Atlas Backflip

State: Joint Angles / Angular Velocity

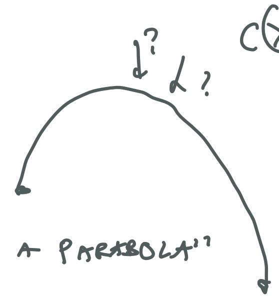
Base Pose

Action: Joint Torques

Transition: Physics

Cost: Penalization of Errors

c(x, u)



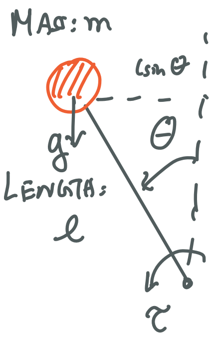
"Track a parabola"

Rotate

GOAL: BUILD A CONTROLLER FOR AN INVERTED PENDULUM

(MDP ROUTE)

DEFINE MDP FOR A PENDULUM



$$\text{TORQUE} = mgl\sin\theta + \tau = I\ddot{\theta}$$

$$= ml^2\ddot{\theta}$$

$$\ddot{\theta} = \frac{g}{l}\sin\theta + \frac{\tau}{ml^2}$$

EULER-LAGRANGE

$$L = \underset{\text{(kinetic energy)}}{T} - \underset{\text{(potential energy)}}{V}$$

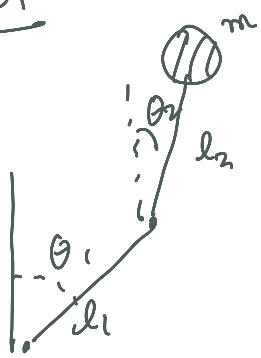
$$= \frac{1}{2} ml^2\dot{\theta}^2 - mgl\cos\theta$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = \tau$$

$$\frac{d}{dt}\left(ml^2\dot{\theta}\right) + mgl\sin\theta = \tau$$

$$\boxed{ml^2\ddot{\theta} + mgl\sin\theta = \tau}$$

ACROBOT



LET'S LINEARIZE ABOUT $(\theta \approx 0, \dot{\theta} \approx 0)$

$$\ddot{\theta} = \frac{g}{l} \sin\theta + \frac{\tau}{ml^2}$$

$$\ddot{\theta} \approx \frac{g}{l} \theta + \frac{\tau}{ml^2} \qquad \text{(Dynamics)}$$

STATE:

$$X_t = \begin{bmatrix} \theta_t \\ \dot{\theta}_t \end{bmatrix}$$

ACTIONS:

$$u_t = \tau$$

DYNAMICS:

$$X_{t+1} = \underbrace{f}_{\text{LINEAR}}(x_t, u_t)$$

$$\begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix}_{t+1} = \underbrace{\begin{bmatrix} 1 + \frac{g}{l}\left(\frac{1}{2}\Delta t^2\right) & \Delta t \\ \frac{g}{l}\Delta t & 1 \end{bmatrix}}_{A} \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix}_t + \frac{1}{ml^2} \underbrace{\begin{bmatrix} \frac{\Delta t^2}{2} \\ \Delta t \end{bmatrix}}_{B} [\tau]_t$$

$$\ddot{\theta} \approx \frac{g}{l} \theta + \frac{\tau}{ml^2}$$

Δt : Timestep

$$X_{t+1} = A x_t + B u_t$$

$$\dot{\theta}_{t+1} = \dot{\theta}_t + \ddot{\theta} \Delta t$$

$$\theta_{t+1} = \theta_t + \dot{\theta}_t \Delta t + \frac{1}{2}\Delta t^2 \ddot{\theta}_t$$

$\theta=0,0$

Cost

"State error" "Control effort" Meaning?

$$C(x,u) = x^T Q x + u^T R u$$

$$= \begin{bmatrix} \theta & \dot{\theta} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix} + \tau^2$$

$$e = \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix} - \begin{bmatrix} \theta_{ref} \\ \dot{\theta}_{ref} \end{bmatrix} = \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix}$$

$$= \theta^2 + \dot{\theta}^2 + \tau^2$$

$$Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix} \qquad Q_2 = \begin{bmatrix} 1 & \to 0 \\ 0 \swarrow & 0 \end{bmatrix}$$

$$Q = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \qquad C(x,u) = -\theta^2 + \dot{\theta}^2$$

Not Convex cost!

Q must be Symmetric Positive Semidefinite $Q \succeq 0$, $x^T Q x \geq 0$

R " " " " Definite $u^T R u > 0$

Value Iteration via Dynamic Programming

$$V_t(x_t) = \min_{u_t} \left[c(x_t, u_t) + V_{t+1}(x_{t+1}) \right]$$

T-1

$$V_{T-1}(x_{T-1}) = \min_{u_{T-1}} \left[c(x_{T-1}, u_{T-1}) + 0 \right]$$

$$= \min_{u_{T-1}} \left[x_{T-1}^T Q x_{T-1} + u_{T-1}^T R u_{T-1} \right]$$

$$\begin{tikzcd} x_{T-1} \arrow[r] & x_T \\ u_{T-1} \arrow[ur] & \end{tikzcd}$$

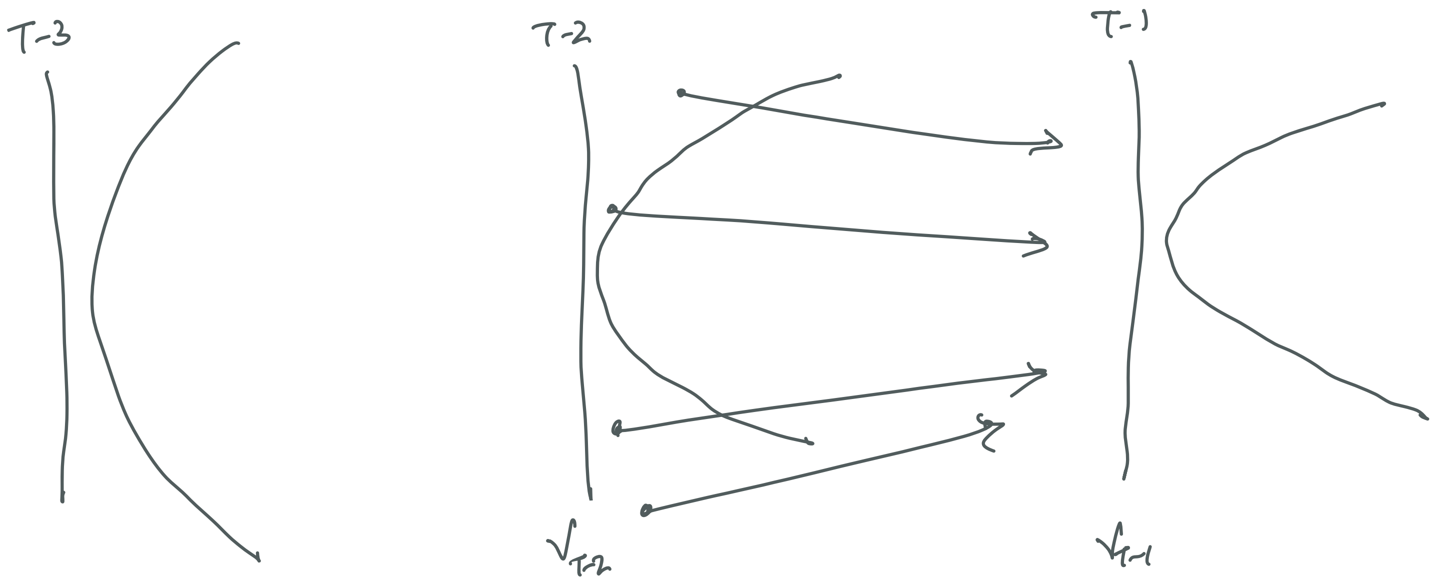
$$= \frac{\partial(\cdot)}{\partial u_{T-1}} = 0 \qquad \boxed{2u_{T-1}^T R = 0}$$

$$u_{T-1} = 0$$

$$V_{T-1}(x_{T-1}) = x_{T-1}^T Q x_{T-1} \qquad \text{"Quadratic"}$$

$$\triangleq x_{T-1}^T V_{T-1} x_{T-1}$$

"The Trick"



Arbitrary timestep t

$$V_t(x_t) = \min_{u_t} \left[c(x_t, u_t) + V_{t+1}(x_{t+1}) \right]$$

$$= \min_{u_t} \left[x_t^T Q x_t + u_t^T R u_t + \underline{\underline{x_{t+1}}}^T V_{t+1} \underline{\underline{x_{t+1}}} \right]$$

Substitute

$$x_{t+1} = A x_t + B u_t$$

$$= \min_{u_t} \left[x_t^T Q x_t + u_t^T R u_t + (A x_t + B u_t)^T V_{t+1} (A x_t + B u_t) \right]$$

$$\frac{\partial (\cdot)}{\partial u_t} = 0$$

$\frac{\partial}{\partial x}(Ax) = A$

$\frac{\partial}{\partial x}(x^T A) = A^T$

$$\cancel{2} u_t^T R + \cancel{2}\left((A x_t + B u_t)^T V_{t+1} B \right) = 0$$

$$R u_t + B^T V_{t+1} (A x_t + B u_t) = 0$$

$$R u_t + B^T V_{t+1} A x_t + B^T V_{t+1} B u_t = 0$$

$$(R + B^T V_{t+1} B) u_t = - B^T V_{t+1} A x_t$$

$$u_t = \underbrace{\boxed{- (R + B^T V_{t+1} B)^{-1} B^T V_{t+1} A}}_{K_t} x_t$$

$$V_t(x_t) = x_t^T \underbrace{\Big(\underbrace{Q}_{\text{one step cost}} + \underbrace{K_t^T R K_t}_{\text{Warped control effort}} + \underbrace{(A+BK_t)^T \underbrace{V_{t+1}}_{\text{Future value}} (A+BK_t)}_{\text{Pull back operator}} \Big)}_{V_t} x_t$$