10\20\08

Markov Decision Proceses

A process produces an infinite sequence of costs

P1: 6 6 6 6 6 6 6 6

P2: 5 5 5 5 5 5 5 5

P2 has a smaller average cost per period

P3 4 6 4 6 4 6 4 6 4 6

Disount factor: $\alpha < 1$

$$C_0 \quad C_1 \quad C_2 \quad \cdots \quad C_n \quad \cdots$$

$$NPV = C_0 + \alpha C_1 + \alpha^2 C_2 + \alpha^3 C_3 + \cdots$$

Net Present Value

Dynamic Programming over ∞ horizon.

Model

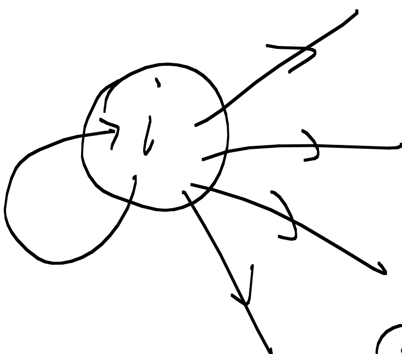
n states $1, 2, \ldots, n$

c_{ij} = cost of goi

Discount factor α

$i = 1, 2, \ldots, n$

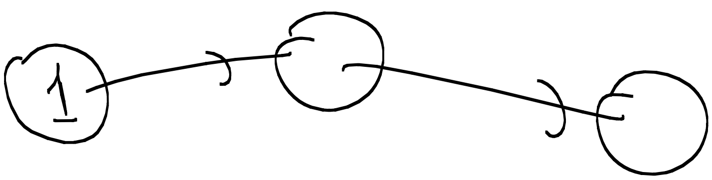
generate a sequence of costs



$1 \longrightarrow 3 \longrightarrow 9 \longrightarrow 6$ ——

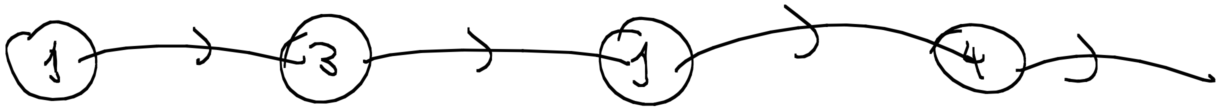
$c_{1,3} + \alpha c_{3,9} + \alpha^2 c_{9,6} + \cdots$

Goal: Minimise cost.



Aim to find best sequence of moves from 1.

∞ # of ways of proceeding.



Non-optimal to go $1 \to 3$
& $1 \to 4$ in same sequence

A *policy* $\pi : V \longrightarrow V$

When at i, go to $\pi(i)$

Problem: find optimal π.

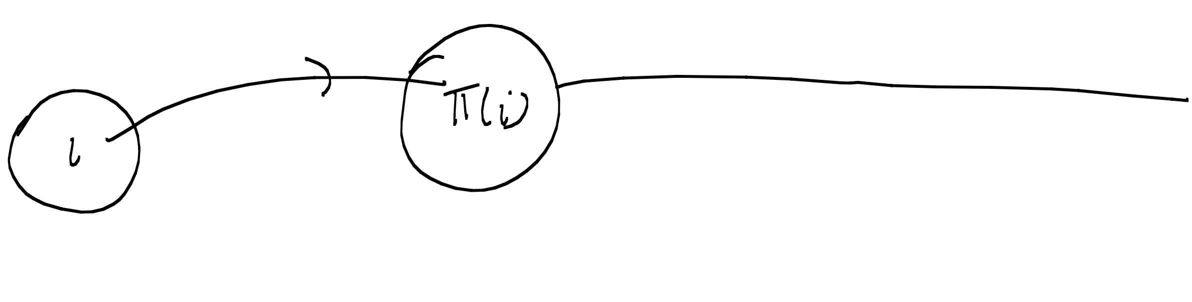
Fix π

$$\begin{bmatrix} 3 & 2 & 1 & 0 \\ -3 & 4 & 2 & -1 \\ 3 & 6 & 5 & 4 \\ 2 & 1 & 2 & 1 \end{bmatrix}$$

y_i = cost starting at i

Evaluate π by solving

$$y_i = c_{i,\pi(i)} + \alpha\, y_{\pi(i)}, \quad i \in V \qquad \alpha = \frac{1}{2}$$



$$\begin{bmatrix} 3 & 2 & 1 & 0 \\ -3 & 4 & 2 & -1 \\ 3 & 6 & 5 & 4 \\ 2 & 1 & 2 & 1 \end{bmatrix}$$

$\alpha = \frac{1}{2}$

i	1	2	3	4
$\pi(i)$	1	3	4	1

$$y_1 = 3 + \tfrac{1}{2} y_1 \qquad y_1 = 6$$

$$y_2 = 2 + \tfrac{1}{2} y_3 \qquad y_2 = 21/4$$

$$y_3 = 4 + \tfrac{1}{2} y_4 \qquad y_3 = 13/2$$

$$y_4 = 2 + \tfrac{1}{2} y_1 \qquad y_4 = 5$$

Goal: Find π^* such that ??

$$y_i^* \leq y_i \quad \forall\, i, \pi$$

Values using π^*