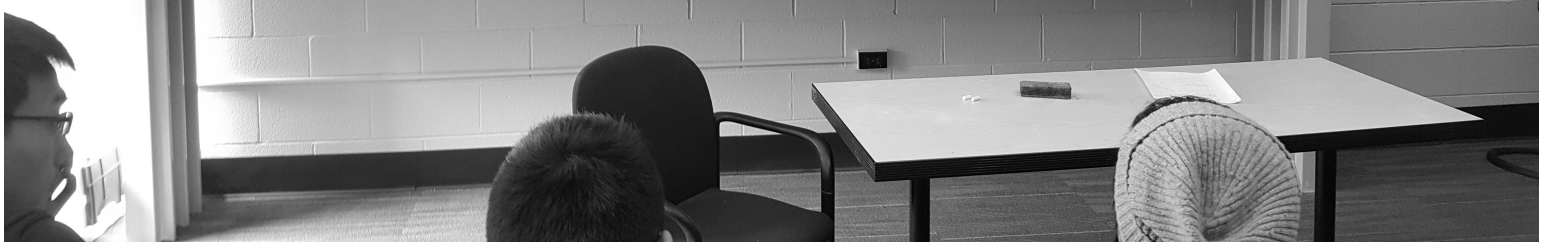
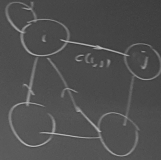


System	Policy Evaluation	Optimality Criterion
	$V = \{\text{States}\}$ $\pi: V \rightarrow V$ is a policy $y_i = c(i, \pi(i)) + \gamma y_{\pi(i)}$	π is optimal iff its values y_i^* satisfy $y_i^* = \min_j (c(i,j) + \gamma y_j^*) \quad \forall i \in V$ (*)

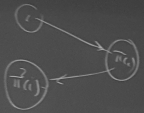


System



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

$\pi: V \rightarrow V$
is a policy



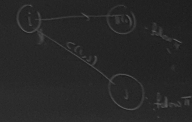
Policy Evaluation

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

Optimality Criterion

π is optimal iff its values y_i^* satisfy

$y_i^* = \min_j c(i,j) + \alpha y_j^*, \forall i \in V$



given y_i
cost is $c(i,j) + \alpha y_j$
if $< y_i$ then π can be improved

\Rightarrow (*) is necessary

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

$C(i,j)$	↓				
5	3	4	6		
2	1	7	2		
3	4	8	3		
1	4	3	7		
	1	2	3	4	
π	4	3	3	4	

Policy Evaluation

$y_1 = 6 + \frac{1}{2} y_4 = 13$
 $y_2 = 7 + \frac{1}{2} y_2 = 15$
 $y_3 = 8 + \frac{1}{2} y_3 = 16$
 $y_4 = 7 + \frac{1}{2} y_4 = 14$

Matrix Solver

$$\begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix}$$

Check for optimality

min $5 + \frac{1}{2} y_1 = 11.5$
 $7 + \frac{1}{2} y_2 = 10.5$
 $4 + \frac{1}{2} y_3 = 12$
 $6 + \frac{1}{2} y_4 = 13$

New Policy

$\hat{\pi}(i)$	1	2	3	4
	2			

System $\Pi: V \rightarrow V$ to a policy

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

Π is optimal iff its values y_i^* satisfy

$y_i^* = \min_j C(i, j) + \alpha y_j^*, \forall i \in V$ (*)

\Rightarrow (*) is necessary

gives y_i

est. is $C(i, j) + \alpha y_j$

If $< y_i$ then Π can be improved

Policy Evaluation

Matrix Solving

Check for optimality

New Policy

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

$C(i, j)$

	1	2	3	4
1	5	3	4	6
2	2	1	7	2
3	3	4	8	3
4	1	4	3	7

$y_1 = 6 + \frac{1}{2} y_4 = 13$

$y_2 = 7 + \frac{1}{2} y_2 = 15$

$y_3 = 8 + \frac{1}{2} y_3 = 16$

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

Matrix Solving

$$\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} -6 \\ -7 \\ -8 \\ -7 \end{bmatrix}$$

Check for optimality

min

$2 + \frac{1}{2} y_1 = 8 \frac{1}{2} *$

$1 + \frac{1}{2} y_2 = 8 \frac{1}{2} *$

$7 + \frac{1}{2} y_3 = 15$

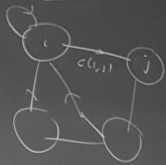
$1 + \frac{1}{2} y_4 = 8$

New Policy

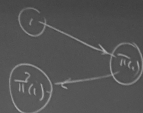
$\hat{\pi}(i) \geq 1$



System



$V = \{ \text{states} \}$
 $\pi: V \rightarrow V$
 is a policy



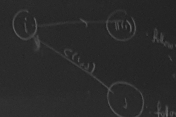
Policy Evaluation

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

Optimality Criteria

π is optimal iff its values y_i^* satisfy

$y_i^* = \min_j c(i, j) + \gamma y_j^*, \forall i \in V$



gives us y_i
 cost is $c(i, j) + \gamma y_j$
 if $< y_i$ then π can be improved

\Rightarrow (*) is necessary

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

$C(i,j)$	5	3	4	6
	2	1	7	2
	3	4	8	3
	1	4	3	7

i	1	2	3	4
π	4	3	3	4

Policy Evaluation

$y_1 = 6 + \frac{1}{2} y_4 = 13$
 $y_2 = 7 + \frac{1}{2} y_2 = 15$
 $y_3 = 8 + \frac{1}{2} y_2 = 16$
 $y_4 = 7 + \frac{1}{2} y_4 = 14$

Matrix Solving

$$\begin{bmatrix} 1 & & & & -2 \\ & 1 - \frac{1}{2} & & & \\ & & 1 - \frac{1}{2} & & \\ & & & 1 - \frac{1}{2} & \\ & & & & 1 \end{bmatrix}$$

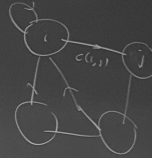
Check for optimality

min $\begin{cases} 5 + \frac{1}{2} y_4 = 9\frac{1}{2} \\ 7 + \frac{1}{2} y_2 = 11\frac{1}{2} \\ 8 + \frac{1}{2} y_2 = 16 \\ 7 + \frac{1}{2} y_4 = 10 \end{cases}$

New Policy

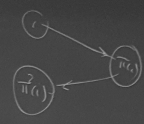
i	1	2	3	4
$\pi(i)$	2	1	1	

System



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

$\pi: V \rightarrow V$
is a policy



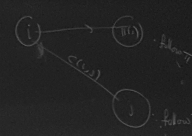
Policy Evaluation

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

Optimality Criterion

π^* is optimal iff its values y^* satisfy

$y_i^* = \min_j C(i, j) + \alpha y_j^*, \forall i \in V$



gives us y_i
cost is $C(i, \pi(i)) + \alpha y_{\pi(i)}$
If $< y_i$ then π can be improved

\Rightarrow (*) is necessary

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

	$C(i, j) \downarrow$			
	5	3	4	6
	2	1	7	2
	3	4	8	3
	1	4	3	7
i	1	2	3	4
π	4	3	3	4

Policy Evaluation

$y_1 = 6 + \frac{1}{2} y_4 = 13$
 $y_2 = 7 + \frac{1}{2} y_3 = 15$
 $y_3 = 8 + \frac{1}{2} y_2 = 16$
 $y_4 = 7 + \frac{1}{2} y_1 = 14$

Matrix Solution

$$\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 6 \\ 7 \\ 8 \\ 7 \end{bmatrix}$$

Check for optimality

$1 + \frac{1}{2} y_1 = 7.5 < 5$
 $4 + \frac{1}{2} y_2 = 12 < 2$
 $3 + \frac{1}{2} y_3 = 11 < 4$
 $7 + \frac{1}{2} y_4 = 14 > 7$

New Policy

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

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

C(u) ↓			
5	3	4	6
2	1	7	2
3	4	8	3
1	4	3	7

Policy Evaluation Matrix Solving Check for optimality New Policy

$y_1 = 6 + \frac{1}{2}y_2 = 13$
 $y_2 = 7 + \frac{1}{2}y_1 = 15$
 $y_3 = 8 + \frac{1}{2}y_2 = 16$
 $y_4 = 7 + \frac{1}{2}y_1 = 14$

$\begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}$

max $\begin{cases} 1 + \frac{1}{2}y_1 = 7.5 \\ 4 + \frac{1}{2}y_2 = 11.5 \\ 3 + \frac{1}{2}y_3 = 11 \\ 7 + \frac{1}{2}y_4 = 14 \end{cases}$

$\hat{\pi}(u) = \begin{matrix} & 1 & 2 & 3 & 4 \\ 1 & 2 & 1 & 1 & 1 \end{matrix}$

Evaluate new policy Check for optimality Verification of optimality criterion

$y_1 = 3 + \frac{1}{2}y_2 = \frac{16}{3}$
 $y_2 = 2 + \frac{1}{2}y_1 = \frac{14}{3}$
 $y_3 = 3 + \frac{1}{2}y_1 = \frac{17}{3}$
 $y_4 = 1 + \frac{1}{2}y_1 = \frac{11}{3}$

Suppose π does not satisfy the criterion
 $y_u > c(u, X(u)) + \alpha y_u$ $u \in U \neq \hat{U}$
 $y_u = \max_v [c(u, v) + \alpha y_v]$ $u \notin \hat{U}$
 New policy $\hat{\pi}(u) = \begin{cases} X(u), u \in \hat{U} \\ \pi(u), u \notin \hat{U} \end{cases}$

$u \in \hat{U} \Rightarrow y_u > c(u, X(u)) + \alpha y_u \Rightarrow \sum_{u \in \hat{U}} y_u > \sum_{u \in \hat{U}} [c(u, X(u)) + \alpha y_u]$
 $\sum_{u \in \hat{U}} y_u > \sum_{u \in \hat{U}} c(u, X(u)) + \alpha \sum_{u \in \hat{U}} y_u$
 $\sum_{u \in \hat{U}} (1 - \alpha) y_u > \sum_{u \in \hat{U}} c(u, X(u))$
 $\sum_{u \in \hat{U}} y_u > \frac{1}{1 - \alpha} \sum_{u \in \hat{U}} c(u, X(u))$
 $\Rightarrow \sum_{u \in \hat{U}} y_u > 0, u \in \hat{U} \Rightarrow \sum_{u \in \hat{U}} y_u > 0, u \notin \hat{U}$

$y_1 = 3 + \frac{1}{2}y_2 = \frac{15}{2}$
 $y_2 = 2 + \frac{1}{2}y_1 = \frac{10}{3}$
 $y_3 = 3 + \frac{1}{2}y_4 = \frac{11}{2}$
 $y_4 = 1 + \frac{1}{2}y_1 = \frac{11}{3}$

Criterion
 ① Suppose π does not satisfy the criterion
 $y_u > c(u, \lambda(u)) + \alpha y_{\lambda(u)} \quad u \in U \neq \emptyset$
 $y_u = \max_{v \in U} [c(u, v) + \alpha y_v]$
 Need policy $\hat{\pi}(u) = \begin{cases} \lambda(u), u \in U \\ \pi(u), u \notin U \end{cases}$

Value of $y_u = c(u, \lambda(u)) + \alpha y_{\lambda(u)}$
 $y_u = y_{\lambda(u)} \Rightarrow y_u \geq \alpha y_{\lambda(u)} \Rightarrow y_u \geq \alpha^2 y_{\lambda(\lambda(u))} \geq \dots$
 $y_u \geq 0, u \in U \Rightarrow y_u \geq 0, u \notin U$

② Suppose π satisfies optimality criterion
 Take any other policy $\hat{\pi}$
 $\hat{y}_u = c(u, \hat{\pi}(u)) + \alpha y_{\hat{\pi}(u)} \leq 16$
 $y_u \leq c(u, \hat{\pi}(u)) + \alpha y_{\hat{\pi}(u)} \leq y_u + 4$
 $\Rightarrow \hat{y}_u = y_u \geq 0$

Check for optimality
 Need policy

