

Mathematics of Reinforcement Learning

Exercise Class 7

Exercise 1 (Existence of a fixed point). Let $\gamma < 1$ and define an operator T acting on functions $w: S \rightarrow \mathbb{R}$ by

$$T[w](s) := \max_{a \in A(s)} \left[r(s, a) + \gamma \sum_{s' \in S} p(s' | s, a) w(s') \right], \quad \text{for } s \in S.$$

Show that T has a unique fixed point $W: S \rightarrow \mathbb{R}$, that is, $W(s) = T[W](s)$ for all $s \in S$.

Hint: Use Banach's fixed point theorem with the supremum norm on the space of functions $w: S \rightarrow \mathbb{R}$. ◇

Exercise 2 (Programming task).

In the Frozen Lake example (see Exercise Class 4), calculate the value function using the fixed point approach. Then, use the value function to compute an optimal strategy.

A corresponding Jupyter notebook file can be found on Moodle. ◇