Stochastic shortest path problems

Dynamics: • $x_{k+1} = \omega_k$ $\quad x_k \in \{1, \dots n, t\}, u_k \in U(i)$

finite sets

Assumption: reach t eventually

• $P(\omega_k = j \mid x_k = i, u_k = u) = p_{ij}(u)$

with $p_{tt} = 1 \quad \forall u \in U(t)$

Cost: For $\pi = \{\mu_0, \mu_1, \dots\}$,

$$J_\pi(i) = \lim_{N \to \infty} E\left(\sum_{k=0}^{N-1} g(x_k, \mu_k(x_k)) \mid x_0 = i\right), i \in S$$

and $g(t, u) = 0 \ \forall u \in U(t)$

$\rightarrow$ optimal cost: $J^*(i) = \min_\pi J_\pi(i)$

(Note: $J_\pi(t) = 0 \ \forall \pi \quad \therefore J^*(t) = 0$)

Result:

A) Given any initial condition $J_0(1), \dots J_0(n)$, the sequence

$$J_{k+1}(i) = \min_{u \in U(i)} \left(g(i,u) + \sum_{j=1}^{n} p_{ij}(u) J_k(i)\right) \quad i \in \{1, \dots n\}$$

converges to optimal $J^*(i)$.

(Note: shortcut, we can include t, $J_0(t) = 0$)

B) $$J^*(i) = \min_{u \in U(i)} \left(g(i,u) + \sum_{j=1}^{n} p_{ij}(u) J^*(j)\right) \quad i \in \{1, \dots n\}$$

$\rightarrow$ stationary optimal policy $\pi = \{\mu, \mu, \dots\}$

Proof:

- cost is bounded

$$|J_\pi(i)| \leq \frac{M}{1-\rho}$$ with $\rho < 1$ probability of NOT reaching t eventually

and M is a constant

A1) bounded tail cost ($\forall$ policies)

$$\left| \lim_{N \to \infty} E\left(\sum_{k=mK}^{N-1} \left(g(x_k, \mu_k(x_k)) \right) \right) \right| \leq \frac{\rho^K \cdot M}{1-\rho}$$

A2) bounded contribution of initial conditions ($\forall$ policies)

$$|E(J_0(x_{mK}))| \leq \rho^K \cdot \max_i |J_0(i)|$$

expected value of terminal cost under π after mK stages

A3) "sandwich"

A4) take limits as $K \to \infty$

$\to$ finally prove B)