Chap3：Pontryagin 极大值原理.

Hamiltonian 方程： $H(x,p,a) := (Mx + N\alpha)\cdot p$ $\alpha \in A = [-1,1]^m$

假设：① 无 running payoff ② 最小化到达时间 ③ 线性问题

p^*：最优协态变量
costate

Thm 3.4 (PMP) 令 α^* 为时间最优控制，x^* 为对应的最优状态.
⇒ 存在 p^* (最优协状态) 使得

(ODE) $$\dot{x}^*(t) = \nabla_p H(x^*(t), p^*(t), \alpha^*(t))$$

协态方程
adjoint equations

(ADJ) $$\dot{p}^*(t) = -\nabla_x H(x^*(t), p^*(t), \alpha^*(t))$$

(PMP) $$H(x^*(t), p^*(t), \alpha^*(t)) = \max_{a\in A} H(x^*(t), p^*(t), \alpha)$$

Thm 4.1 (Euler–Lagrangian 方程)

$$\frac{d}{dt}\left[\nabla_v L(x^*(t), \dot{x}^*(t))\right] = \nabla_x L(x^*(t), \dot{x}^*(t))$$

其中，$L(x,\dot{x})$ 为 Lagrangian 方程.

Thm 4.2 (Hamilton dynamics) $(x(\cdot), p(\cdot))$ 为下列 E-L 方程的解.

$$\begin{cases} \dot{x}(t) = \nabla_p H(x(t), p(t)) \\ \dot{p}(t) = -\nabla_x H(x(t), p(t)) \end{cases}$$

且 $H(x(t), p(t))$ 是关于时间 t 的常数.

最优化问题：

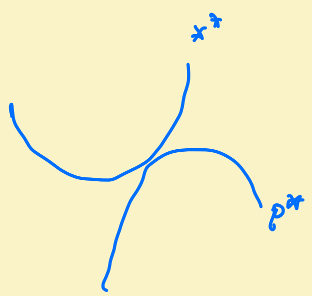
$$\begin{cases} \max\ f(x) \\ \text{s.t.}\ \ g(x) \le 0 \end{cases}$$

$L(x;\lambda) = f(x) - \langle \lambda, g(x) \rangle$

一阶最优性条件：

$$\nabla f(x^*) = \lambda \nabla g(x^*)$$

Lagrangian 乘子.



4.3.1 时间固定，终端自由

控制集： $A \subseteq \mathbb{R}^m$ $\mathcal{A} = \{\alpha(\cdot) \mid [0,+\infty) \to A \mid \alpha(\cdot)$ 是可测的$\}$

状态函数 $f: \mathbb{R}^n \times A \to \mathbb{R}^n$.

(ODE)
$$\begin{cases} \dot{x}(t) = f(x(t), \alpha(t)) & t > 0 \\ x(0) = x_0 \end{cases}$$

收益函数

$$P[\alpha(\cdot)] = \int_0^T \underbrace{r(x(t), \alpha(t))\,dt}_{\text{运行收益}} + \underbrace{g(x(T))}_{\text{终端收益}}$$

问题：找到最优控制 $\alpha^*(\cdot)$ 使得收益函数最大化

$$P[\alpha^*(\cdot)] = \max_{\alpha(\cdot) \in \mathcal{A}} P[\alpha(\cdot)].$$

chap 3

$H(x, p, \alpha) = (Mx + N\alpha) \cdot p$

Hamiltonian 方程： $H(x, p, \alpha) := f(x, a) \cdot p + r(x, a)$

Thm 4.3 (PMP, 时间固定，终端自由版) 令 α^* 为最优控制，x^* 为对应的最优状态，则存在 p^* 使得

(ODE) $$\dot{x}^*(t) = \nabla_p H(x^*(t), p^*(t), \alpha^*(t))$$

(ADJ) $$\dot{p}^*(t) = -\nabla_x H(x^*(t), p^*(t), \alpha^*(t))$$

(PMP) $$H(x^*(t), p^*(t), \alpha^*(t)) = \max_{a \in A} H(x^*(t), p^*(t), \alpha)$$

此外，有 ① Hamiltonian 方程是关于时 t 的常数.

$$t \mapsto H(x^*(t), p^*(t), \alpha^*(t))$$ 为常数.

② 存在终端条件 (terminal condition)

$$p^*(T) = \nabla g(x^*(T))$$

Proof：可以将时间固定，终端自由的最优控制问题改写成如下最优化问题

$$\max_{\alpha} P[\alpha(\cdot)] = \int_0^T r(x(t),\alpha(t))\,dt + g(x(T))$$

$$s.t \quad \dot{x} = f(x,\alpha) \quad t>0$$

则存在伴随变量（可看作 Lagrangian 乘子）

$$\mathcal{L}(x,\alpha;p) = g(x(T)) + \int_0^T r(x(t),\alpha(t))\,dt + \int_0^T p\cdot(f(x,\alpha)-\dot{x})\,dt$$

$$\frac{\delta \mathcal{L}}{\delta x} = \nabla g(x(T))^T \delta x(T) + \int_0^T \nabla_x r\,\delta x\,dt + \int_0^T [\nabla_x f\cdot p - \dot{p}]\,\delta x\,dt - p\cdot\delta x\Big|_0^T = 0 \quad t>0$$

$$\Rightarrow \int_0^T (\nabla_x r + (\nabla_x f)\cdot p - \dot{p})\,\delta x\,dt = \nabla g(x(T))\,\delta x(T) - p^T\delta x(T)$$

$$\Rightarrow \quad \nabla_x r + \nabla_x f\cdot p - \dot{p} = 0 \quad \Rightarrow \quad \dot{p} = -\nabla_x f\cdot p - \nabla_x$$

$$= \nabla_x H(x(t), p(t), \alpha(t))$$

(ADJ)

$$\nabla g(x(T)) - p(T) = 0$$

$$\Rightarrow \quad p(T) = \nabla g(x(T))$$ （终端条件）

终端条件：横截条件（Transversality condition . 4.5）

$$\frac{\delta \mathcal{L}}{\delta p} = \int_0^T (f(x,\alpha)-\dot{x})\cdot\delta p\,dt = 0 \quad \Rightarrow \quad \dot{x} = f(x,\alpha) \quad \text{(ODE)}$$

Hamiltonian 方程关于 t 求导：

$$\frac{d}{dt}H = \nabla_x H\cdot\dot{x} + \nabla_p H\cdot\dot{p} = \nabla_x H\cdot\nabla_p H - \nabla_p H\cdot\nabla_x H = 0$$

因此 $H(x^*(t), p^*(t), \alpha^*(t)) \equiv C$

4.3.2 时间自由，终端固定.

控制集: $A \subseteq \mathbb{R}^m$ $\mathcal{A} = \{\alpha(\cdot) \mid [0,+\infty) \to A \mid \alpha(\cdot)\text{是可测的}\}$

状态函数 $f: \mathbb{R}^n \times A \to \mathbb{R}^n$.

(ODE)
$$\begin{cases} \dot{x}(t) = f(x(t), \alpha(t)) & t>0 \\ x(0) = x_0 \end{cases}$$

收益函数

$$P[\alpha(\cdot)] = \int_0^{\tau[\alpha(\cdot)]} r(x(t), \alpha(t))\,dt$$

→运行收益

τ: 为首次到达状态 x_1 的时间.

问题: 找到最优控制 $\alpha^*(\cdot)$ 使得收益函数最大化

$$P[\alpha^*(\cdot)] = \max_{\alpha(\cdot)\in\mathcal{A}} P[\alpha(\cdot)].$$

Hamiltonian 方程: $H(x,p,\alpha) := f(x,a)\cdot p + r(x,a)$

Thm 4.4 (PMP, 时间自由, 终端固定版) 令 α^* 为最优控制, x^* 为对应的最优状态, 则存在 p^* 使得

(ODE) $$\dot{x}^*(t) = \nabla_p H(x^*(t), p^*(t), \alpha^*(t)).$$

(ADJ) $$\dot{p}^*(t) = -\nabla_x H(x^*(t), p^*(t), \alpha^*(t))$$

(PMP) $$H(x^*(t), p^*(t), \alpha^*(t)) = \max_{a\in A} H(x^*(t), p^*(t), a).$$

此外, 有 $H(x^*(t), p^*(t), \alpha^*(t)) \equiv 0$.

Proof: 对应优化问题:

$$\max_{\alpha(\cdot)} P[\alpha(\cdot)]$$

$$\text{s.t.} \begin{cases} \dot{x} = f(x,\alpha) & t>0 & \to P \\ \underline{x(\tau) = x_1} & & \to A \end{cases}$$

Lagrangian 方程：

$$\mathcal{L}(x,\alpha,\tau;p,\lambda)=\int_0^\tau [r(x,\alpha)+p\cdot(f(x,\alpha)-\dot{x})]dt+\lambda\cdot(x(\tau)-x_1)$$

$$\frac{\delta \mathcal{L}}{\delta x}=\int_0^\tau (\nabla_x r+\nabla_x f\cdot p-\dot{p})\delta x dt - p\cdot\delta x\Big|_0^\tau+\lambda\cdot\delta x(T) = 0$$

$$\Rightarrow \dot{p}=-\nabla_x r-\nabla_x f\cdot p \qquad p(T)=\lambda$$

$$\frac{\delta \mathcal{L}}{\delta p}=\int_0^\tau (f(x,\alpha)-\dot{x})\delta p dt \Rightarrow \dot{x}=f(x,\alpha)$$

$$\frac{\delta \mathcal{L}}{\delta \tau}=[r(x(T),\alpha(\tau))+p(\tau)\cdot f(x(\tau),\alpha(\tau))]\delta\tau$$

$$=H(x(T),p(\tau),\alpha(\tau))\delta\tau=0$$

$$\Rightarrow H(x^*(\tau),p^*(\tau),\alpha^*(\tau))=0 \qquad ①$$

又因为 $\frac{d}{dt}H=0$ $\quad H(x^*(t),p^*(t),\alpha^*(t))\equiv C.\ \forall t$ ②

由 ①-② 可得 $H\equiv 0$。

4.4. 例子.

① 线性时间最优控制.

$$A=[-1,1]^m \qquad \begin{cases}\dot{x}=Mx+N\alpha & t>0\\ x(0)=x_0\end{cases}$$

目标：最短时间到达 x_1

目标函数： $P[\alpha(\cdot)]=-\tau=-\int_0^\tau 1dt=\int_0^\tau -1dt$，此时运行收益 $r=-1$.

因此，Hamiltonian 方程： $H(x,p,a)=f(x,a)\cdot p+r(x,a)$

$$=(Mx+Na)p-1$$

根据 PMP 原理，可得.

ODE： $\dot{x}=\nabla_p H=Mx+Na$

ADJ： $\dot{p}=-\nabla_x H=-M^T p$

PMP: $H(x^*, p^*, \alpha^*) = \max_{a\in A} H(x^*, p^*, a)$

$$\alpha^* \in \arg\max_a (Mx + Na)p - 1$$

$$\Leftrightarrow \alpha^* \in \arg\max_a Nap$$

$$a_i^*(t) = \text{sign}\,(N^T p(t)_i) \quad i=1,\dots,m$$

② 生产—消费分配.

$x(t)$: t 时刻的产出.

$\alpha \in [0,1]$: 再投资的比例. 则 $(1-\alpha(t))$ 为 t 时刻的消费比例.

$$\begin{cases} \dot{x}(t) = \alpha(t)\cdot x(t) & t>0 \\ x(0) = x_0 > 0 \end{cases}$$

收益函数: $P[\alpha] = \int_0^T (1-\alpha(t))\, x(t)\, dt$. $= \int_0^T r(x,a)\,dt + g(x(T))$

Hamiltonian 方程: $H(x,p,a) = f(x,a)\,p + r(x,a)$

$$= pxa + (1-a)x = x + ax(p-1)$$

PMP:

(ODE) $\dot{x} = \nabla_p H = ax \qquad a\in[0,1]$

(ADJ) $\dot{p} = -\nabla_x H = -(1+\alpha(p-1))$

(PMP) $\alpha^* \in \arg\max_a H(x,p,a) = \arg\max_a ax(p-1)$ $\quad x>0$

终端条件 $p(T) = 0$ $\quad p(T) = \nabla_x g(x(T))$

求解最优控制 α:

$$\alpha(t) \begin{cases} 0 & \text{if } p-1 \le 0 \\ 1 & \text{if } p-1 > 0 \end{cases}$$

求解协状态 p, 有

$$\begin{cases} \dot{p}(t) = -1 - \alpha(t)(p(t) - 1) & \leftarrow \text{ADJ} \\ p(T) = 0 & \leftarrow \text{终端条件.} \end{cases}$$

因为 $p(T) = 0 < 1$, 且 T 的一个邻域内, 都有 $p(t) < 1$, 因此 $\alpha(t) = 0$

根据 (ADJ) 有 $\dot{p}(t) = -1$ 即 $\frac{dp}{dt} = -1 \Rightarrow p(t) = C - t$

联立 $\begin{cases} p(T) = 0 \\ p(t) = C - t \end{cases} \Rightarrow C = T \Rightarrow p(t) = T - t$ 当 $p(t) \le 1$

即 $t \in [T-1, T]$ 有 $p(t) = T - t$

当 $t\in[0, T-1)$ 有 $p(t)>1$ 此时 $\alpha(t)=1$

综上
$$\alpha(t)=\begin{cases}0 & \text{if } t\in[T-1, T]\\ 1 & \text{if } t\in[0, T-1)\end{cases}$$

即存在一个最优的转换时间 $t^*=T-1$

③ 一维线性二次正则项
$$\begin{cases}\dot{x}(t)=x(t)+\alpha(t)\\ x(0)=x_0\end{cases} \qquad A=\mathbb{R}$$

收益函数:
$$P[\alpha(\cdot)]=-\int_0^T\left(x^2(t)+\alpha^2(t)\right)dt$$

Hamiltonian 方程:
$$H(x,p,\alpha)=fp+r$$
$$=(x(t)+\alpha(t))\,p(t)-(x^2(t)+\alpha^2(t))$$

PMP:

(ODE) $\dot{x}=\nabla_p H=x+\alpha$

(ADJ) $\dot{p}=-\nabla_x H=-(p-2x)=2x-p$

(PMP) $\alpha\in\arg\max_a H(x,p,a)$
$$=\alpha(t)p(t)-\alpha^2(t)$$

终端条件: $p(T)=\nabla_x g(x(T))=0$

求最优控制 α: $\nabla_\alpha H=p-2\alpha=0$
$$\Rightarrow \alpha^*(t)=\frac{1}{2}p^*(t)$$

$$\begin{cases}\dot{p}=2x-p\\ p(T)=0\end{cases} \qquad \dot{x}=x+\alpha=x+\frac{1}{2}p$$

得到新的动力系统:
$$\begin{cases}\dot{x}=x+\frac{1}{2}p\\ \dot{p}=2x-p\end{cases} \iff \begin{pmatrix}\dot{x}\\ \dot{p}\end{pmatrix}=\underbrace{\begin{pmatrix}1 & \frac{1}{2}\\ 2 & -1\end{pmatrix}}_{=M}\begin{pmatrix}x\\ p\end{pmatrix}$$

可得系统一般解： $\begin{pmatrix} x(t) \\ p(t) \end{pmatrix} = e^{tM} \begin{pmatrix} x_0 \\ p_0 \end{pmatrix}$ ← x_0 已知 ← 需要求 p_0

Riccati 变换 $d(t) := \dfrac{p(t)}{x(t)}$ 则

$$\alpha(t) = \frac{1}{2} p(t) = \frac{1}{2} d(t) x(t)$$

$$\begin{cases} \dot{x} = x + \frac{1}{2} p = x + \frac{1}{2} d x \\ \dot{p} = 2x - p = 2x - d x \end{cases}$$

$$\dot{d} = \frac{\dot{p} x - p \dot{x}}{x^2} = \frac{\dot{p}}{x} - \frac{p \dot{x}}{x^2} = (2-d) - d\left(1 + \frac{d}{2}\right)$$

$$= 2 - 2d - \frac{1}{2} d^2$$

$$d(T) = \frac{p(T)}{x(T)} = 0$$

可得 Riccati 方程 $\begin{cases} \dot{d} = 2 - 2d - \frac{1}{2} d^2 \\ d(T) = 0 \end{cases}$

找 α 的显式解．令 $z = d + 2$ (对 d 作平移)

代入 $d = z - 2$，$\dot{z} = \dot{d}$

$$\dot{z} = 4 - \frac{1}{2} z^2$$

$$\Rightarrow \frac{dz}{dt} = \frac{8 - z^2}{2} \Rightarrow \frac{2}{8 - z^2} dz = dt$$

$$\int \frac{dx}{c^2 - x^2} = \frac{1}{2c} \ln\left|\frac{c+x}{c-x}\right|$$

$$\Rightarrow \frac{1}{2\sqrt{2}} \ln \frac{2\sqrt{2} + z}{2\sqrt{2} - z} = t + C \Rightarrow \frac{2\sqrt{2} + z}{2\sqrt{2} - z} = K e^{at}$$

$$z(t) = 2\sqrt{2} \tanh\left(\sqrt{2} t + c'\right)$$

由于 $d(T) = 0 \Rightarrow z(T) = 2$

$$z(T) = 2\sqrt{2} \tanh\left(\sqrt{2} T + c'\right) = \frac{1}{\sqrt{2}}$$

$$c' = \operatorname{arctanh}\left(\frac{1}{\sqrt{2}}\right) - \sqrt{2} T$$

$$d(t) = -2 + 2\sqrt{2} \tanh\left(\sqrt{2}(t - T) + \operatorname{arctanh} \frac{1}{\sqrt{2}}\right)$$

$$\alpha^* = \frac{d(t)}{2} x(t)$$