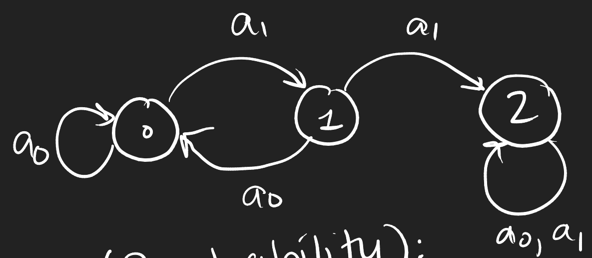
example 3 state 2 action MDP deterministic



$$r(s,a) = \begin{cases} 1 & s=0 \\ 0 & \text{otherwise} \end{cases}$$

Definition (Reachability):

- state s' is <u>reachable</u> from s if $\exists\, a_0, \dots, a_{H-1}$ for finite H such that
$$\mathbb{P}(s_H = s' \mid s_0 = s, a_0, \dots, a_{H-1}) > 0$$
- MDP is <u>reachable</u> if all states are reachable from any initial state

Theorem (Discrete MDP Reachability)

Given $\mathcal{S}, \mathcal{A}, P$ construct a graph with vertices = states and a directed edge $s \to s'$ if $P(s'|s,a) > 0$ $\exists\, a$.
Then MDP is reachable if graph is fully connected

Example: $s \in \mathbb{R}^2$ $a \in \mathbb{R}^1$

$$s_{t+1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} s_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} a_t$$

$$s_t^{(1)} = (1/2)^t s_0^{(1)}$$

Define (Reachable for Deterministic MDP) f dynamics

- s' is reachable from s if $\exists\, a_0, \dots, a_{H-1}$ for finite H such that $s_H = s'$ for $s_{t+1} = f(s_t, a_t)$, $s_0 = s$
- The MDP (dynamics) is reachable if all states are reachable from any state

Theorem (Linear reachability)

A linear system $s_{t+1} = A s_t + B a_t$ is reachable if

$$\text{rank}\big(\underbrace{[B \;\; AB \;\; A^2B \; \cdots \; A^{n_s-1}B]}_{C \in \mathbb{R}^{n_s \times n_a \cdot n_s}}\big) = n_s$$

Proof: Recalling $s_t = A^t s_0 + \underbrace{\sum_{k=0}^{t-1} A^k B a_{t-k-1}}$ (HW1)

$t = n_s$

$$s_{n_s} - A^{n_s} s_0 = \underbrace{\begin{bmatrix} B & AB & \cdots & A^{n_s-1}B \end{bmatrix}} \begin{bmatrix} a_{n_s-1} \\ \vdots \\ a_0 \end{bmatrix}$$

if C is full rank, can solve for $a_0, \dots, a_{n_s-1}$ $\quad a \in \mathbb{R}^{n_a}$

3) Limitations in Observation

Markovian assumption

$$\mathbb{P}(s_{t+1} = s' \mid \underbrace{s_0, \dots, s_t, a_0, \dots, a_t}) = \mathbb{P}(s_{t+1} = s' \mid s_t, a_t)$$

1) Delays

$$= \mathbb{P}(s_{t+1} = s' \mid s_t, a_{t-D})$$

2) Partial observation

$o_t = g(s_t)$ if g invertible ✓

otherwise, g not invertible or noisy, correct approach to plan with

$$\mathbb{P}(s_t = s \mid a_0, \dots, a_{t-1}, o_0, \dots, o_{t-1})$$

- Linear-Gaussian → Kalman filter
- otherwise → particle filtering for example

Another idea: (HW1)

$$s_t \approx \begin{bmatrix} a_t \\ \vdots \\ a_{t-H} \\ o_t \\ \vdots \\ o_{t-H} \end{bmatrix}$$

4) Model Mis-specification & robustness

What if we compute π for a slightly incorrect $P / f, r/c$

Example $\min \mathbb{E}_w\left[\sum_t \|s_t - b a_t\|_2^2 \right.$
$\left. \text{s.t. } s_{t+1} = b a_t + w_t \right]$ $\quad s, a, w \in \mathbb{R}$

optimal $a_t = s_t / b$ under this,

$s_{t+1} = 1 \cdot s_t + w_t$ marginally stable

what if the dynamics & cost actually $\tilde{b}$

$$s_{t+1} = \left(\frac{\tilde{b}}{b}\right) s_t + w_t \qquad c_t = \|s_t - \frac{\tilde{b}}{b} s_t\|_2^2 = (1 - \tilde{b}/b)^2 \|s_t\|_2^2$$

$\tilde{b} > b \rightarrow$ unstable, infinite cumulative cost

Arbitrarily small errors in b lead to arbitrarily bad performance.