Lecture 8: Optimal Control Fundamentals

Last time:

Intro to Optimal Control

- Control design as solving an optimization problem
- Quality of the design evaluated via some metric of "cost"

Today:

- Discrete time optimal control
- Bellman's principle of optimality
- Continuous time analog
- Linear Quadratic Regulator (LQR)

- Control Design as optimization.

  - System: $\dot{x} = F(t, x, u)$ (*) $\quad x(t) \in \mathbb{R}^n$ (state)

    $x(t_0) = x_0 \qquad u(t) \in \mathbb{R}^m$ (control) $\left[u(\cdot) : \mathbb{R} \to \mathbb{R}^m \right]$

    $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ (dynamics)

    $x_0 \in \mathbb{R}^n$ initial condition.

  - Cost Functional:

    Running Cost

    $$V(u(\cdot)) = \int_{t_0}^{t_f} l(t, x(t), u(t))\, dt + l_f(x(t_f))$$

    Terminal Cost

    where $x(t)$ is the solution to (*) when applying $u(\cdot)$ as input.

  - Optimal control problem:

    $$\min_{u(\cdot)} V(u(\cdot))$$

    $$\text{s.t.} \quad \left. \begin{array}{l} u(t) \in \mathcal{U}_t \\ x(t) \in \mathcal{X}_t \end{array} \right]$$

    sets of admissible controls & states

    $$x(t_0) = x_0$$

Pendulum Case:

System: $I\ddot{\theta} = \tau + mgl\sin\theta \Rightarrow x = \begin{bmatrix}\theta \\ \dot{\theta}\end{bmatrix} \quad \dot{x} = \underbrace{\begin{bmatrix}\dot{\theta} \\ \frac{1}{I}(\tau + mgl\sin\theta)\end{bmatrix}}_{f(x,u)}$

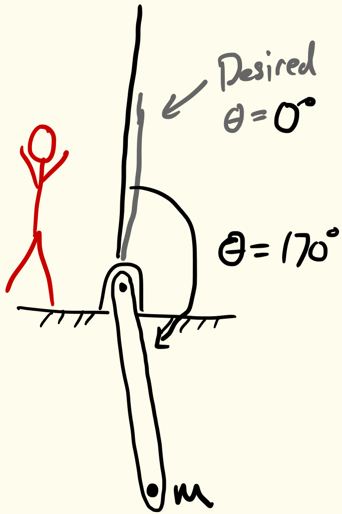
$\tau = u$

Cost Function: $\min_{u(\cdot)} \int_{t_0}^{\infty} \left[\theta^2_{(t)} + \dot{\theta}^2_{(t)}\right] + \tau^2_{(t)}\, dt$

Optimal Control problem: $\min_{u(\cdot)} \int_{t_0}^{\infty} \left[\theta^2(t) + \dot{\theta}^2(t) + \tau^2(t)\right] dt$

$\text{s.t. } |\tau(t)| \le \tau_{max}$

$\theta_{min} < \theta < \theta_{max}$



Terminology:

Continuous Time:

$$V(u(\cdot)) = \int_{t_0}^{t_f} \ell(t, x(t), u(t))\,dt + \ell_f(x(t_f))$$

Discrete time

$$V(u[\cdot]) = \sum_{k=0}^{N} \ell(k, x[k], u[k]) + \ell_N(x[N])$$

- Problems of this form are said to be in Bolza form
  - Lagrange problem if $\ell_f = 0$
  - Mayer problem if $\ell = 0$

- Problem said to be
  - Infinite Horizon if $t_f = \infty$
  - Finite Horizon o/w

Assertion: Every OCP in Bolza Form can be reformulated as an OCP in Mayer Form with the same optimal $U(\cdot)$

Bolza Form:

$$\min_{U(\cdot)} \int_{t_0}^{t_f} \ell(t, x(t), u(t))\, dt + \ell_f(x(t_f))$$

$$\text{s.t. } x(t_0) = x_0$$

Convert running cost into terminal cost:

Trick: New Variable x'

$$x'(t) = \begin{bmatrix} x(t) \\ T(t) \end{bmatrix} = \begin{bmatrix} x(t) \\ \int_{t_0}^{t} \ell(t,x,u)\,dt \end{bmatrix}$$

Total of running cost so far

Alternate Problem:

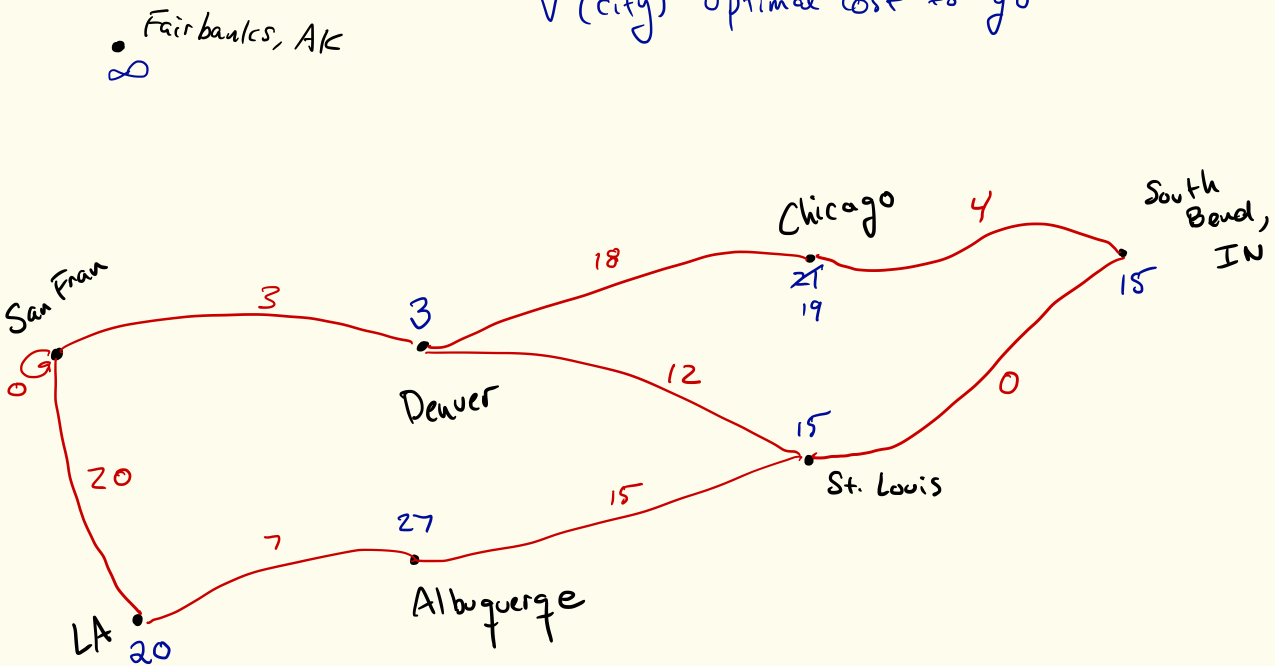
$$\min_{U(\cdot)} \int_{t_0}^{t_f} 0\, dt + \underbrace{T(t_f) + \ell_f(x(t_f))}_{\ell_f'(x'(t_f))}$$

$$\dot{x}' = \begin{bmatrix} f(t,x,u) \\ \ell(t,x,u) \end{bmatrix} \qquad x'(t_0) = \begin{bmatrix} x_0 \\ 0 \end{bmatrix}$$

Same optimal $u^*(\cdot)$!

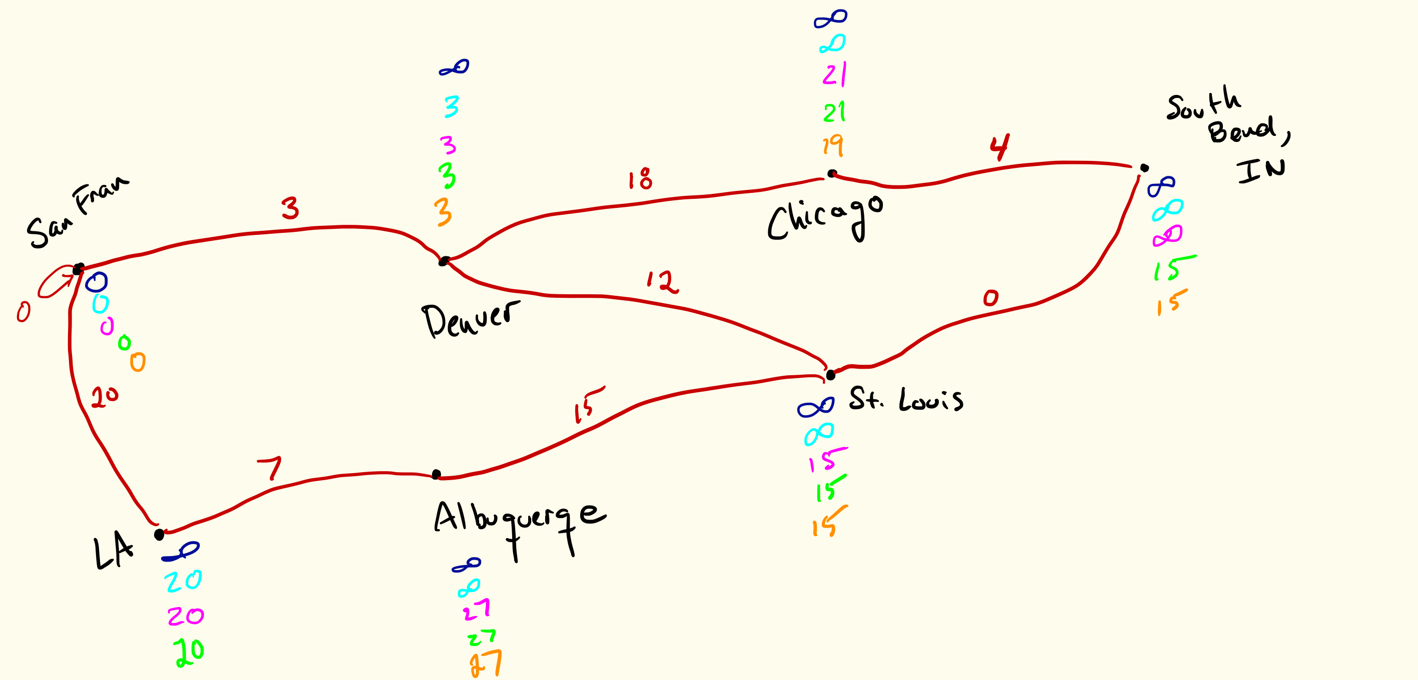
Problem: Road trip. 10 days to the Bay.

$V^*(\text{city})$ optimal cost to go



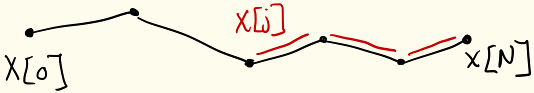
Problem: Road trip. Let $V^*(k, \text{city})$ denote the optimal cost-to-go if you are in "city" on day k (10 days max)

$$V^*(10, \text{city}) \to V^*(9, \text{city}) \to V^*(8, \text{city}) \to V^*(7, \text{city}) \to V^*(6, \text{city})$$



Principle of optimality:

$$V(u[\cdot]) = \sum_{k=0}^{N} \ell(t, x[k], u[k])\, dt + \ell_N(x[N])$$



Given an optimal trajectory

subtrajectories are optimal too

Optimal Cost to Go

$$V^*(k_0, x_0) : \min_{u[\cdot]} V(u[\cdot]) \quad \text{s.t. } x[k_0] = x_0$$

Bellman's Equation

Bellman's equation propagates the optimal cost to go backward in time

$$V^*(k, x) = \min_u \left[\ell(k, x, u) + V^*(k+1, F(k, x, u)) \right]$$

$$V^*(x) = \min_u \left[\ell(x, u) + V^*(F(x, u)) \right]$$

What you pay now

Lowest you could pay after that

∞ horizon time invariant case.

Value iteration Algorithm: Discrete States, discrete Controls, ∞ horizon

$$x[n+1] = f(x[n], u[n]) \qquad V(u[\cdot]) = \sum_{k=1}^{\infty} \ell(x[k], u[k])$$

Goal
- Find V^* iteratively
- Let V^i denote your guess of V^* at iteration i
- Iterate to improve your guess

① Start with $V^0 = 0$ [or some other guess]

② $V^{i+1}(x) = \min_u \left[\ell(x,u) + V^i(f(x,u)) \right]$

③ Repeat to convergence

Challenge #1: Storage

How do you store V?

① On a grid

Grid size $= (\#\text{ divisions})^{\text{State Dimension}}$

"Curse of dimensionality"

② Approximate V w/ some function approximator

$\Rightarrow$ convergence difficult

$\Rightarrow$ This is what most of Reinforcement Learning is doing (Broadly so)

Challenge #2: Convergence

- May not converge
- Iteration procedure may have fixed points that are not V^*

Example: Discrete approx to continuous dynamics

Suppose a continuous time problem

Dynamics
$$\dot{x} = f_c(x,u)$$

Cost
$$\int_0^{\infty} l_c(x,u)\,dt$$

Denotes continuous

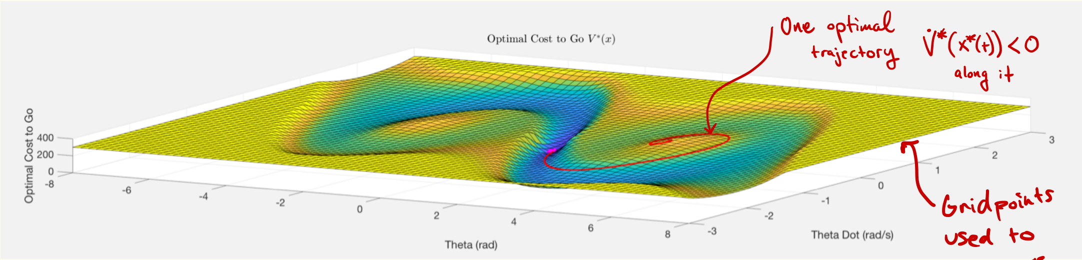
Transition to Discrete $\Rightarrow$ (Approx)

$$x[k+1] = \overbrace{x[k] + f_c(x,u)\Delta t}^{f_d(x,u)}$$

$$l_d(x,u) = l_c(x,u)\Delta t$$

Pendulum case:

$$\frac{d}{dt}\begin{bmatrix}\theta \\ \dot{\theta}\end{bmatrix} = \begin{bmatrix}\dot{\theta} \\ \tau + \sin\theta\end{bmatrix}$$



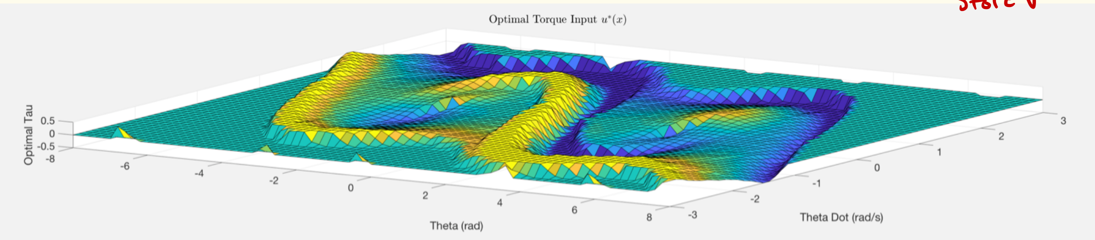
One optimal trajectory

$\dot{V}^*(x^*(t)) < 0$ along it

Gridpoints used to store V^*

$$\min_{u(\cdot)} \int_{t_0}^{\infty} \left[\theta^2(t) + \dot{\theta}^2(t) + 150\,\tau^2(t)\right] dt$$

$$|\tau(t)| \le 0.5\ \text{Nm}$$



Continuous Time Version of Bellman's equation:

$\dot{x} = f(t,x,u)$

State after Δt

$$0 = \lim_{\Delta t \to 0} \min_u \left[\frac{\ell(t,x,u)\Delta t + V^*(t+\Delta t, x + f(t,x,u)\Delta t) - V^*(t,x)}{\Delta t} \right]$$

$$= \lim_{\Delta t \to 0} \min_u \left[\frac{\ell(t,x,u)\cancel{\Delta t} + \cancel{V^*(t,x)} + \frac{\partial V^*}{\partial t}\cancel{\Delta t} + \frac{\partial V^*}{\partial x} f(t,x,u)\cancel{\Delta t} - \cancel{V^*(t,x)}}{\cancel{\Delta t}} \right]$$

V^* must satisfy

$$\boxed{0 = \min_u \left[\ell(t,x,u) + \frac{\partial V^*}{\partial t} + \frac{\partial V^*}{\partial x} f(t,x,u) \right]}$$

Hamilton-Jacobi-Bellman equation (HJB)

Denote $u^*(x)$ as the optimal control policy

Along optimal trajectories:

$$0 = \ell(t,x,u^*(x)) + \frac{\partial V^*}{\partial t} + \frac{\partial V^*}{\partial x} f(t,x,u^*(x))$$

$$-\ell(t,x,u^*(x)) = \dot{V}^*(t,x,u^*(x))$$

Along optimal trajectories, the Rate of Decrease in V^* must be equal to the running cost.

Summary:

- Optimal cost to go = min [what I pay now + what I pay later]

- Bellman Equation (discrete time)

$$V^*(k,x) = \min_u \left[\ell(k,x,u) + V^*(k+1, f(k,x,u)) \right]$$

- HJB (continuous time)

$$0 = \min_u \left[\ell(t,x,u) + \frac{\partial V^*}{\partial t} + \frac{\partial V^*}{\partial x} f(t,x,u) \right]$$