ECEN 743: Reinforcement Learning
Assignment 3

---

1. (2 points) Let $x, y \in \mathbb{R}^n$. The triangle inequality states that $\|x+y\| \leq \|x\| + \|y\|$. Use this to show that $\|x-y\| \geq \|x\| - \|y\|$.

2. (2 points) Consider an MDP with discount factor $\gamma \in (0,1)$. Show that

$$\sup_{\pi} \|V_\pi\|_\infty \leq \frac{\max_{s,a} |r(s,a)|}{(1-\gamma)}.$$

3. (2 points) Show that the Bellman operator T is a monotone operator, i.e, for any $V_1, V_2 \in \mathbb{R}^{|\mathcal{S}|}$ with $V_1 \geq V_2$ (elementwise), $TV_1 \geq TV_2$.

4. (3 points) Consider the function $f : \mathbb{R}^n \to \mathbb{R}^n, f(u) = Au$, where $A \in \mathbb{R}^n \times \mathbb{R}^n$. Assume that the row sums of A is strictly less than 1, i.e., $\sum_j |a_{ij}| \leq \alpha < 1$. Show that $f(\cdot)$ is a contraction mapping with respect to $\|\cdot\|_\infty$.

5. (4 points) Let $\mathcal{U}$ be a given set, and $g_1 : \mathcal{U} \to \mathbb{R}$ and $g_2 : \mathcal{U} \to \mathbb{R}$ be two real-valued functions on $\mathcal{U}$. Also assume that both functions are bounded. Show that

$$|\max_u g_1(u) - \max_u g_2(u)| \leq \max_u |g_1(u) - g_2(u)|$$

6. (7 points) Consider the value iteration algorithm $V_{k+1} = TV_k$, with an arbitrary V_0, where T is the Bellman operator.

   (a) Show that, for $n > m$

$$\|V_m - V_n\|_\infty \leq \frac{\gamma^m}{(1-\gamma)} \|V_0 - V_1\|_\infty.$$

   (b) Let V^* be the optimal value function. Show that

$$\|V_m - V^*\|_\infty \leq \frac{\gamma^m}{(1-\gamma)} \|V_0 - V_1\|_\infty.$$

   (c) Show that

$$\|V_m - V^*\|_\infty \leq \frac{\gamma}{(1-\gamma)} \|V_{m-1} - V_m\|_\infty.$$

7. (5 points) Let $\bar{Q}$ be such that $\|\bar{Q} - Q^*\|_\infty \leq \epsilon$, where Q^* is the optimal Q-value function. Let $\bar{\pi}$ be the greedy policy with respect to $\bar{Q}$, i.e., $\bar{\pi}(s) = \arg\max_a \bar{Q}(s,a)$. Show that

$$\|V^* - V_{\bar{\pi}}\|_\infty \leq \frac{2\epsilon}{(1-\gamma)}$$