**MS-E2191 Graduate Seminar on Operations Research**

Fall 2020

Homework 11

Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Please fill in the missing parts of the proof below.

**Proposition:** **Convergence of the DP Algorithm.** For any bounded function $J:X\rightarrow R$, we have

$J^{\*}\left(x\right)=\lim\_{N\to \infty }(T^{N}J)(x),  ∀x\in X$.

**Proof:** For every positive integer $N$, initial state $x\_{0}\in X$, and optimal policy $π=\{μ\_{0},μ\_{1},…\}$, we break the cost into portions acquired from the first $N$ stages and from the remaining stages

$$J\_{π}\left(x\_{0}\right)=\lim\_{K\to \infty }E\{\sum\_{k=0}^{K}α^{k}g(x\_{k},μ\_{k}\left(x\_{k}\right),w\_{k})\} $$

$=\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_$ + $\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_$.

By assuming that the cost is bounded, $\left|g\left(x\_{k},μ\_{k}\left(x\_{k}\right),w\_{k}\right)\right|\leq M$, we obtain

$|\lim\_{K\to \infty }E\{\sum\_{k=N}^{K}α^{k}g\left(x\_{k},μ\_{k}\left(x\_{k}\right),w\_{k}\right)|\leq $ $\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_$ = $\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_$.

Using the relations above it follows that

$J\_{π}\left(x\_{0}\right)-\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_$ - $α^{N}\max\_{x\in X}|J\left(x\right)|$

$$\leq E\{α^{N}J\left(x\_{N}\right)+\sum\_{k=0}^{N-1}\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\}$$

$\leq J\_{π}\left(x\_{0}\right)+\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_$ $+ α^{N}\max\_{x\in X}|J\left(x\right)|$.

By taking the minimum over $π$, we obtain for all $x\_{0}$ and $N$

$$J^{\*}\left(x\_{0}\right)-\frac{α^{N}M}{1-α}-α^{N}\max\_{x\in X}|J\left(x\right)|$$

$$\leq (T^{N}J)(x\_{0})$$

$\leq \\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_\\_$.

By taking the limit $N\rightarrow \infty $, the result follows. **Q.E.D.**