Chp4 Pontryagin 极大值原理：时间T固定，终端自由

$$\begin{cases} \dot{x}(t) = f(x(t), \alpha(t)) & 0 < t \le T \\ x(0) = x_0 \end{cases}$$

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

Thm: (PMP 时间T固定 终端自由) 设 $\alpha^*(\cdot)$ 是最优控制，$x^*(\cdot)$ 为对应状态轨迹，则存在协态变量 p^* 满足

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

终端条件：$p^*(T) = \nabla g(x^*(T))$

Chp6 微分博弈设定

$$\begin{cases} \dot{x}(s) = f(x(s), \alpha(s), \beta(s)) & t < s \le T \\ x(t) = x \end{cases}$$

Payoff: $P_{x,t}[\alpha(\cdot), \beta(\cdot)] = \int_t^T r(x(s), \alpha(s), \beta(s)) + g(x(T))$

设定：双人微分零和博弈

下值函数 $v(x,t) = \inf_{\Psi \in \mathcal{B}} \sup_{\alpha(\cdot) \in A} P_{x,t}[\alpha(\cdot), \Psi[\alpha(\cdot)]]$

上值函数 $u(x,t) = \sup_{\Phi \in \mathcal{A}} \inf_{\beta(\cdot) \in B} P_{x,t}[\Phi[\beta(\cdot)], \beta(\cdot)]$

博弈有值 $v(x,t) = u(x,t)$

Isaacs 条件

$$\left.\begin{aligned}&\max_{a\in A}\min_{b\in B}\{f(x,a,b)\cdot p + r(x,a,b)\}\\ =\ &\min_{b\in B}\max_{a\in A}\{f(x,a,b)\cdot p + r(x,a,b)\}\end{aligned}\right\}\ (1)$$

引入 Hamiltonian:

$$H(x,p,a,b) := f(x,a,b)\cdot p + r(x,a,b)$$

式 (1) 可以写成

$$H(x,p) = \max_{a\in A}\min_{b\in B} H(x,p,a,b) = \min_{b\in B}\max_{a\in A} H(x,p,a,b)$$

在 Isaacs 条件成立条件下，设 α^*, β^* 为最优控制.

Thm (PMP, 微分博弈) 设 Isaacs 条件成立. (α^*, β^*) 为最优控制对时，对应的轨线 x^*. 令

$$p^*(t) = \nabla_x v(x^*(t), t)$$

则 $(x^*, p^*, \alpha^*, \beta^*)$ 满足

(ODE) $$\begin{aligned}\dot{x}^*(t) &= \nabla_p H(x^*(t), p^*(t), \alpha^*(t), \beta^*(t))\\ &= f(x^*(t), \alpha^*(t), \beta^*(t))\end{aligned}$$

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

(Saddle point condition) $$H(x^*, p^*, \alpha, \beta^*) \le H(x^*, p^*, \alpha^*, \beta^*) \le H(x^*, p^*, \alpha^*, \beta)$$

终端条件 $$p^*(T) = \nabla_x g(x^*(T))$$

$$\begin{cases} H(x^*, p^*, \alpha^*, \beta^*) = \max_{\alpha\in A} H(x^*, p^*, \alpha, \beta^*)\\ H(x^*, p^*, \alpha^*, \beta^*) = \min_{\beta\in B} H(x^*, p^*, \alpha^*, \beta)\end{cases}$$

Proof: 根据 Hamiltonian 的定义

$$H(x,p,a,b) = f(x,a,b)\cdot p + r(x,a,b)$$

可得 $\nabla_p H(x,p,a,b) = f(x,a,b)$

对 Isaacs 方程 $v_t + H(x,\nabla_x v) = 0$ 对 x 求梯度

$$\nabla_x v_t(x,t) + \nabla_x H(x,\nabla_x v) = 0$$

记 $p(x,t) = \nabla_x v(x,t)$

$$\nabla_x H(x,p) = H_x(x,p) + H_p(x,p)\cdot \nabla_x p(x,t)$$

代回可得

$$\cancel{\nabla_x v_t}\ p_t + H_x + H_p\cdot\nabla_x p = 0$$

考虑最优轨线 x^*，则 $p^*(t) = p(x^*(t),t)$，对 $p^*(t)$ 关于 t 求导，

$$\dot{p}^*(t) = p_t(x^*(t),t) + \nabla_x p(x^*(t),t)\cdot\dot{x}^*(t)$$

代回可得

$$\dot{p}^*(t) = -H_x(x^*(t),p^*(t)) + \nabla_x p(x^*(t),t)\cdot\left[\dot{x}^*(t) - H_p(x^*(t),p^*(t))\right]$$

由于 $\dot{x}^*(t) = H_p(x^*,p^*)$，可得

$$\dot{p}^*(t) = -H_x(x^*(t),p^*(t))$$

终端条件：$v(x,T) = g(x)$，且 $p^*(t) = \nabla_x v(x(t),t)$，可得

$$p^*(T) = \nabla_x v(x(T),t) = \nabla_x g(x(T))$$

6.4. 消耗战，攻击模型

状态：$x_1(t)$：玩家 I 在 t 时刻的资源（武器，燃料补给）

$x_2(t)$：玩家 II 在 t 时刻的资源

控制：控制集 $A = B = [0,1]$

$\alpha(t)\in A$：玩家 I 用于消耗战的比例。

$1-\alpha(t)$：玩家 I 用于直接攻击的比例。

$\beta(t)$：相同含义

参数：$m_1 > 0$：玩家 I 的资源生产率

$m_2 > 0$：玩家 II 的资源生产率

$c_1>0$：玩家II对玩家I的消耗能力

$c_2>0$：玩家I对玩家II的消耗能力

假设 $c_2>c_1$：即，玩家I在削弱对手战争潜力上的效率更强

$$\begin{cases} \dot{x}_1(t) = m_1 - c_1\beta(t)x_2(t) \\ \dot{x}_2(t) = m_2 - c_2\alpha(t)x_1(t) \end{cases}$$

收益函数 $P[\alpha(\cdot),\beta(\cdot)] = \int_0^T [\underbrace{(1-\alpha(t))x_1(t)}_{\text{玩家I的直接攻击收益}} - \underbrace{(1-\beta(t))x_2(t)}_{\text{玩家II的直接攻击损失}}]dt$

玩家I：P最大化，玩家II：P最小化。

求解：Hamiltonian 为 $H(x,p,a,b) = p\cdot f(x,a,b) + r(x,a,b)$

$$f(x,a,b) = \begin{pmatrix} m_1 - c_1 b x_2 \\ m_2 - c_2 a x_1 \end{pmatrix} \qquad r(x,a,b) = (1-a)x_1 - (1-b)x_2$$

$$\Rightarrow H(x,p,a,b) = p_1\cdot(m_1 - c_1 b x_2) + p_2(m_2 - c_2 a x_1) + (1-a)x_1 - (1-b)x_2$$

$$= (x_1 - x_2) + m_1 p_1 + m_2 p_2 + a(-x_1 - c_2 x_1 p_2) + b(x_2 - c_1 x_2 p_1)$$

令 $A(x,p) = -x_1(1+c_2p_2)$ $\qquad B(x,p) = x_2(1-c_1p_1)$

$C(x,p) = (x_1 - x_2) + m_1p_1 + m_2p_2$

由于 a,b 与 $A(x,p)$, $B(x,p)$ 的依赖关系线性，可得

$$a^*(x,p) = \begin{cases} 1 & A(x,p)\geq 0 \\ 0 & A(x,p)<0 \end{cases} \qquad b^*(x,p) = \begin{cases} 1 & B(x,p)<0 \\ 0 & B(x,p)\geq 0 \end{cases}$$

又因为 $H(x,p,a,b) = C(x,p) + a\cdot A(x,p) + b\cdot B(x,p)$

$$\max_{a\in A}\min_{b\in B} H(x,p,a,b) = \max_{a\in A}\min_{b\in B}[C(x,p) + a\cdot A(x,p) + b\cdot B(x,p)]$$

$$= C(x,p) + \max_{a\in A}[a\cdot A(x,p)] + \min_{b\in B}[b\cdot B(x,p)]$$

$$= \min_{b\in B}\max_{a\in A} H(x,p,a,b)$$

因此，Isaacs 条件成立.

应用极大值原理： $\dot{p}(t) = -\nabla_x H(x, p, \alpha, \beta)$ 可得

$$\dot{p}_1(t) = \alpha(t) - 1 + c_2 \cdot p_2(t) \cdot \alpha(t) = -1 + (1 + c_2 p_2(t))\alpha(t)$$
$$\dot{p}_2(t) = 1 - \beta(t) + c_1 \cdot p_1(t) \cdot \beta(t) = 1 + (-1 + c_1 p_1(t))\beta(t)$$

终端条件：
$$p(T) = \nabla_x g(x(T)) = 0 \quad \Rightarrow \quad p_1(T) = p_2(T) = 0$$

令：
$$s_1(t) = -1 - c_2 p_2(t) \qquad s_2(t) = 1 - c_1 p_1(t)$$

可得
$$\alpha(t) = \begin{cases} 1 & s_1(t) \geq 0 \\ 0 & s_1(t) < 0 \end{cases} \qquad \beta(t) = \begin{cases} 0 & s_2(t) \geq 0 \\ 1 & s_2(t) < 0 \end{cases}$$

由于 $A(x,p) = -x_1(1 + c_2 p_2)$ $\qquad B(x,p) = x_2(1 - c_1 p_1)$
$$\alpha(t) = \begin{cases} 1 & A(x,p) \geq 0 \\ 0 & A(x,p) < 0 \end{cases} \qquad \beta(t) = \begin{cases} 0 & B(x,p) \geq 0 \\ 1 & B(x,p) \leq 0 \end{cases}$$

且 $x_1, x_2 > 0$，那么 $\operatorname{sign}(A(x,p)) = \operatorname{sign} s_1$
$\operatorname{sign}(B(x,p)) = \operatorname{sign} s_2$

接下来推导演化方程：
$$\dot{s}_1(t) = -c_2 \dot{p}_2(t) = -c_2(1 - \beta + c_1 p_1 \beta)$$
$$\dot{s}_2(t) = -c_1 \dot{p}_1(t) = -c_1(\alpha - 1 + c_2 p_2 \alpha)$$

$$\dot{p}_1(t) = \alpha(t) - 1 + c_2 \cdot p_2(t) \cdot \alpha(t)$$
$$\dot{p}_2(t) = 1 - \beta(t) + c_1 \cdot p_1(t) \cdot \beta(t)$$

终端条件： $s_1(T) = -1 - c_2 p_2(T) = -1$
$s_2(T) = 1 - c_1 p_1(T) = 1$

整理可得
$$\begin{cases} \dot{s}_1 = c_2(-1 + \beta s_2) \\ \dot{s}_2 = c_1(1 + \alpha s_1) \end{cases} \qquad \begin{cases} s_1(T) = -1 \\ s_2(T) = 1 \end{cases}$$

$$\begin{cases} s_1(T) = 4 \\ s_2(T) = 1 \end{cases}$$

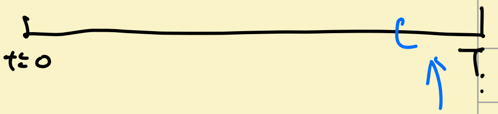
根据终端条件可知，

$$\alpha(T) = \beta(T) = 0$$

即双方都把所有资源用于直接攻击。

$$\begin{cases} \dot{s}_1 = c_2(-1+\beta s_2) \\ \dot{s}_2 = c_1(1+\alpha s_1) \end{cases}$$

代入



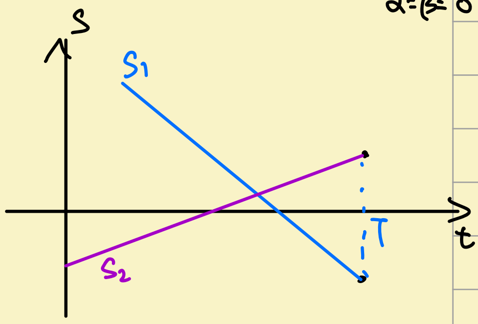
在末段区间内 $[t^*, T]$ 可取 $\alpha=\beta=0$，可得

$$\dot{s}_1 = -c_2 \qquad \dot{s}_2 = c_1$$

$\alpha=\beta=0$ 结合终端条件取积分得

$$s_1(t) = \int_t^T \dot{s}_1\, dt = 4 + c_2(T-t)$$

$$s_2(t) = \int_t^T \dot{s}_2\, dt = 1 + c_1(t-T)$$



令 $\begin{cases} s_1(t) = 0 \Rightarrow t = T - \frac{1}{c_2} \\ s_2(t) = 0 \Rightarrow t' = T - \frac{1}{c_1} \end{cases}$

由于 $c_2 > c_1$，$t' < t$，即玩家Ⅱ更早开始直接攻击。

第二阶段：$\alpha=1$, $\beta=0$

代入可得．$\begin{cases} \dot{s}_1(t) = c_2(-1+0\cdot s_2) = -c_2 \\ \dot{s}_2(t) = c_1(1+1\cdot s_1) \end{cases}$ $s_1(t^*) = 0$, $s_2(t^*) = 1 - \frac{c_1}{c_2}$

积分可得

$$s_1(t) = \int_t^{t^*} \dot{s}_1\, dt = -1 + c_2(T-t)$$

$$s_2(t) = \int_t^{t^{**}} \dot{s}_2\, dt = 1 - \frac{c_1}{2c_2} - \frac{c_1 c_2}{2}(t-T)^2$$

令 $s_2(t^{**}) = 0 \Rightarrow t^{**} = T - \frac{1}{c_2}\sqrt{\frac{2c_2}{c_1} - 1}$

第三阶段：$\alpha=1$, $\beta=1$