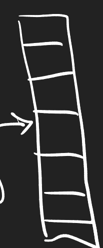
1) State distribution & Transition matrices

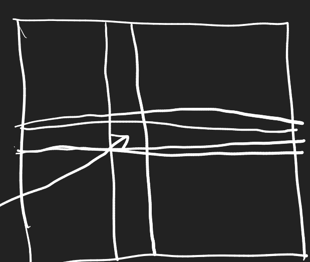
In discrete setting (i.e. $\mathcal{S}$ and $\mathcal{A}$ have finite elements) functions & probability distributions can be represent "tabular form"

State $s_0 \sim \mu_0$  $\in \mathbb{R}^S$ vector representation d_0

$s_0 \in \mathcal{S}$

probability of $s_0 = s$

$d_0 \in \mathbb{R}^S$

recall P^π  $\in \mathbb{R}^{S \times S}$

$P(s'|s, \pi(s))$

$$\left(\mathbb{E}_{s' \sim P(s,\pi(s))}[V^\pi(s')] = \underbrace{\sum_{s' \in \mathcal{S}} P(s'|s,\pi(s))\, V^\pi(s')}_{\langle P^\pi_s,\ V^\pi \rangle} \right)$$

$s_1 \sim P(s_0, \pi(s_0)),\ s_0 \sim \mu_0$

$$d_1 = (P^\pi)^T d_0$$

2) Continuous Control

Historically "optimal control problem"

$$\mathcal{M} = \{\mathcal{S}, \mathcal{A}, (f, \mathcal{D}), c, H, \mu_0\}$$

$\mathcal{S} \subseteq \mathbb{R}^{n_s}$

$\mathcal{A} \subseteq \mathbb{R}^{n_a}$

dynamics

$f : \mathcal{S} \times \mathcal{A} \times \mathcal{W} \to \mathcal{S}$

$s_{t+1} = f(s_t, a_t, w_t)$

$w_t \sim \mathcal{D} = \Delta(\mathcal{W})$

sometimes, $f = (f_0, f_1, \dots f_{H-1})$

cost function (to minimize)

$c : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$

$c = (c_0, c_1, \dots, c_{H-1}, c_H)$

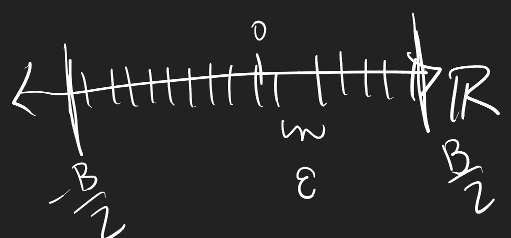
$c_H : \mathcal{S} \to \mathbb{R}$

Optimal Control Problem

$$\min_{\pi} \mathbb{E}\left[\sum_{t=0}^{H-1} c(s_t, a_t) + c_H(s_H) \;\middle|\; \begin{array}{l} s_{t+1} = f(s_t, a_t, w_t) \\ a_t = \pi_t(s_t) \\ s_0 \sim \mu_0 \\ w_t \sim \mathcal{D} \end{array} \right]$$

Discretization?

Exponential dependence on n_s & n_a



Represent functions parametrically

e.g. $f_\theta(x) = \theta^T x$ $\quad \theta \in \mathbb{R}^d$

parameter

3) Linear Dynamics

$$s_{t+1} = A s_t + B a_t + w_t$$

Gaussian

$w_t \sim \mathcal{N}(0, \sigma^2 I)$

$A \in \mathbb{R}^{n_s \times n_s} \quad B \in \mathbb{R}^{n_s \times n_a} \quad w_t \in \mathbb{R}^{n_s}$

ex Robot moves in 1D by applying force left (neg) or right (pos) of any magnitude

" $\underset{a_t}{\underline{\text{force}}} = \underset{m}{\text{mass}} \times \text{acceleration}$" $\text{accel.} = \frac{a_t}{m}$ ← action

$dt=1$ $\text{accel.} = \frac{V_{t+1} - V_t}{dt} = V_{t+1} - V_t$

$$V_{t+1} = 1V_t + \frac{a_t}{m}$$

position $V_t = \frac{P_{t+1} - P_t}{1}$ → $P_{t+1} = P_t + V_t = [1 \quad 1] S_t$

$$S_t = \begin{bmatrix} P_t \\ V_t \end{bmatrix} \in \mathbb{R}^2$$

$$S_{t+1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} S_t + \begin{bmatrix} 0 \\ 1/m \end{bmatrix} a_t$$

Trajectory → $S_{t+1} = AS_t + Ba_t + W_t$

→ $S_t = A^t S_0 + \sum_{k=0}^{t-1} A^k (Ba_{t-k-1} + W_{t-k-1})$

if $a_t = \pi(S_t) = KS_t$ $K \in \mathbb{R}^{n_a \times n_s}$

$$S_{t+1} = (A+BK)S_t + W_t$$

$$S_t = (A+BK)^t S_0 + \sum_{k=0}^{t-1} (A+BK)^k W_{t-k-1}$$

4) Stability

as $t \to \infty$, what does the system do?

For linear systems,

consider $S_{t+1} = AS_t$ $\left(\begin{matrix} a_t = 0 \\ w_t = 0 \end{matrix} \ \forall t \right)$

1) stable: $s_t \to 0$ as $t \to \infty$

2) unstable: $\|s_t\| \to \infty$ as $t \to \infty$

3) marginally unstable: when system is not stable or unstable

<u>Spectral Radius</u>

Define $\rho(A) = \max_{i=1,\dots,n} |\lambda_i(A)|$ largest magnitude eigenvalue

$A \in \mathbb{R}^{n \times n}$

<u>Theorem</u> (linear system stability)

The dynamics $s_{t+1} = A s_t$ are

1) stable if $\rho(A) < 1$
2) unstable if $\rho(A) > 1$
3) marginally unstable if $\rho(A) = 1$

<u>Proof</u>: (preview)

$A = VDV^{-1}$

$\tilde{s} = V^{-1} s$

$\tilde{s}_{t+1} = D \tilde{s}_t$

$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_{n_s} \end{bmatrix}$