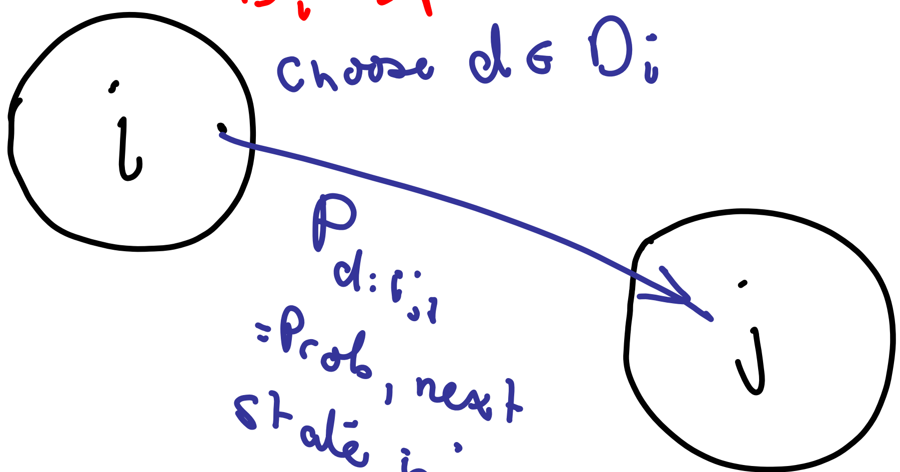
9/17/10

States

$D_i = \{\text{possible decisions}\}$

choose $d \in D_i$



$P_{d:i,j}$ = Prob, next state is j

$c_{d,i}$ = expected cost of the transition.

Goal: find decisions that minimize expected discounted cost from each state

Production problem: with random demand

state i = current inventory

decision x = amount to make

(expected) cost & next state is random

A policy π defines a decision $\pi(i) \in D_i$ for all i. — Defines a Markov chain with transition matrix P_π say.

$y_i = y_i(\pi) =$ expected discounted cost, starting at i

$$= c_{\pi(i),i} + \alpha \sum_j P_{\pi(i):i,j}\, y_j$$

first transition

$$\underline{y} = \underline{c}_\pi + \alpha P_\pi \underline{y}$$

$$\underline{y} = \underline{c}_\pi + \alpha P_\pi \underline{y}$$

$$\underline{y} = (I - \alpha P_\pi)^{-1} \underline{c}_\pi$$

$$= \sum_{j=0}^{\infty} (\alpha P_\pi)^j \underline{c}_\pi$$

Observe $(I - \alpha P_\pi)^{-1} \geq 0$

Optimality Condition

$$y_i = \min_d \left\{ c_{d;i} + \alpha \sum_j P_{d;i,j}\, y_j \right\}$$

Suppose $\nearrow$ holds and $\hat{\pi}$ is any other policy.

$$\hat{y}_i = c_{i,\hat{\pi}(i)} + \alpha \sum_j P_{\hat{\pi}(i);\, i,j}\, \hat{y}_j$$

$$y_i \leq c_{i,\hat{\pi}(i)} + \alpha \sum_j P_{\hat{\pi}(i):\, i,j}\, y_j$$

$$\hat{y}_i = c_{i,\hat{\pi}(i)} + \alpha \sum_j P_{\hat{\pi}(i):\, i,j}\, \hat{y}_j$$

$$y_i \leq c_{i,\hat{\pi}(i)} + \alpha \sum_j P_{\hat{\pi}(i):\, i,j}\, y_j$$

$$\underbrace{\hat{y}_i - y_i}_{\xi_i} \geq \alpha \sum_j P_{\hat{\pi}(i):\, i,j} \left(\hat{y}_j - y_j\right)$$

$$\underline{\xi} \geq \alpha P_{\hat{\pi}}\, \underline{\xi}$$

or $$\left(I - \alpha P_{\hat{\pi}}\right) \underline{\xi} \geq 0$$

$$\left(I - \alpha P_{\hat{\pi}}\right) \xi \geq 0$$

Since $\left(I - \alpha P_{\hat{\pi}}\right)^{-1} \geq 0$

we have

$$\left(I - \alpha P_{\hat{\pi}}\right)^{-1} \left(I - \alpha P_{\hat{\pi}}\right) \xi \geq 0$$

$$\xi \geq 0$$