<u>10\22\08</u>

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

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

α = discount factor.

Policy: $\pi : V \to V$

When in i, go next to $\pi(i)$, indefinitely.

y_i = discounted cost starting at i

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

Solve these equations to get the y_i:

$$A = \begin{array}{cc} & \pi(i) \\ i & \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & -\alpha \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix} \end{array}$$

NON-SINGULAR

Want to find optimal $\bar{\pi}$: $\bar{\pi}^*$, y^*

$$y^*_i \leq y_i \quad \forall i, \forall \bar{\pi}$$

$y^*_{i,n}$ = min cost, starting at i, making n moves

$$= \min_j \{ C_{ij} + \alpha y^*_{j,n-1} \}$$

Let $n \to \infty$

$$y^*_i = \min_j \{ C_{ij} + \alpha y^*_j \} \quad **$$

Π^* is optimal $\iff$ (**) holds.

(1) Suppose

$$y_i^* = \min_j \left[c_{ij} + \alpha y_j^* \right] \quad (**)$$

Let Π be any policy with corresponding y.

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

$$y_i^* \leq c_{i,\pi(i)} + \alpha y^*_{i,\pi(i)}$$

$$\underbrace{y_i - y_i^*}_{z_i} \geq \alpha \left[y_{i,\pi(i)} - y^*_{i,\pi(i)} \right]$$

$$
\begin{aligned}
Z_i &\geq \alpha Z_{i,\pi(i)} \\
&\geq \alpha^2 Z_{i,\pi^2(i)} \\
&\;\vdots \\
&\geq \alpha^k Z_{i,\pi^k(i)} \to 0
\end{aligned}
$$

$$\Rightarrow \quad Z_i \geq 0$$