Math 5588

Recall: Optimal Control

Dynamics

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

$\alpha : [0,T] \to A$ control

$x : [0,T] \to \mathbb{R}^n$ state of system

Cost

$$C_{x,t}(\alpha) = \int_t^T r(x(s), \alpha(s))\,ds + g(x(T))$$

$r(x,a)$ = running cost

$g(x)$ = Terminal cost

Value Function

$$u(x,t) = \min_{\alpha : [0,T] \to A} C_{x,t}(\alpha)$$

We proved last time that $u(x,t)$ is the viscosity solution of the Hamilton-Jacobi-Bellman Equation

$$(HJB) \begin{cases} u_t + H(\nabla u, x) = 0, & 0 \le t < T \\ u(x,T) = g(x), & x \in \mathbb{R}^n \end{cases}$$

where
$$H(p,x) = \min_{a \in A} \{ p \cdot f(x,a) + r(x,a) \}.$$

We'll show now how this recovers the Hopf-Lax formula for special choices of f and r. We choose $A = \mathbb{R}^n$,

$$r(x,a) = L(a) \quad \text{and} \quad f(x,a) = -a$$

Then

$$H(p,x) = \min_{a \in \mathbb{R}^n} \{ -p \cdot a + L(a) \} = -\max_{q \in \mathbb{R}^n} \{ p \cdot q - L(q) \}.$$

The <u>Legendre transform</u> on $\mathbb{R}^n$ is

$$L^*(p) = \max_{q \in \mathbb{R}^n} \{ p \cdot q - L(q) \}.$$

Hence

$$H(p,x) = -L^*(p)$$

Thus, for

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

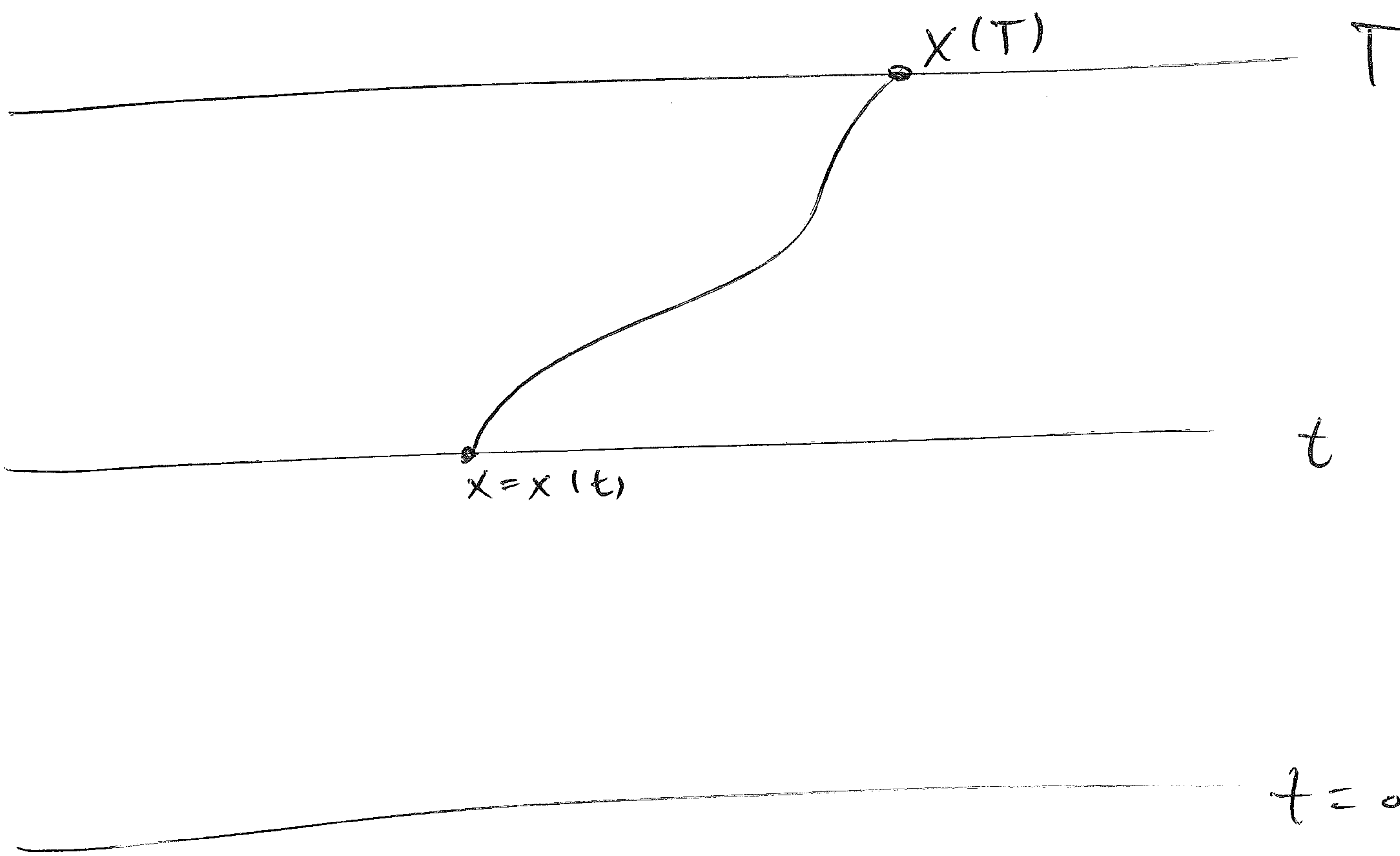
$$u(x,t) = \min_{\alpha : [0,T] \to \mathbb{R}^n} \int_t^T L(\alpha(s))\, ds + g(x(T))$$

Solves

$$\begin{cases} u_t - L^*(\nabla u) = 0, & \cancel{s \le t < T} \quad 0 \le t < T \\ u(x,T) = g(x). \end{cases}$$

By choosing $\alpha(s) := -\dot{x}(s)$ for a path $x:[0,T] \longrightarrow \mathbb{R}^n$ with $x(t)=x$ we can control the system to any path we choose. Hence

$$u(x,t) = \min_{\substack{x:[0,T]\longrightarrow \mathbb{R}^n \\ x(t)=x}} \left\{ \int_t^T L(-\dot{x}(s))\,ds + g(x(T)) \right\}$$



We claim that for L convex, the optimal path from x to $x(T)$ is always the straight line

$$x(s) = x\left(\frac{T-s}{T-t}\right) + x(T)\left(\frac{s-t}{T-t}\right)$$

Here $\dot{x}(s) = \dfrac{x(T) - x}{T-t}$ and

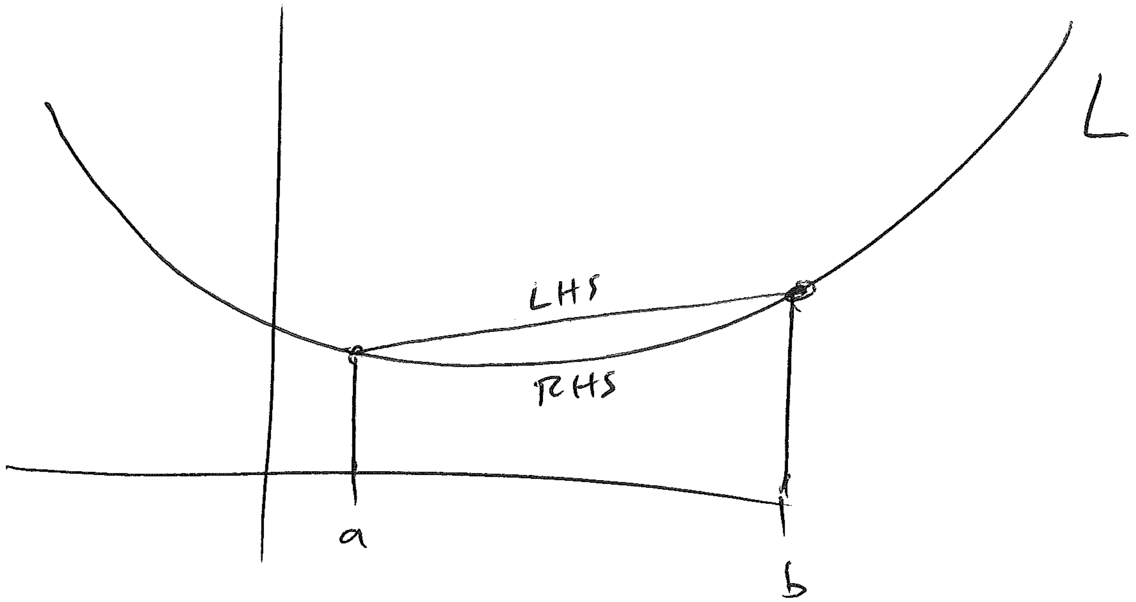
$$\int_t^T L(-\dot{x}(s))\,ds = \int_t^T L\left(\frac{x - x(T)}{T-t}\right) ds$$

$$= (T-t)\,L\left(\frac{x - x(T)}{T-t}\right)$$

To see that this is optimal we use Jensen's inequality

$$\frac{1}{T-t}\int_t^T L(y(s))\,ds \geq L\left(\frac{1}{T-t}\int_t^T y(s)\,ds\right)$$

which holds when L is convex



If $n=1$ and $a \le y(s) \le b$ then the average $\frac{1}{T-t}\int_t^T L(y(s))\,ds$ lies above the graph of L between a and b, while $L\left(\frac{1}{T-t}\int_t^T y(s)\,ds\right)$ lies on the graph, as $a \le \frac{1}{T-t}\int_t^T y(s)\,ds \le b$.

Applying Jensen's inequality here

$$\int_t^T L(-\dot{x}(s))\,ds = (T-t)\frac{1}{T-t}\int_t^T L(-\dot{x}(s))\,ds$$

$$\geq (T-t)\,L\left(\frac{1}{T-t}\int_t^T -\dot{x}(s)\,ds\right)$$

$$= (T-t)\,L\left(\frac{x(t)-x(T)}{T-t}\right)$$

$$= (T-t)\,L\left(\frac{x-x(T)}{T-t}\right)$$

for <u>any</u> path $x(s)$. Thus straight lines are minimal, and we only need to minimize over the terminal point

$$y = x(T).$$

Hence,

$$u(x,t) = \min_{y\in\mathbb{R}^n}\left\{(T-t)\,L\left(\frac{x-y}{T-t}\right) + g(y)\right\}.$$

solves

$$\begin{cases} u_t - L^*(Du) = 0, & 0 \le t < T \\ u(x,T) = g(x). \end{cases}$$

Now we reverse time to get an initial value problem

$$v(x,t) := \cancel{} = u(x, T-t)$$

Then $v_t(x,t) = -u_t(x, T-t)$

$$\nabla v(x,t) = Du(x, T-t)$$

and so
$$\begin{cases} v_t + L^*(\nabla v) = 0, & t > 0 \\ v(x,0) = g(x) \end{cases}$$

and $v(x,t) = \min\limits_{y \in \mathbb{R}^n} \left\{ t L\left(\frac{x-y}{t}\right) + g(y) \right\}.$

This is exactly the multi-dimensional version of the Hopf-Lax formula!

Math 5588

Differential Games

Differential games are optimal control problems where we now have two players A and B, both with their own controls $\alpha(s)$ and $\beta(s)$, and the players have competing objectives.

Example (Homicidal Chauffeur)

A homicidal chauffeur driving a car is attempting to run over a pedestrian. The car can travel much faster than the pedestrian, but has a minimal turning radius, while the pedestrian is highly maneuverable but slow.

(Aside: In other applications car = missile
pedestrian = airplane).

Let P = Homicidal chauffeur (pursuer)

E = pedestrian (evader).

(x_p, y_p) = position of pursuer $(x_p(t), y_p(t))$

(x_e, y_e) = position of evader $(x_e(t), y_e(t))$.

A simple model is that pursuer moves with speed 1 and can turn with a maximum turning radius of R.

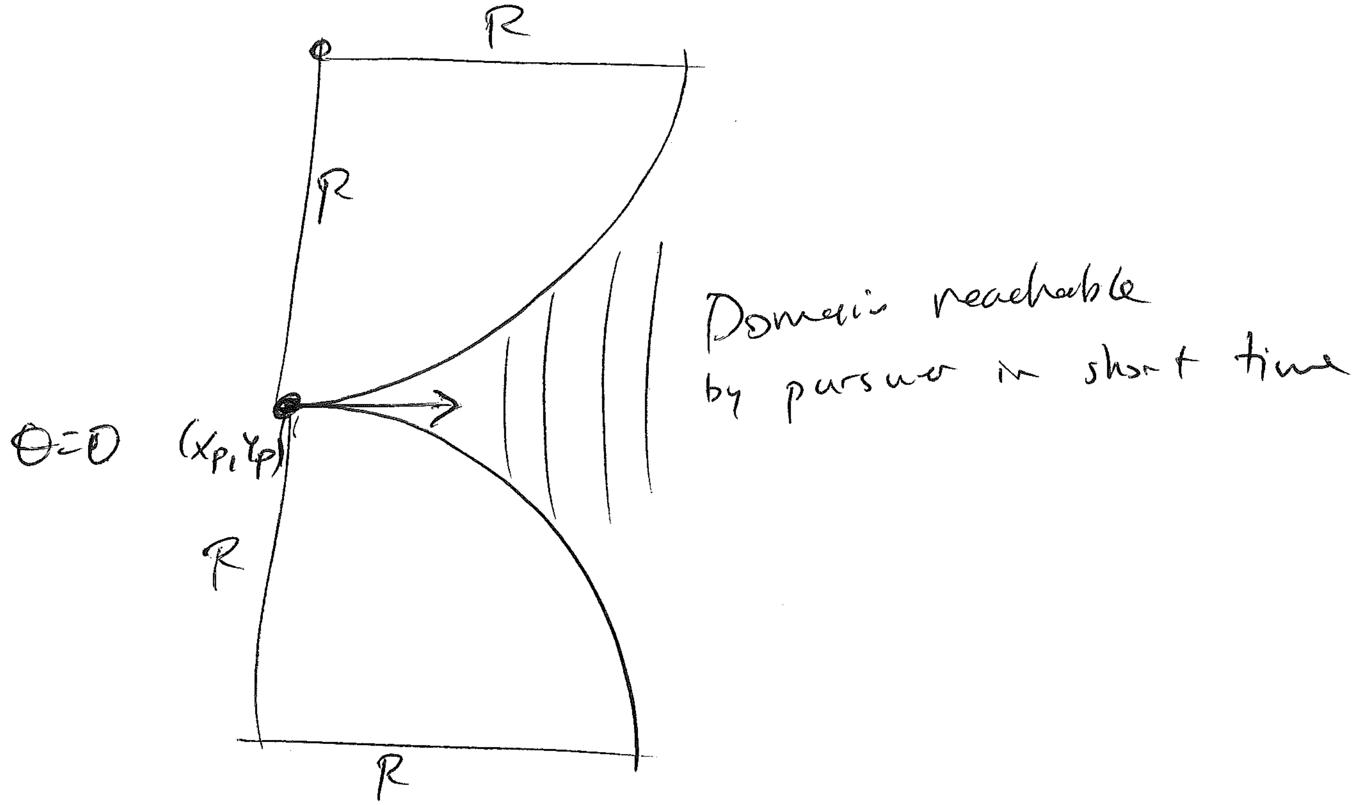
Hence

$$P: \begin{cases} \dot{x}_p(s) = \cos\theta \\ \dot{y}_p(s) = \sin\theta \\ \dot{\theta}(s) = \dfrac{\alpha(s)}{R} \end{cases}$$

$(x_p(s), y_p(s), \theta(s))$ is the state of the pursuer.

$\alpha(s)$ = control of pursuer, and

$$\boxed{|\alpha(s)| \leq 1}$$



The evader E is highly maneuverable, but has slow maximum speed.

$$E: \begin{cases} \dot{x}_e(s) = \beta_1(s) \\ \dot{y}_e(s) = \beta_2(s) \end{cases} \qquad \boxed{\beta_1^2 + \beta_2^2 \le \varepsilon}$$

State of evader is (x_e, y_e)

and $\beta(s) = (\beta_1(s), \beta_2(s))$ is

the control for evader.

The game terminates when

$$(x_p - x_e)^2 + (y_p - y_e)^2 \leq \delta^2,$$

that is, the pursuer comes within distance $\delta > 0$ of the evader.

<u>Question</u>: Under what circumstances can the purser always catch the evader, and when can the evader indefinitely avoid the homicidal chauffeur?

A general <u>differential game</u> involves a system

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

$x(s) \in \mathbb{R}^n$ = state of system

$\alpha : [0,T] \longrightarrow A$ is player 1 control

$\beta : [0,T] \longrightarrow B$ is player 2 control.

We have a cost

$$C_{x,t}(\alpha, \beta) = \int_t^T r(x(s), \alpha(s), \beta(s))\, ds + g(x(T))$$

$r(x,a,b)$ = running cost

$g(x)$ = terminal cost.

Player 1's goal is to <u>minimize</u> $C_{x,t}$ while player 2's goal is to <u>maximize</u> $C_{x,t}$

This is a <u>zero-sum</u> game, meaning any gain by one player is a loss by the other.

<u>Strategies</u>

Let
$$\mathcal{A} = \{\alpha : [0,T] \longrightarrow A\}$$
$$\mathcal{B} = \{\beta : [0,T] \longrightarrow B\}.$$

<u>Definition</u>: A <u>strategy</u> for player 1 is a map

$$S_1 : \mathcal{B} \longrightarrow \mathcal{A}$$

That is, player 1 observes player 2's control choice $\beta \in \mathcal{B}$ and decides on what $\alpha \in \mathcal{A}$ to choose.

__Def__: A strategy $S_1 : \mathcal{B} \longrightarrow \mathcal{A}$ is __non-anticipating__ if for any $0 < t \leq T$ and $\beta, \tilde{\beta} \in \mathcal{B}$ with $\beta(s) = \tilde{\beta}(s)$ for all $s \leq t$ we have

$$S_1[\beta](s) = S_1[\tilde{\beta}](s), \quad s \leq t$$

That is, at time t, player 1's strategy can only make use of previous knowledge of player 2's control β.

Let

$$\Delta_1 = \left\{ \text{set of } \underline{\text{non-anticipating}} \ \underline{\text{strategies}} \text{ for player 1} \right\}.$$

Non-anticipating strategies for player 2 are defined accordingly and

$\Delta_2 = \{$ non-anticipating strategies for player 2 $\}$

Definition The lower value of the game is

$$\underline{u}(x,t) = \min_{S \in \Delta_1} \max_{\beta \in B} C_{x,t}(S(\beta), \beta)$$

The upper value of the game is

$$\overline{u}(x,t) = \max_{S \in \Delta_2} \min_{\alpha \in A} C_{x,t}(\alpha, S(\alpha)).$$

It turns out that $\underline{u} \leq \overline{u}$, but this is not obvious yet.

Def: If $\overline{u} = \underline{u}$ we say the game has a value

$$u := \overline{u} = \underline{u}.$$

Note: $x(s)$ depends on $\alpha(s)$ and $\beta(s)$!

Thm (Dynamic programming principle).

For any $0 < t < t+h < T$ we have

$$\bar{u}(x,t) = \max_{S \in \Delta_2} \min_{\alpha \in A} \left\{ \int_t^{t+h} r(x(s), \alpha(s), S[\alpha](s))\, ds + \bar{u}(x(t+h), t+h) \right\}$$

and

$$\underline{u}(x,t) = \min_{S \in \Delta_1} \max_{\beta \in B} \left\{ \int_t^{t+h} r(x(s), S[\beta](s), \beta(s))\, ds + \underline{u}(x(t+h), t+h) \right\}$$

We omit the proof, which is similar to the dynamic programming principle in optimal control.

The dynamic programming principle allows us to derive PDEs for $\bar{u}$ and $\underline{u}$.