Methods to compute value function — continuous time

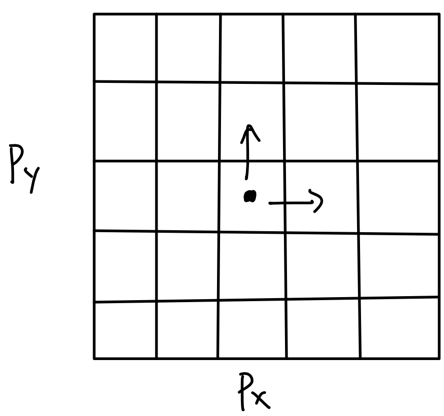
HJB-PDE : $\frac{\partial V(x,t)}{\partial t} + \min_u \{ L(x,u) + \frac{\partial V(x,t)}{\partial x} \cdot f(x,u) \} = 0$

$$V(x,T) = \ell(x)$$

(1) Closed-form computation of V

- LQR
- can be very restrictive

(2) Mesh/grid-based methods : These methods solve the PDE numerically over a grid of state and time.



$\frac{\partial V}{\partial x}$, $\frac{\partial V}{\partial t}$ are approximately numerically over the grid

- level set toolbox / helperOC (Matlab)
- BEACLS (C++)
- OptimizeDP (Python)
- julia Reach (Julia)

(3) Function approximation-based methods

- represent the value function as a parametrized function. compute the parameters to satisfy the PDE as closely as possible.

$$\theta^* = \arg\min_\theta \left(\left\| \frac{\partial V_\theta(x,t)}{\partial t} + \min_u \{ L(x,u) + \frac{\partial V_\theta(x,t)}{\partial x} \cdot f(x,u) \} \right\| + \lambda \left\| V_\theta(x,t) - \ell(x) \right\| \right)$$

Physics-informed Neural Network

Deep Reach

other approximation :

- polynomials
- zonotopes
- ellipsoids

Robust Optimal Control → zero-sum dynamic games

$$\min_{u(\cdot)\in\Gamma_u} \max_{d(\cdot)\in\Gamma_d} J(x(t), u(\cdot), d(\cdot), t)$$

$$\text{s.t.} \quad \dot{x}(s) = f(x(s), u(s), d(s))$$

$$\underline{u} \le u(s) \le \overline{u}$$

$$\underline{d} \le d(s) \le \overline{d}$$

→ this is a game between control and disturbance

→ dynamic game (evolving over time)

→ zero-sum game

- Control is trying to minimized the worst-case cost incurred under uncertainty.

instance of zero-sum dynamic games

Rufus Isaacs (1950)

- The order of play is important max min ≠ min max

- d has a slight informational advantage

Importance of information pattern

B_1

S1	S2
2000	10

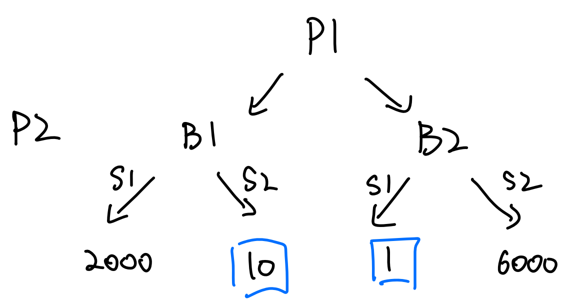
B_2

S1	S2
1	6000

P1: choose which box to open (maximize reward)

P2: choose slot (minimize reward)

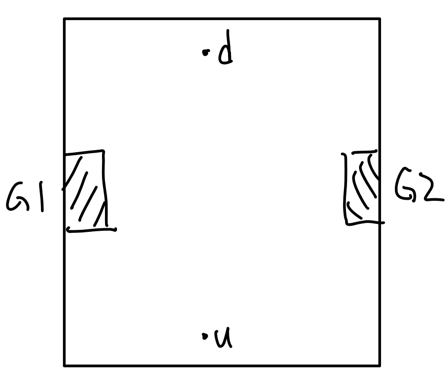
suppose P1 goes first



$$\max_{u\in\{b1,b2\}}\left(\min_{d\in\{s_1,s_2\}} R(u,d)\right) = 10$$

play first second

P1 should pick B1, P2 picks S2.



$u \to$ wants to reach G1/G2 without being intercepted by d

$d \to$ intercept u

(1) $\min_{u(\cdot)} \max_{d(\cdot)} J \cdots \}$ overly pessimistic

(2) $\max_{d(\cdot)} \min_{u(\cdot)} J \cdots \}$ overly optimistic

open-loop formulation

close-loop formulation

At any given time t, only information upto time t is known.
Such strategies are called non-anticipative strategies.

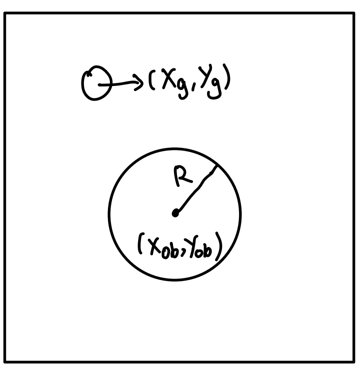
$$V(x,t) = \min_{u(\cdot)\in\Gamma_u} \max_{d(\cdot)\in\Gamma_d} J(x, u(\cdot), d(\cdot), t)$$

$$= \min_{u(\cdot)\in\Gamma_u} \max_{d(\cdot)\in\Gamma_d} \{ L(x, u(t), d(t))\delta + V(x(t+\delta), t+\delta) \}$$

$$\boxed{\begin{aligned} &\frac{\partial V(x,t)}{\partial t} + \min_{u(t)} \max_{d(t)} \{ L(x, u(t), d(t)) + \frac{\partial V(x,t)}{\partial x} \cdot f(x,u,d) \} = 0 \\ &V(x,T) = \ell(x) \end{aligned}}$$

Hamilton - Jacobi - Isaacs PDE

Demo:



Dynamics:

$$\dot{P}_x = V_x + d_x$$
$$\dot{P}_y = V_y + d_y$$

Control: V_x, V_y $\quad |V_x|, |V_y| \leq 1$ m/s

disturbance: d_x, d_y $\quad |d_x|, |d_y| \leq 0.2$ m/s

Cost function:

$$J(x, u(\cdot), d(\cdot), t) = \int_{s=t}^{T} \left(\text{dist}^2(x(s), \text{goal}) + \lambda \cdot \text{obs-penetration}(x(s)) \right) ds$$
$$+$$
$$\left(\text{dist}^2(x(T), \text{goal}) + \lambda \cdot \text{obs-pene}(x(T)) \right)$$

$$\text{dist}^2(x(s), \text{goal}) = (P_x(s) - X_g)^2 + (P_y(s) - Y_g)^2$$

$$\text{obs-pene} = \max\{R - \text{sqrt}((P_x(s) - X_{obs})^2 + (P_y(s) - Y_{obs})^2), 0\}$$

$$\text{Hamiltonian} = L(x) + \min_{u(t)} \max_{d(t)} \underbrace{\frac{\partial V(x,t)}{\partial x} \cdot f(x,u,d)}$$

$$\Downarrow$$

$$\begin{bmatrix} P_1(x) & P_2(x) \end{bmatrix} \begin{bmatrix} V_x + d_x \\ V_y + d_y \end{bmatrix}$$

$$= P_1(x)V_x + P_2(x)V_y + P_1(x)d_x + P_2(x)d_y$$