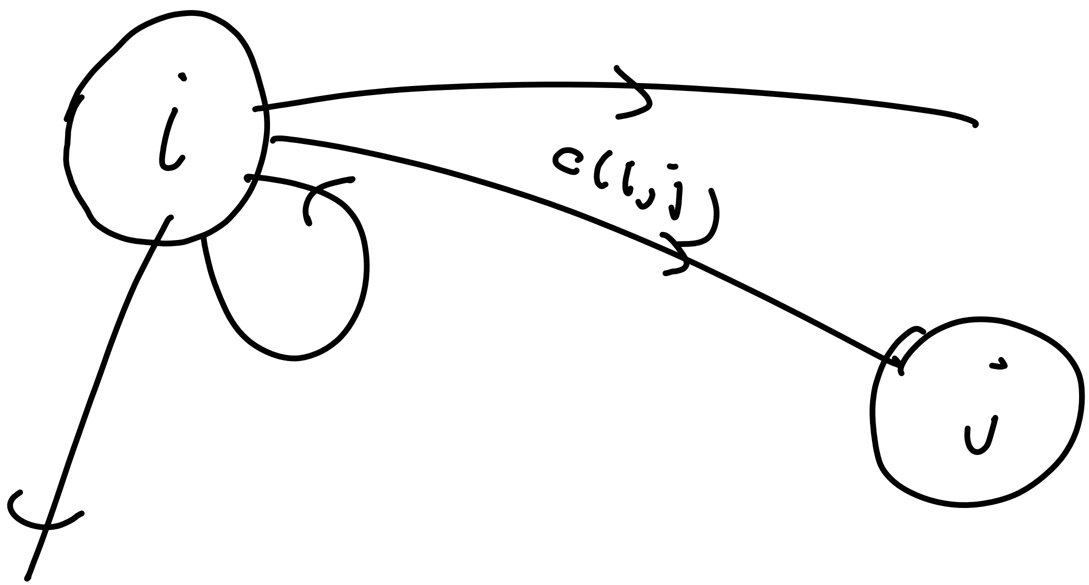
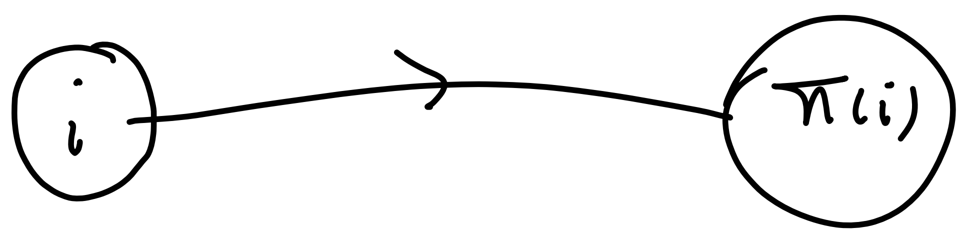
9\16\11

α = discount factor.



A policy $\pi : V \to V$



Policy Evaluation

y_i = N.P.V. of follow π, starting at i

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

cost of period 1 — discounted cost of rest of sequence

Solve the equations to get the y's.

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

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

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

$$y_1 = 5 + \frac{1}{2} y_2 = 72/7$$

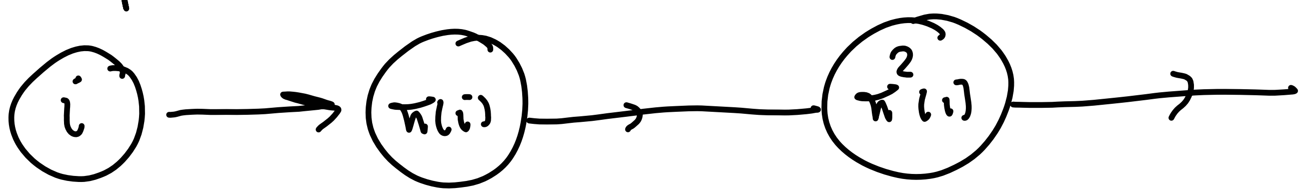
$$y_2 = 5 + \frac{1}{2} y_3 = 74/7$$

$$y_3 = 6 + \frac{1}{2} y_1 = 78/7$$

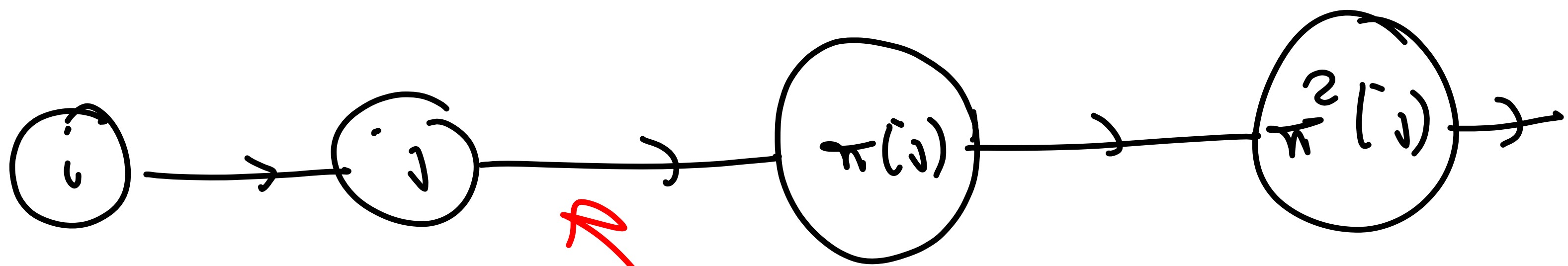
$$y_4 = 7 + \frac{1}{2} y_4 = 14$$

Optimality Test:

Compare:



with



cost of ↑ $= c(i,j) + \alpha y_j$

Let $I = \{ i : y_i = \min_j [c(i,j) + \alpha y_j] \}$

$y_i \geq$ smallest of these always

If $I = V$, optimal; otherwise make $\pi(i) =$ minimiser, $i \notin I$

$$c(1,1) + \alpha y_1 = 3 + \tfrac{1}{2} \cdot \tfrac{72}{7} = \tfrac{57}{7}$$
$$c(1,2) + \alpha y_2 = 5 + \tfrac{1}{2} \tfrac{74}{7} = \tfrac{72}{7}$$
$$c(1,3) + \alpha y_3 = 2 + \tfrac{1}{2} \tfrac{78}{7} = \tfrac{53}{7} \quad *$$
$$c(1,4) + \alpha y_4 = 1 + \tfrac{1}{2} \cdot 14 = 8$$

$$c(2,1) + \alpha y_1 = 4 + \tfrac{1}{2} \tfrac{72}{7} = \tfrac{64}{7}$$
$$c(2,2) + \alpha y_2 = 2 + \tfrac{1}{2} \cdot \tfrac{74}{7} = \tfrac{51}{7} \quad *$$
$$c(2,3) + \alpha y_3 = 5 + \tfrac{1}{2} \tfrac{78}{7} = \tfrac{74}{7}$$
$$c(2,4) + \alpha y_4 = 3 + \tfrac{1}{2} \cdot 14 = 10$$

$$c(3,1) + \alpha y_1 = 6 + \tfrac{1}{2} \tfrac{72}{7} = \tfrac{78}{7}$$
$$c(3,2) + \alpha y_2 = 1 + \tfrac{1}{2} \tfrac{74}{7} = \tfrac{44}{7} \quad *$$
$$c(3,3) + \alpha y_3 = 3 + \tfrac{1}{2} \tfrac{78}{7} = \tfrac{60}{7}$$
$$c(3,4) + \alpha y_4 = 4 + \tfrac{1}{2} \cdot 14 = 11$$

$$c(4,1) + \alpha y_1 = 1 + \tfrac{1}{2} \tfrac{72}{7} = \tfrac{43}{7} \rightarrow *$$
$$c(4,2) + \alpha y_2 = 3 + \tfrac{1}{2} \tfrac{74}{7} = \tfrac{58}{7}$$
$$c(4,3) + \alpha y_3 = 5 + \tfrac{1}{2} \tfrac{78}{7} = \tfrac{74}{7}$$
$$c(4,4) + \alpha y_4 = 7 + \tfrac{1}{2} 14 = 14$$

NEW POLICY

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

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

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

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

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

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

$$y_3 = 1 + \frac{1}{2} y_2 = 3$$

$$y_4 = 1 + \frac{1}{2} y_1 = 11/4$$

$$C(1,1) + \alpha y_1 = 3 + \frac{1}{2} \cdot \frac{7}{2} = \frac{19}{4}$$
$$C(1,2) + \alpha y_2 = 5 + \frac{1}{2} \cdot 4 = 7$$
$$C(1,3) + \alpha y_3 = 2 + \frac{1}{2} \cdot 3 = 7/2$$
$$C(1,4) + \alpha y_4 = 1 + \frac{1}{2} \cdot \frac{11}{4} = \frac{19}{8} \; *$$

$$C(2,1) + \alpha y_1 = 4 + \frac{1}{2} \cdot \frac{7}{2} = \frac{23}{4}$$
$$C(2,2) + \alpha y_2 = 2 + \frac{1}{2} \cdot 4 = 4 \; *$$
$$C(2,3) + \alpha y_3 = 5 + \frac{1}{2} \cdot 3 = \frac{13}{2}$$
$$C(2,4) + \alpha y_4 = 3 + \frac{1}{2} \cdot \frac{11}{4} = \frac{35}{8}$$

$$C(3,1) + \alpha y_1 = 6 + \frac{1}{2} \cdot 7/2 = 31/4$$
$$C(3,2) + \alpha y_2 = 1 + \frac{1}{2} \cdot 4 = 3 \; *$$
$$C(3,3) + \alpha y_3 = 3 + \frac{1}{2} \cdot 3 = 9/2$$
$$C(3,4) + \alpha y_4 = 24 + \frac{1}{2} \cdot \frac{11}{4} = 43/8$$

$$C(4,1) + \alpha y_1 = 1 + \frac{1}{2} \cdot \frac{7}{2} = 11/4 \; *$$
$$C(4,2) + \alpha y_2 = 3 + \frac{1}{2} \cdot 4 = 5$$
$$C(4,3) + \alpha y_3 = 5 + \frac{1}{2} \cdot 3 = 13/2$$
$$C(4,4) + \alpha y_4 = 7 + \frac{1}{2} \cdot \frac{11}{4} = 67/8$$

NEW POLICY

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

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

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

$y_1 = 1 + \frac{1}{2}(1 + \frac{1}{2}$

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

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

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

$$y_3 = 1 + \frac{1}{2} y_2 = 3$$

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

$C(1,1) + \alpha y_1 = 3 + \frac{1}{2} \cdot 2 = 4$

$C(1,2) + \alpha y_2 = 5 + \frac{1}{2} \cdot 4 = 7$

$C(1,3) + \alpha y_3 = 2 + \frac{1}{2} \cdot 3 = 7/2$

$C(1,4) + \alpha y_4 = 1 + \frac{1}{2} \cdot 2 = 2$ *

$C(2,1) + \alpha y_1 = 4 + \frac{1}{2} \cdot 2 = 5$

$C(2,2) + \alpha y_2 = 2 + \frac{1}{2} \cdot 4 = 4$ *

$C(2,3) + \alpha y_3 = 5 + \frac{1}{2} \cdot 3 = 13/2$

$C(2,4) + \alpha y_4 = 3 + \frac{1}{2} \cdot 2 = 4$

$C(3,1) + \alpha y_1 = 6 + \frac{1}{2} \cdot 2 = 7$

$C(3,2) + \alpha y_2 = 1 + \frac{1}{2} \cdot 4 = 3$ *

$C(3,3) + \alpha y_3 = 3 + \frac{1}{2} \cdot 3 = 9/2$

$C(3,4) + \alpha y_4 = 4 + \frac{1}{2} \cdot 2 = 5$

$C(4,1) + \alpha y_1 = 1 + \frac{1}{2} \cdot 2 = 2$ *

$C(4,2) + \alpha y_2 = 3 + \frac{1}{2} \cdot 4 = 5$

$C(4,3) + \alpha y_3 = 5 + \frac{1}{2} \cdot 3 = 11/2$

$C(4,4) + \alpha y_4 = 7 + \frac{1}{2} \cdot 2 = 8$

CURRENT POLICY IS OPTIMAL