) + \mathcal{L}^\alpha H(t,x) \} = \Psi(t,x,g) + \mathcal{L}^g H(t,x)$, then $H = V$ is the value function, and the optimal control α^* exists, $\alpha^* = g$.

Prf. "$H \geq V$":

Choose arbitrary $\alpha \in A$, fix (t,x). Apply Itô on H.

$$H(T, X_T^\alpha) \overset{(*)}{=} H(t,x) + \int_t^T \left(\frac{\partial H}{\partial t} + \mathcal{L}^\alpha H \right)(s, X_s^\alpha)\, ds + \int_t^T -\, dB_s.$$

- H solves HJB $\Rightarrow$ $\frac{\partial H}{\partial t} + \mathcal{L}^\alpha H + \Psi^\alpha \leq 0$. $\forall \alpha$.

  Therefore, $\left(\frac{\partial H}{\partial t} + \mathcal{L}^\alpha H \right)(s, X_s^\alpha) \leq -\Psi^\alpha(s, X_s^\alpha)$ for all s.

- Bdr condition: $H(T, X_T^\alpha) = \phi(X_T^\alpha)$, thus.

  $$\phi(X_T^\alpha) \leq H(t,x) + \int_t^T -\Psi^\alpha(s, X_s^\alpha)\, ds + \int_t^T -\, dB_s$$

  Taking expectation $\mathbb{E}_{t,x}$:

$$H(t,x) \geq \mathbb{E}_{t,x}\left[\int_t^T \psi^\alpha(s, X_s^\alpha)\,ds + \phi(X_T^\alpha)\right] = J^\alpha(t,x).$$

Taking sup :

$$H(t,x) \geq \sup_\alpha J^\alpha = V(t,x).$$

" $H \leq V$ ".

Take specifically $\alpha = g$. similarly, since

$$\frac{\partial H}{\partial t} + \psi^\alpha + \mathcal{L}^g H = 0.$$

(*) yields

$$H(t,x) = \mathbb{E}_{t,x}\left[\int_t^T \psi^g(s, X_s^g)\,ds + \phi(X_T^g)\right] = J^g(t,x).$$

Trivially, $V(t,x) \geq J^g(t,x) = H(t,x)$.

$\Longrightarrow$ $V = H$, and g is the optimal control. $\boxtimes$

Remark 9.8. (On HJB).

i). If we consider an "inf" problem instead of "sup":

$$V(t,x) = \inf_\alpha J^\alpha(t,x).$$

then HJB still holds with the term

$$\inf_\alpha \{ \psi^\alpha(t,x) + \mathcal{L}^\alpha V(t,x) \}.$$

ii). HJB generalises to the "indefinite time horizon" case naturally, with τ the hitting time of Γ_T.

$\boxed{\text{How to use HJB in 3 steps.}}$

1. Write down HJB corresponding to your problem.
2. Fix (t,x), find where $\sup_\alpha \left(\psi^\alpha(t,x) + \mathcal{L}^\alpha V(t,x) \right)$ is attained.

   call the solution α^*, our candidate for the optimal control.

3. plug α^* in HJB, solve the resulting PDE, use the verification thm to identify V and α^*.

But how to solve the PDE in 3?

— In general very difficult.

— make ansatz.

— consider structural properties of ϕ and ψ.

"Guess" and "Verify"