Problem Set no. 2

Given: November 29, 2020
Due: December 18, 2020

Exercise 2.1 (a) Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the *value iteration operator* of a discounted TBSG on n states, i.e., $T(\mathbf{y})_i = \min_{a\in A_i} c_a + \lambda \sum_{j\in S} p_{a,j} y_j$, if $i \in S_0$, and $T(\mathbf{y})_i = \max_{a\in A_i} c_a + \lambda \sum_{j\in S} p_{a,j} y_j$, if $i \in S_1$, for $\mathbf{y} = (y_i) \in \mathbb{R}^n$, where $0 < \lambda < 1$ is the *discount factor*. Prove that T is a *contraction*, i.e., that $||T(\mathbf{x}) - T(\mathbf{y})||_\infty \le \lambda ||\mathbf{x} - \mathbf{y}||_\infty$ for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

(b) Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the *value iteration operator* of a *stopping* TBSG on n states, i.e., $T(\mathbf{y})_i = \min_{a\in A_i} c_a + \sum_{j\in S} p_{a,j} y_j$, if $i \in S_0$, and $T(\mathbf{y})_i = \max_{a\in A_i} c_a + \sum_{j\in S} p_{a,j} y_j$, if $i \in S_1$, for $\mathbf{y} = (y_i) \in \mathbb{R}^n$. Prove that $T^{(n)}$, i.e., T iterated n times, is a *contraction*, i.e., there exists $0 < \lambda < 1$ such that $||T^{(n)}(\mathbf{x}) - T^{(n)}(\mathbf{y})||_\infty \le \lambda ||\mathbf{x} - \mathbf{y}||_\infty$ for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Exercise 2.2 Construct a stopping TBSG on which there is a sequence of alternating improving switches by both players that cycles.

Exercise 2.3 Prove Lemma 3 in slide 28 of lecture 3: Let $\pi_0 = (\sigma_0, \tau_0)$, where τ_0 is a best response to σ_0. Let π_1 be a profile obtained by one iteration of Howards algorithm. Then $\mathbf{y}^{\pi_1} \le T(\mathbf{y}^{\pi_0})$.

Exercise 2.4 Give a full proof of the theorem on slide 34 of lecture 3: Let Γ be a TBSG. Then, there exists positional strategies σ^* and τ^* of the two players, and $\lambda^* = \lambda^*_\Gamma$, such that σ^* and τ^* are optimal for every discount factor $\lambda^* \le \lambda < 1$.

Exercise 2.5 Show that the value of a MPG played on a graph $G = (V, E, c)$, where $V = V_0 \cup V_1$ and $c : E \to \mathbb{R}$, that starts at $s \in V$ is at *most* μ if and only if there exists a subset $U \subseteq V$ such that $s \in U$ and a potential function $h : U \to \mathbb{R}$ such that (1) If $u \in U \cap V_0$, then there exists an edge $(u, v) \in E$ such that $v \in U$ and $h(u) + c(u, v) \le h(v) + \mu$. (2) If $u \in U \cap V_1$, then for every edge $(u, v) \in E$ we have $v \in U$ and $h(u) + c(u, v) \le h(v) + \mu$.

Exercise 2.6 Let $G = (V, E, c)$, where $V = V_0 \cup V_1$ and $c : E \to [-W, W]$ be an MPG. Let $\mathit{Val}(u)$ be the value of the MPG that starts at u. Let $\mathit{Val}_k(u)$ be the value of the game that starts at u and goes on for exactly k steps. (The game always goes on for k steps, even if a cycle is formed.) Show that for every $u \in V$, we have $\mathit{Val}(u) - \frac{2nW}{k} \le \frac{\mathit{Val}_k(u)}{k} \le \mathit{Val}(u) + \frac{2nW}{k}$. (Hint: show that using an optimal positional strategy for player 0 in the infinite game insures that $\mathit{Val}_k(u) \le (k - n)\mathit{Val}(u) + nW$, where $n = |V|$. Recall that such an optimal positional strategy for player 0 ensures that every cycle formed has an average cost of at most $\mathit{Val}(u)$.)

Exercise 2.7 Use the result of the previous exercise to show that value iteration can be used to obtain an $O(mn^3W)$ to obtain the values of all vertices in a MPG $G = (V, E, c)$ where $c : E \to \{-W, \ldots, 0, \ldots, 1\}$. (I.e., all edge costs are integers of absolute value at most W.) (Hint: the value of each vertex is a rational number with denominator at most n.)