9\16\09

$$i \xrightarrow{c_{ij}} j$$

$V = \{\text{states}\}$

$\alpha =$ discount factor

Policy: $\pi : V \longrightarrow V$

when in i, go to $\pi(i)$.

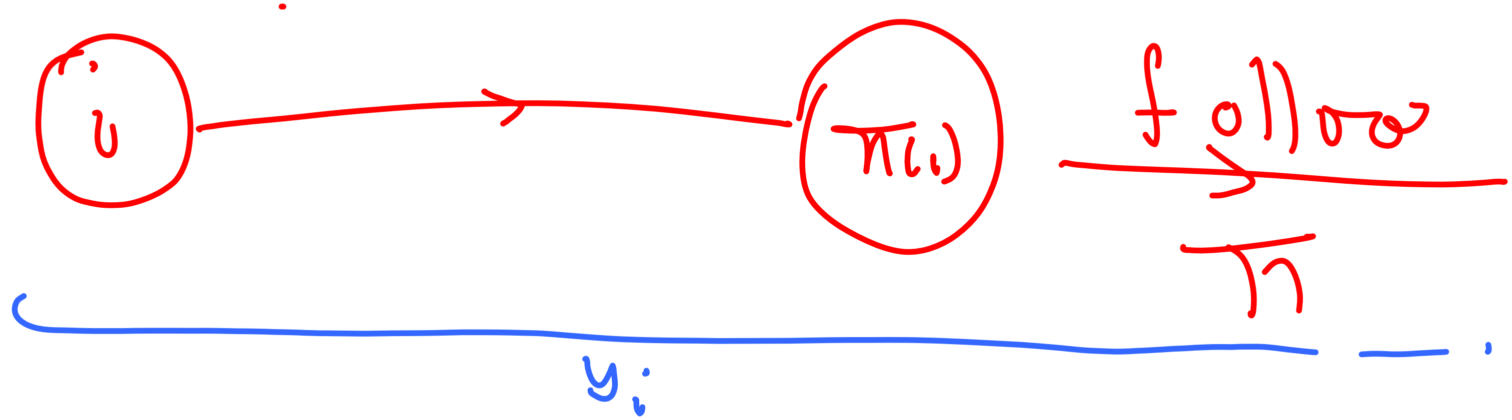
Policy Evaluation:

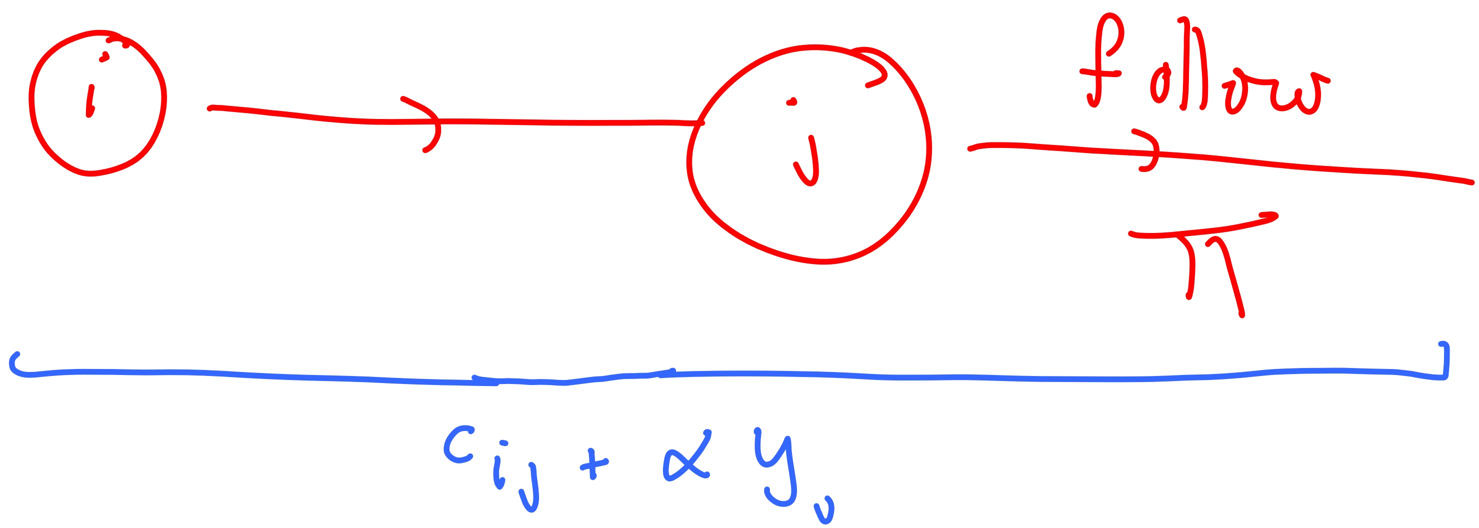
y_i = discounted cost of π starting at i

$$= c_{i,\pi(i)} + \alpha\, y_{\pi(i)}$$

Problem: find π that simultaneously minimize $y_i : i \in V$

Policy Improvement.





Compare

y_i with $\underbrace{c_{ij} + \alpha y_j, \quad \forall j}_{}$

if min. $< y_i$

switch $\pi(i)$ to min.

Compute

$$M_i = c_{i,\mu(i)} + \alpha y_{\mu(i)} = \min_j \left[c_{i,j} + \alpha y_j \right]$$

if $M_i < y_i$, switch from $\pi(i)$ to $\lambda(i)$

Example

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

$$\begin{bmatrix} 3 & 7 & 4 & 2 \\ 1 & 5 & 6 & 3 \\ 2 & 8 & 4 & 7 \\ 5 & 9 & 3 & 1 \end{bmatrix}$$

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

$$y_1 = 7 + \tfrac{1}{2} y_2 = 41/3$$
$$y_2 = 6 + \tfrac{1}{2} y_3 = 40/3$$
$$y_3 = 8 + \tfrac{1}{2} y_2 = 44/3$$
$$y_4 = 9 + \tfrac{1}{2} y_2 = 47/3$$

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

$$\begin{bmatrix} 3 & 7 & 4 & 2 \\ 1 & 5 & 6 & 3 \\ 2 & 8 & 4 & 7 \\ 5 & 9 & 3 & 1 \end{bmatrix}$$

$y_1 = 41/3$

$y_2 = 40/3$

$y_3 = 44/3$

$y_4 = 47/3$

Policy Improvement:

$i = 1$:

$3 + \frac{1}{2} y_1 = 59/6$ *

$7 + \frac{1}{2} y_2 = 41/3$

$4 + \frac{1}{2} y_3 = 34/3$

$2 + \frac{1}{2} y_4 = 59/6$ *

$i = 2$:

$1 + \frac{1}{2} y_1 = 47/6$ *

$5 + \frac{1}{2} y_2 = 35/3$

$6 + \frac{1}{2} y_3 = 40/3$

$3 + \frac{1}{2} y_4 = 65/6$

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

$$\begin{bmatrix} 3 & 7 & 4 & 2 \\ 1 & 5 & 6 & 3 \\ 2 & 8 & 4 & 7 \\ 5 & 9 & 3 & 1 \end{bmatrix}$$

$y_1 = 41/3$ $\quad y_3 = 44/3$

$y_2 = 40/3$ $\quad y_4 = 47/3$

Policy Improvement:

$i = 3$:

$$2 + \tfrac{1}{2} y_1 = 53/6 \quad *$$
$$8 + \tfrac{1}{2} y_2 = 44/3$$
$$4 + \tfrac{1}{2} y_3 = 34/3$$
$$7 + \tfrac{1}{2} y_4 = 89/6$$

$i = 4$:

$$5 + \tfrac{1}{2} y_1 = 71/6$$
$$9 + \tfrac{1}{2} y_2 = 47/3$$
$$3 + \tfrac{1}{2} y_3 = 31/3$$
$$1 + \tfrac{1}{2} y_4 = 53/6 \quad *$$

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

$$\begin{bmatrix} 3 & 7 & 4 & 2 \\ 1 & 5 & 6 & 3 \\ 2 & 8 & 4 & 7 \\ 5 & 9 & 3 & 1 \end{bmatrix}$$

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

$$y_1 = 3 + \frac{1}{2} y_1 = 6$$

$$y_2 = 1 + \frac{1}{2} y_1 = 4$$

$$y_3 = 2 + \frac{1}{2} y_1 = 5$$

$$y_4 = 1 + \frac{1}{2} y_4 = 2$$

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

$$\begin{bmatrix} 3 & 7 & 4 & 2 \\ 1 & 5 & 6 & 3 \\ 2 & 8 & 4 & 7 \\ 5 & 9 & 3 & 1 \end{bmatrix}$$

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

$y_1 = 6, \quad y_2 = 4, \quad y_3 = 5, \quad y_4 = 2$

$i = 1$

$$3 + \tfrac{1}{2} y_1 = 6$$
$$7 + \tfrac{1}{2} y_2 = 9$$
$$4 + \tfrac{1}{2} y_3 = 13/2$$
$$2 + \tfrac{1}{2} y_4 = 3 \quad *$$

$i = 2$

$$1 + \tfrac{1}{2} y_1 = 4 \quad *$$
$$5 + \tfrac{1}{2} y_2 = 7$$
$$6 + \tfrac{1}{2} y_3 = 17/2$$
$$3 + \tfrac{1}{2} y_4 = 4 \quad *$$

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

$$\begin{bmatrix} 3 & 7 & 4 & 2 \\ 1 & 5 & 6 & 3 \\ 2 & 8 & 4 & 7 \\ 5 & 9 & 3 & 1 \end{bmatrix}$$

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

$y_1 = 6, \quad y_2 = 4, \quad y_3 = 5, \quad y_4 = 2$

$i = 3$

$$2 + \tfrac{1}{2} y_1 = 5 \quad *$$
$$8 + \tfrac{1}{2} y_2 = 10$$
$$4 + \tfrac{1}{2} y_3 = 13/2$$
$$7 + \tfrac{1}{2} y_4 = 8$$

$i = 4$

$$5 + \tfrac{1}{2} y_1 = 8$$
$$9 + \tfrac{1}{2} y_2 = 11$$
$$3 + \tfrac{1}{2} y_3 = 11/2$$
$$1 + \tfrac{1}{2} y_4 = 2 \quad *$$

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

$$\begin{bmatrix} 3 & 7 & 4 & 2 \\ 1 & 5 & 6 & 3 \\ 2 & 8 & 4 & 7 \\ 5 & 9 & 3 & 1 \end{bmatrix}$$

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

$$y_1 = 2 + \frac{1}{2} y_4 = 3$$

$$y_2 = 1 + \frac{1}{2} y_1 = 5/2$$

$$y_3 = 2 + \frac{1}{2} y_1 = 7/2$$

$$y_4 = 1 + \frac{1}{2} y_4 = 2$$