B9120-001 Dynamic Programming | Winter 2020

Prof. Daniel Russo

Homework Assignment 6: Due Monday March 23

Approximate Fixed Point Operators.

Let $\mathcal{J} \subset \mathbb{R}^n$ and let $T : \mathcal{J} \to \mathcal{J}$ be a contraction operator with modulus α is some norm $\|\cdot\|$. That is, $\|TJ - TJ'\| \leq \alpha\|J - J'\|$ for all $J, J' \in \mathcal{J}$. Let J^* denote its unique fixed point. Let $\hat{T} : \mathcal{J} \to \mathcal{J}$ represent an approximation to T. Define the induced norm

$$\|\hat{T} - T\| = \sup_{J \in \mathcal{J}} \|\hat{T}J - TJ\|.$$

Part (a) Suppose $\hat{J} = \hat{T}\hat{J}$ is a fixed point of $\hat{T}$. Show

$$\|\hat{J} - J^*\| \leq \frac{\|\hat{T} - T\|}{1 - \alpha}$$

Part (b) No longer assume $\hat{T}$ has a fixed point. Consider the iteration $\hat{J}_{k+1} = \hat{T}\hat{J}_k$ for $k = 0, 1, 2, \cdots$. Show

$$\limsup_{k \to \infty} \|\hat{J}_k - J^*\| \leq \frac{\|\hat{T} - T\|}{1 - \alpha}$$

Remark: *For concreteness, one can have in mind for this problem that* $\mathcal{J} = \{J \in \mathbb{R}^n : \|J\|_\infty \leq \frac{M}{1-\alpha}\}$ *denotes a bounded subset of cost-to-go functions,* T *is the Bellman optimality operator and* $\hat{T}$ *is an approximation to* T *that is consistent with the apriori bounds on* J_μ*. For example,* $\hat{T}J(i) = \min_u \hat{g}(i,u) + \sum_j \hat{P}_{ij}(u)J(j)$ *could be a Bellman operator for an MDP whose transition probabilities were estimated from data.*

Asynchronous value iteration

Consider the iteration

$$J_{k+1}(x) \leftarrow \begin{cases} TJ_k(x) & \text{if } x = x_k \\ J_k(x) & \text{if } x \neq x_k \end{cases}$$

for some sequence of states $(x_i : i = 0, 1, 2, \cdots)$. Suppose the state space is finite and each state is updated infinitely often. Show that $\|J_k - J^*\|_\infty \to 0$.

Hint: We know that

$$J^* - \|J_0 - J^*\|_\infty e \preceq J_0 \preceq J^* - \|J_0 - J^*\|_\infty e,$$

where e denotes a vector of all ones. Show that if a state x has been updated *at least once* by time k, then

$$J^*(x) - \alpha\|J^* - J_0\|_\infty \leq J_k(x) \leq J^*(x) + \alpha\|J_0 - J^*\|_\infty.$$